UNIVERSITA' DEGLI STUDI DI VERONA

DEPARTMENT OF

Foreign Languages and Literatures

DOCTORAL PROGRAM IN

Linguistics

In collaboration with the

Freie Universität Bozen

CYCLE 37 /YEAR 2021-2022

THESIS TITLE

Typologically-driven parallels between nominal modifiers and arguments in the clause.

A formal account

S.S.D. GLOT-01/A

Coordinator: Prof. Stefan Rabanus

Stefan Rabanus

Tutor: Prof. Andrea Padovan

Andrea Padovan

Doctoral Candidate: Emanuele Bernardi

Emanuele Bernardi

This work is licensed under a Creative Commons CC BY-NC-ND 4.0 DEED
Attribution – NonCommercial - NoDerivatives 4.0 International License, Italy.
To read a copy of the license, visit the web page:
https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en

Attribution - You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial - You may not use the material for commercial purposes.

No derivatives - If you remix, transform, or build upon the material, you may not distribute the modified material.

Typologically-driven parallels between nominal modifiers and arguments in the clause. A formal account – Emanuele Bernardi
PhD thesis
Verona, 15/10/2025

Abstract

Greenberg's (1966) Universal 20 is a widely discussed typological statement that is concerned with the crosslinguistic syntactic variation of demonstratives, cardinal numerals and adjectives modifying the noun. The exact formulation of this universal hinges on the collection of data from the widest range of natural languages possible. The theoretical explanation of the observed patterns is guided instead by the presence of two extremes. First, there are many unattested permutations. Second, a fistful of highly frequent orders manage to cover a large part of the observed crosslinguistic variation. Moreover, the theoretical debate on Greenberg's U20 turns out to be related to the variability among languages with respect to the order of the verb and its arguments in mono-transitive clauses. In fact, the existence and the nature of correlations between the nominal and the clausal domains are relevant to both syntacticians and typologists who argue that natural languages are ruled by some universal principles. In particular, the generative tradition addresses the relation between the nominal and the clausal domains as a debate on the NP- vs. DP-hypothesis (Blümel & Holler 2021). The present work will however follow Corver (2013) in arguing that such a debate can be better termed as an argument on the Lexical vs. Functional Head Hypothesis for the nominal domain (LHH vs. FHH).

The present dissertation tackles Greenberg's U20 and the potential parallels with the clausal domain by crossreferencing data that are publicly available. The generalizations that can be subsequently observed suggest that many theoretical assumptions previously made in the literature must be recast. Although the nominal and the clausal domains appear to affect the word order of one another, they do not show the parallel syntactic behaviors predicted by either the DP-/FH hypothesis or the Consistent Head Serialization stated by Primus (2001: 855). In this respect, the analysis of U20 proposed by Roberts (2017b) has a twofold merit. On the one hand, it provides some insights that set the stage for the data collection presented in this work, while on the other hand it suggests that the correlations between the nominal and

the clausal domains can be captured by means of some macroparameters *à la* M. C. Baker (2008a).

However, neither Roberts (2017b) nor any other syntactic analysis presently available manages to account for all the facts. To provide a better analysis of Greenberg's U20, the present dissertation follows Chiang & Scheffler (2008) in adopting a computationally empowered version of Tree Adjoining Grammars (TAGs). The 2-delayed 2-Multi-Component TAG subsequently proposed manages to overcome the limits that Longenbaugh (2014) and Georgi (2019) point out with respect to the alternate TAG approach formulated by Frank (2002) *et seq*. By virtue of the adopted formalism, the present work not only provides an adequate account of Greenberg's U20 that supports the FHH, but it also solves theoretical issues related to TAGs, like the missing link problem (Kallmeyer & Kuhlmann 2012, Chen-Main & Joshi 2014). By providing many insights on syntactic matters related to both the nominal and the clausal domains, 2-delayed 2-Multi-Component TAGs finally allow to set a rich research agenda for the future.

Acknowledgments

This doctoral thesis is the result of a work process that, in addition to countless instructive experiences and new insights, has also brought an astonishing number of fruitful discussions and valuable encounters. The people with whom I had the opportunity to exchange ideas during my doctoral studies and those who supported me have made a significant contribution to this work.

First and foremost, I would like to express my sincere gratitude to my doctoral supervisor Andrea Padovan, who provided me with advice and support throughout my studies and constantly supervised my work. The regular meetings with him provided me with helpful input and motivated me to deepen my knowledge and get the job done.

My heartfelt thanks, then, go to Andreas Blümel for his support before, during and after my stay at the Georg-August-Universität in Göttingen. I greatly appreciated his help in arranging my time abroad, his extraordinary hospitality, his passionate interest in my work and his patience in clarifying theoretical aspects that were unclear to me.

I am also deeply grateful to my reviewers Cristina Guardiano and Gaetano Fiorin for their helpful comments and enriching suggestions on my work. The care and enthusiasm with which they prepared their reviews filled me with great joy.

Similarly, I owe special thanks to Diego Gabriel Krivochen for his valuable feedback on several drafts that I wrote. He stands out for always responding with enviable speed and acumen, as well as for laying out several ideas that play a pivotal role in my dissertation.

I also wish to extend my appreciation to the people that reviewed my referenced papers (Bernardi 2022, 2023, 2025) and to the editor that always got in touch with me to coordinate the publishing processes, i.e. Mihaela Tănase-Dogaru. The works just mentioned represented crucial intermediate steps that fostered the content of the present dissertation.

I would also like to hugely thank Laura Narisano and David Adger, who made me aware of the existence of Grambank (Skirgård, Haynie, Blasi, et al. 2023, Skirgård, Haynie, Ham-

marström, et al. 2023) and of Georgi (2019), respectively. I do struggle to imagine what appendix A and chapters 4 and 5 would have looked like, had I not dedicated a lot of time thinking of and working on those two references. Similar thanks then go to Guglielmo Cinque, who shared with me Cinque (2023) before it got published, as well as Cinque (Forthcoming).

I am also grateful to a large number of people that friendly and patiently discussed the content of my studies with me, shared relevant works with me (or made me find them) and gave me valuable help with related matters of the academic life. If I had to split these aspects into separate acknowledgments, the following names would occur more than once: Chiara Melloni, Alessandra Tomaselli, Stefan Rabanus, Birgit Alber, and Katharina Hartmann.

I would also like to thank those whom I regularly met at various seminar cycles. In particular, my gratitude goes to the many people organizing and/or attending the Reading Group at the University of Verona and the Conversazioni Linguistiche at the University of Trento. The seminars and the subsequent get-togethers were an important point of reference during my doctoral studies.

For the extensive comments and/or the meetings oriented at improving the content of this dissertation, my heartfelt thanks go to the following people: Timm Lichte, Giorgio Satta, Ermenegildo Bidese, Patrizia Cordin, Marco Magnani, Roland Hinterhölzl, Roberto Zamparelli, Mark Van de Velde, Gerhard Jäger, Tommaso Mattiuzzi and Romano Madaro.

Similar thanks go to the yet unmentioned professors and researchers who welcomed me and/or helped me with my work during my time in Göttingen: Hedde Zeijlstra, Sascha Alexeyenko, Marco Coniglio, Markus Steinbach, Vassilios Spyropoulos, Michael J.M. Barrie, Benjamin Bruening, Balkız Öztürk, Miriam S. Schmidt, Gautam Ottur.

For the help with the tutorships that I conducted, for the in-person meetings and/or for sharing important materials and presentations, I would also like to thank Valerio Pisaniello, Stella Merlin, Serena Dal Maso, Martina Irsara, Jelena Zivojnovic, Veronica Girolami, Eleonora Morei, Arianna Compostella, Marta Tagliani, Michela Redolfi, Sara Cavaglià.

For the kind e-mails and for the Over-The-Air help more generally, I want to thank Mark deVries, Hans Broekhuis, Karsten Schmidtke-Bode, Simone Mattiola, Irina E. Cristea, Julie Doner, Carolin Reinert, Elli Tourtouri, Hunter P. Youngquist, Ilaria Morgani, Stefano Peric, Letizia Savino, Jonas J.J. Klöker, John T.M. Merrill and Viktoria Apel.

I would also like to express my gratitude to my fellow doctoral students Tamara Bassighini, Elena Ferrato, Giulia D'Agostino, Katharina Knapp, Lorenzo Ferroni and Salvatore Gi-

anninò, both for the exchange and support in numerous academic and bureaucratic matters, and for the carefree hours we spent together, which I always remember fondly.

Special thanks also go to the people that I met at the conferences I joined, as well as to those attending the tutorships that I conducted and the LiveLunches I organized. In particular, I would like to express my gratitude to the following people that I met during the PhD program: Gloria Mambelli, Giorgia Andreolli, Anne Kruijt, Chiara Bernardini, Giada Blasut, Marco Fresolone, Pietro Costa León, Tor H. Solhaug, Qingin Tan, Husnain Raza, Daniel Bejarano Bejarano, Mohsen Shaeyan, Farbod Khouzani, Melanie Hobich, Anke Stakemann, Helmut Weiß. My gratitude also goes to Andrea (and his friends) and Daniel, two wonderful flatmates that I had the great pleasure to meet, as well as to the most exquisite landlady ever, Livia.

A special mention also goes to the professors that helped me when the present research was far from being real. Thus, I would like to thank Federico Gobbo for making me love (the) syntax (of natural language), as well as Cecilia Andorno, Mario Squartini and Cristiano Chesi, who supervised my BA thesis, by MA thesis and my applications for a PhD, respectively. I would also like to express my gratitude to Christopher Owen. I am afraid I never thanked him properly, even though he greatly assisted me with my BA thesis in Turin. I wish to make amends by acknowledging him here.

Last but certainly not least, I would like to express my gratitude to the people who helped and supported me throughout this process with far more patience and hope than I did have.

I would like to thank my friend Mattia and my friends from Caraglio and its city belt: Giacomo, Stefano, Gabriele, Marika, Marta, Federico, Fabio, Erik. I would also like to thank several people I am no longer in contact with, as they helped me choose to embark on this journey.

Special thanks from the bottom of my heart then go to my parents Aldo and Susanna, to my sister Sonia, her husband Valerio and my nephew *Frank*. I am also truly grateful to the wind underneath my wings, i.e. Tamara, as well as to her parents and to the relatives and the friends she loves.

My final thoughts go to my grandmother Catterina, who is surely enjoying this achievement, as well as to the other relatives that I carry in my heart.

Contents

1 Introduction **13**

2 Syntactic analyses of typological observations **17**

2.1 The Pāṇinian approach 17

2.2 The widespread focus on head-dependent pairs 20

2.2.1 Dominance and Harmony 20

2.2.2 The SV/VS and OV/VO pairs 24

2.2.3 From the NP to the DP hypothesis 29

2.2.4 On the necessity of functional projections above NP 35

2.3 Antisymmetric approaches 47

2.3.1 Kayne (1994) 48

2.3.2 The Final-Over-Final Condition (FOFC) 54

2.3.3 FOFC as a consequence of antisymmetry 60

2.4 Approaches that reject antisymmetry 63

2.4.1 Symmetric base-generated structures 65

2.4.2 SOV as the base-generated order 69

3 Greenberg's Universal 20 **79**

3.1 Definitions 79

3.1.1 On demonstratives 81

3.1.2 On cardinal numerals 91

3.1.3 On adjectives 100

3.1.4 On nouns 115

3.2 Previous analyses 123

3.2.1 Cinque (2005) 126

3.2.2 Cinque (2023) . . . 131
3.2.3 Roberts (2017b) . . . 137
3.2.4 Abels & Neeleman (2012) and related works . . . 142
3.2.5 Bruening (2018a) and Medeiros (2018) . . . 148
3.2.6 Other relevant works on U20 . . . 153

4 Novel typological observations 162
4.1 Constructing the database . . . 162
4.1.1 Sources and methods: WALS as a starting point . . . 163
4.1.2 Sources and methods: Beyond WALS . . . 165
4.1.3 One first generalization . . . 169
4.1.4 Mono-transitive clauses: back to six types? . . . 170
4.2 A comparison between the present sample and others . . . 171
4.2.1 Glottolog over Ethnologue . . . 171
4.2.2 U20 in Cinque's sample, Dryer's, and in Appendix A . . . 174
4.2.3 Different samples w.r.t. the linear order in the clause . . . 176
4.3 An analysis of the data . . . 178
4.3.1 The cluster in Roberts (2017b), extended . . . 179
4.3.2 Four orders with peculiar properties . . . 180
4.3.3 Two more groupings . . . 183
4.3.4 Some further remarks . . . 185
4.4 Interim conclusions . . . 189
4.5 On the need of a different theoretical approach . . . 191
4.5.1 Beyond a linear statement . . . 192
4.5.2 A parametric restatement . . . 193
4.5.3 Different structures despite the same paremeter setting . . . 199
4.5.4 The conceptual basis for a different solution . . . 204

5 Tree Adjoining Grammars 209
5.1 Setting the stage . . . 209
5.1.1 Context-freeness: Tree Substitution Grammars . . . 212
5.1.2 Mild context-sensitivity: Tree Adjoining Grammars . . . 216
5.1.3 Derivation trees . . . 220

5.1.4 An aside: Two remarks on derivation trees . . . 223
5.1.5 On the interaction between Merge and TAG operations, pt. 1 . . . 228
5.2 L-TAGs and U20 . . . 239
5.2.1 The interface between TAGs and semantics . . . 240
5.2.2 An L-TAG adaptation of Bruening's analysis of U20 . . . 257
5.2.3 An L-TAG adaptation of Abels and Neeleman's analysis of U20 . . . 262
5.3 Empowering the formalism . . . 266
5.3.1 Multi-Component TAGs . . . 268
5.3.2 On the interaction between Merge and TAG operations, pt. 2 . . . 277
5.3.3 Delayed tree-locality . . . 291
5.4 Fixing the problems of L-TAGs . . . 301
5.4.1 Reversing the perspective on adjoining . . . 302
5.4.2 A rejoinder to Georgi (2019) . . . 309
5.4.3 On the interaction between Merge and TAG operations, pt. 3 . . . 324
5.4.4 An alternative based on features and syncategorematicity . . . 340
5.4.5 Some final remarks on the clausal domain . . . 363
5.5 A consistent (MC-)TAG account of U20 . . . 369
5.5.1 A 2-delayed 2-MC-TAG adaptation of pied-piping for U20 . . . 369
5.5.2 On the elementary tree that each modifier of U20 anchors . . . 380

6 Conclusions **389**

A Typological data **393**
A.1 All the languages in the sample, listed for family and genus . . . 394
A.2 N-A-Num-Dem . . . 442
A.3 Dem-Num-A-N . . . 452
A.4 Num-N-A-Dem . . . 459
A.5 Dem-N-A-Num . . . 465
A.6 Dem-Num-N-A . . . 470
A.7 N-A-Dem-Num . . . 473
A.8 N-Dem-Num-A . . . 475
A.9 Dem-A-N-Num . . . 477
A.10 N-Num-A-Dem . . . 478

A.11 Num-A-N-Dem 479
A.12 Dem-N-Num-A 481
A.13 N-Dem-A-Num 481
A.14 A-N-Num-Dem 482
A.15 A-N-Dem-Num 483

Bibliography **484**

Glossary

1	first person
2	second person
3	third person
ACC	accusative
ADV	adverbial
AGR	agreement
ART	article
AUX	auxiliary
CLF	sortal classifier
COMP	complementizer
COP	copula
CS	construct state
DAT	dative
DEF	definite
DEM	demonstrative
DET	determiner
DIST	distal
EXCL	exclusive
F	feminine
FUT	future
GEN	genitive
INDF	indefinite
INF	infinitive
INS	instrumental
IPFV	imperfective
M	masculine
MENS	mensural classifier
N	neuter
NEG	negative
NOM	nominative
PFV	perfective
PL	plural
POSS	possessive
POT	potential
PROX	proximal
PRS	present
PST	past
PTCP	participle
Q	question particle
REFL	reflexive
REL	relative
SBJV	subjunctive
SG	singular
STAT	stative

Chapter 1

Introduction

Greenberg's (1966) Universal 20 is a widely discussed typological statement concerned with the crosslinguistic syntactic variation of demonstratives, cardinal numerals and adjectives modifying the noun. Among the various universals proposed in Greenberg (1966–1963), his Universal 20 stands out for two reasons. First, it tries to generalize over the order of four elements belonging to a single domain without resorting to any implicational relation. Second, it soon turned out to exclude some orders that are actually attested. Still, some principle along those lines stands and asks for a reformulation, as some orders remain unattested even if more and more data are collected crosslinguistically (Dryer 2018a,b, Cinque 2021). These facts have fostered subsequent research aimed at assessing the empirical breadth of the attested permutations on the one hand, and at providing on the other hand a potential explanation of the patterns, which are further characterized by the fact that a fistful of highly frequent orders manages to cover a large part of the observed crosslinguistic variation.

An adequate account of U20 is inevitably related to the verification of another hypothesis that is relevant to both typologists and syntacticians, namely that there is some compelling parallel between the clausal and the nominal domains. Given that in the generative tradition it is assumed that syntactic structures are strictly endocentric (see Fukui & Narita 2014, Lichte 2021 for discussion), the relation between the nominal and the clausal domains boils down to a debate on the NP- vs. DP-hypothesis (Blümel & Holler 2021, a.o.). According to the former view, the noun is the head of the phrase in which it is contained, and this principle makes the structure of the nominal domain completely different from the structure that is assumed for the clause. Supporters of the DP-hypothesis argue instead that the two domains closely parallel each other, even though the exact details in this respect vary depending on the

author. The present work will follow Corver (2013) in arguing that such a debate can be better termed as an argument on the Lexical vs. Functional Head Hypothesis for the nominal domain (LHH vs. FHH). In order to remain agnostic with respect to this debated property of the whole nominal construct, Stowell (1989: 232), Bošković (2005) and Giusti (in press) choose to talk about *Common Noun Phrases*, *traditional NPs* and *nominal expressions*, respectively. The expression *nominal domain* will be used in the present work with a similar theory-neutral intent.

As has been stated above, an adequate analysis of U20 can be relevant to both typologists and syntacticians. Chapter 2 will therefore be concerned with a confrontation between the research traditions started by Chomsky (2002–1957, 1965) and Greenberg (1966–1963). The discussion in section 2.1 will first concentrate on how one can benefit from both traditions. Then, section 2.2 will illustrate that the notion of headedness has played a relevant role in shaping both typological and generative analyses. In the latter, headedness has played a fundamental role in defining a shift from the LHH to the FHH in both the nominal and the clausal domains (see sections 2.2.3 and 2.2.4). The typological tradition has often assumed, instead, two opposing structures that are ruled by a principle that Primus (2001: 855) calls the *Consistent Head Serialization*, which is argued to capture the crosslinguistic frequence of certain patterns across both the nominal and the clausal domains (see sections 2.2.1 and 2.2.2).

As section 2.3 will show, the antisymmetric approach (Kayne 1994) proposes to account for the empirical facts of interest to typologists by assuming a single syntactic structure in which the elements can be remerged in highly constrained manners. In particular, sections 2.3.2 and 2.3.3 will linger on the *Final-Over-Final Condition* (Sheehan et al. 2017), which is an offshot of antisymmetry that is argued to effectively combine the insights from both the Chomskyan and the Greenbergian traditions (Biberauer, Holmberg & Roberts 2014: 170). However, the generative literature has also provided a series of competing alternatives that aim to capture the same (or a broader range of) facts by means of different theoretical starting points, which will be discussed in section 2.4.

Chapter 3 will get into the heart of the discussion by focusing on Greenberg's U20. Section 3.1 will deal with the alternate structures that the generative tradition has provided to account for the relevant morphosyntactic properties of demonstratives, cardinal numerals, adjectives and nouns, while section 3.2 will provide a survey of the existing analyses of U20, so as to assess the manners in which the empirical and theoretical content of Greenberg's

original formulation has been updated in the literature. In this respect, three works are particularly relevant. On the one hand, Dryer (2018a) and Cinque (2023) provide two competing analyses of U20 by grounding them on distinct language samples that are collected by resorting to overlapping methodologies (see section 3.2.6), i.e. Dryer (2018b) and Cinque (2021). On the other hand, Roberts (2017b) focuses on two specific permutations of U20 and tries to explain their formal properties by sketching yet another parallel with the structure of the clause. Instead of focusing on the high occurrence of verb-initial languages in such a subset of U20, the present work will place emphasis on the surprising rarity of verb-final languages therein, a fact which is overlooked in Roberts (2017b).

The content of chapter 4 extends and refines Bernardi (2025). It is also closely related to what is presented in appendix A, namely a convenience sample (Körtvélyessy 2017: 127) aimed at collecting languages that have been classified in the literature with respect to both the basic order concerning Greenberg's U20 (as per Dryer 2018b, Cinque 2021) and the unmarked order of arguments in mono-transitive clauses. Section 4.1 will present how the data have been collected and classified, while section 4.2 will illustrate that the results are in agreement with samples of comparable sizes. The core idea motivating this effort is that Roberts (2017b) may have provided evidence of a negative correlation between certain permutations of U20 and verb-finality in the clause, and the analysis of appendix A which is done in section 4.3 will assess whether this observation still holds in a larger sample and if comparable points can be raised with respect to the other attested permutations of U20. The four major observations that will consequently emerge will be summarized in a rather pre-theoretic sense in section 4.4, whereas section 4.5 will argue that a parameterization is the better strategy to capture the facts in an adequate, meaningful and a concise way. However, none of the existing analyses of U20 seems well-suited to provide a satisfactory account of the observed correlations, and section 4.5 will argue that this is due to some theoretical premises that need to be dispensed with.

Thus, chapter 5 will assess whether the mentioned limitations can be overcome by adopting an alternate formalism, namely Tree Adjoining Grammars (TAGs), which suffer in turn from two problems. First, many syntacticians are not used to TAGs, and second, such a formalism bears its own theory-internal issues, which have been argued to represent a compelling reason for the rejection of the approach (Georgi 2019). Section 5.1 will introduce the basic properties of the formalism, and section 5.2 will show that those are already suffi-

cient to recast two previous analyses of U20 in a different framework. However, such TAG adaptations do not improve some key aspects of the original proposals, and neither do they manage to overcome the theory-internal difficulties just mentioned.

Section 5.3 will be concerned with the proposals that have been put forth in the literature to fix the problems related to TAGs. These attempts will provide the theoretical toolkit that allows to put forth a novel promising version of TAGs, to be presented in section 5.4. After solving the major problems of TAGs and introducing some further formal refinements, section 5.4 will eventually pave the way for a satisfactory TAG account of Greenberg's U20, to be discussed in section 5.5. As a result, the FHH will be supported on a twofold ground. On the one hand, the LHH fails to predict the generalizations that emerge in chapter 4. On the other hand, the most satisfactory TAG account of U20 that will be provided crucially hinges on the need of some functional projection topping off the structure of the nominal domain.

Chapter 2

Syntactic analyses of typological observations

The present chapter will be concerned with a confrontation between the research traditions started by Chomsky (2002–1957, 1965) and Greenberg (1966–1963), as they show similarities and differences. The discussion in section 2.1 will first concentrate on how one can benefit from both traditions, then section 2.2 will illustrate that the notion of headedness has played a relevant role in shaping typological as well as generative analyses of empirical facts. The discussion in section 2.3 will focus on the antisymmetric approach, a line of research whose theoretical premises are argued to be capable of capturing many relevant typological facts. A series of competing alternatives argue that similar results can be obtained by means of different theoretical starting points, which will be discussed in section 2.4.

2.1 The Pāṇinian approach

Both Wiltschko (2014: 300-301) and Pleshak & Polinsky (Forthcoming: Sec. 2.2) report the popular idealization according to which formal syntacticians focus on an in-depth investigation of a single language, and draw generalizations based on the related analysis. However, the interest among generative syntacticians with respect to crosslinguistic variation can be traced back to the Principles and Parameters approach (Chomsky 1981), and "it has come to a degree of convergence of interest and method with linguistic typology as practiced by linguists like Joseph Greenberg" (M. C. Baker 2012: Sec. 1). The points of contact between the research traditions started by Chomsky (2002–1957, 1965) and Greenberg (1966–1963)

are motivated by the fact that they both aim at "establishing general-universal truths [...] valid in all past, present and future languages" (Alfieri, Arcodia & Ramat 2021: 23).[1] Moreover, both generative grammar and Greenberg's typology emerged in response to scholars who were claiming that languages vary arbitrarily, in a manner that is not constrained by any universal principle (see Kashyap 2019: 776 and references therein).

Apparently, history repeats itself, as nowadays any view that sees crosslinguistic comparison as a worthwhile effort must necessarily take distance from Evans & Levinson (2009). The latter argue that no valid language universal can be discovered, because there exists none. Wiltschko (2014: Sec. 1.3-1.4, 6.3, 7.3) calls this stance the *No base hypothesis*, and its major problem is twofold. On the one hand, it predicts a way greater degree of crosslinguistic variability than it is actually found, while on the other hand it gives up the very premises of any confrontation among two or more languages (Alfieri, Arcodia & Ramat 2021, Wiltschko 2021b).

At the same time, anyone who follows Graffi (1980), Cinque (2013), Biberauer, Holmberg & Roberts (2014) and related works in arguing that formal analyses can benefit from typological observations must also take distance from Newmeyer (2010: 217). In fact, the latter argues that "typological evidence is not a reasonable heuristic for grammar construction, either directly or indirectly". Contrary to what Newmeyer (2010: 217) claims, there is a growing body of works supporting the idea that a set of principles or parameters internal to grammatical theory can derive typological generalizations, as sections 2.3 and 2.4 will discuss. Moreover, the empirical observations that will be presented in chapter 4 with respect to Greenberg's Universal 20 will be formulated in section 4.5.2 and section 5.5.1 along the macroparametric lines advocated by M. C. Baker (2008a: Sec. 1).

However, the level of theoretical abstraction and the types of explanations provided to account for similarities and differences among languages represent a point of attrition between the generative and the typological approaches. As Pleshak & Polinsky (Forthcoming) observe, "[n]owhere do typology and formal comparative syntax differ more than in the nature of explanations for linguistic phenomena". Instead of being discouraged by the divergences that characterize the extreme takes within the two traditions, one can however benefit from what is observed by Kashyap (2019: 772):

> [L]anguage typology is not a linguistic theory or framework; rather, it is a method

[1] The claim quoted from Alfieri, Arcodia & Ramat (2021: 23) is addressed as *uniformitarianism* in works like Croft (2003: 49-50), Roberts (2017c) and Walkden (2019).

> of linguistic analysis, which necessarily involves cross-linguistic comparison of linguistic items, meanings, and related phenomena. Typological investigations can be carried out within any theoretical framework [...].

What Kashyap (2019) describes is perfectly in tune not only with proposals that pursue their own ways to reconcile typology with a formal approach (see e.g. M. C. Baker 2012), but also with any work that tries to integrate empirical generalizations made available by typologists so as to support or falsify a theoretical analysis (see Greenberg 1990: 703, 706, Park, Bahrudin & Han 2020 and Blümel & Holler 2021: 8 for related discussion). In fact, the number of works in the generative tradition that try to back up their theoretical claims by means of a typologically diversified number of languages is steadily increasing. One can mention as a first set of examples those works that (in spite of the issues outlined by Gianollo, Guardiano & Longobardi 2008: 113) use WALS (Dryer & Haspelmath 2013) or related typological sources as a first exploratory device to support their own theoretical claims, namely Julien (2002), Cinque (2005), Cecchetto (2013), Richards (2016), Sheehan et al. (2017), and Mahowald, Jurafsky & Norris (2021), among others.

In addition to those, one can mention works that create typological samples of their own. A first way to do this is by having the latter balanced along the lines of Rijkhoff & Bakker (1998) and related works. This is for instance the case of Tvica (2017), whose focus is the Rich Agreement Hypothesis (to be further discussed in section 4.5.3). Alternatively, one can try to provide a convenience sample of the phenomenon at hand by collecting data from as many languages as possible in which it is attested (see Körtvélyessy 2017: 127). Examples in this sense are Cinque (2021, 2023), which will play a central role in section 3.2.2 and chapter 4, Höhn's (2017) analysis of adnominal pronoun constructions (see section 3.1.1), and the account for Anti-Agreement provided by Baier (2018).

The majority of the formal approaches that have been mentioned so far represent instances of the *the Pāṇinian approach*, whose methodology is presented by Bickel (2015: Sec. 1) as follows:

> [S]eek maximally general statements, try to explain away as many counterexamples as possible by reanalyzing them as falling outside the scope of the statement, and then account for the remaining exceptions by specifying their conditions in terms that are again as general as possible. [... A]s Chomsky (1965) noted long ago, the only possible justification consists in demonstrating that the generaliza-

> tion can be formally derived from (i.e. proven in) a restrictive theory of possible grammars. This has a crucial implication: for any such derivation to be possible, the vocabulary of the generalizations must be identical to the vocabulary of the theory.

The present work will try to follow the crucial aspects just quoted, while trying, at the same time, to avoid falling into the problem pointed out by Leivada (2020: 368), i.e. adding so many stipulations that some universal principle is eventually captured, though in a manner that is by no means insightful.

2.2 The widespread focus on head-dependent pairs

The present section will illustrate that both the typological and the generative traditions have been concerned with the debate on whether the nominal and the clausal domains parallel each other or not. In section 2.2.1, it will be shown that the typological inquiries have been guided by two terms first proposed in Greenberg (1966–1963), i.e. Dominance and Harmony. In section 2.2.2 a summary will be provided with respect to the ways in which typologists have subsequently tried to account for the empirical facts related to the clausal word order. This will give an opportunity to report some criticisms raised by syntacticians. From section 2.2.3 on, the readers will be introduced into a historical overview of the formal apparatus adopted by generative syntacticians to account for the properties of the clause and of the nominal domain. The way in which the structure of the latter is articulated is a matter of an ongoing debate that section 2.2.4 will try to summarize.

2.2.1 Dominance and Harmony

Greenberg (1966–1963: 97) is probably the best quote to show how the notion of Dominance is applied in typological inquiries:

> [I]n the tetrachoric table resulting from the alternative for each of the combinations there is a single blank. [...] A dominant order may always occur, but its opposite, the recessive, occurs only when a harmonic construction is likewise present. [...] The entry with zero is always the recessive one for each construction, and the two constructions involved are disharmonic with each other.

Greenberg (1966–1963: 97) explicitly focuses on tetrachoric tables because the generalizations with respect to word order that he formulates rely on pairs of linguistic elements, such as the position of adpositions with respect to the rest of the noun phrase they introduce, the order of a mono-transitive verb and its direct object, the position of a relative clause with respect to the noun it modifies, and so on (see Cinque 2023: Sec. 4.2 for some recent discussion). Greenberg (1966–1963) shows that several facts of word order can be seen as an outcome of a material implication that involves two logically independent pairs (Croft 2003: 54), so that several works have broadened this methodology since (see a.o. Hawkins 1983, Dryer 1997, 2013a). A concrete example in this sense is the relation between the order of Object and Verb and the one of Noun and Relative Clause. The interaction of these two typological dimensions gives rise to four possibilities illustrated in table 2.1.

	N-Rel	**Rel-N**
OV	OV & N-Rel	OV & Rel-N
VO	VO & N-Rel	~~VO & Rel-N~~

Table 2.1: Tetrachoric table inferred from Dryer (2013e).

As Dryer (2013e) shows, the concurrence of VO and Rel-N is extremely rare crosslinguistically, being attested in only 5 languages out of 666 unambiguous cases. The presence of some exception does not represent a problem for Greenberg (1966–1963), who claims that typologically insightful correlations can be exceptionless universals as well as statistical (see also Bickel 2015: 903). Therefore, the interaction of the OV/VO and N-Rel/Rel-N pairs can be seen as three cases where at least one dominant (i.e. logically possible) construction is present, be it OV or N-Rel. Croft (2003: 60) further clarifies that the tetrachoric in table 2.1 can be captured by two logically equivalent material implications, the former having the dominant construction OV as its implicatum (1a) and the latter being its contraposition (1c) where the implicatum is N-Rel.

(1) a. Rel-N → OV ≡
b. ¬ OV → ¬ Rel-N ≡ [contrapositive]
c. VO → N-Rel

The "logical factor of a zero in [a] tetrachoric table" mentioned in the quotation above from Greenberg (1966–1963: 97) is represented by the lower right case in table 2.1, because

VO and Rel-N are negatively correlated crosslinguistically. This factor, which both (1a) and (1c) manage to capture, makes them "disharmonic with each other" as per Greenberg (1966–1963: 97) and also determines that VO and Rel-N are recessive with respect to OV and N-Rel, respectively. Therefore, the frequencies observable from Dryer (2013e), reported in table 2.2 on this page with respect to table 2.1, can also be captured by the concept of Harmony.

	N-Rel	**Rel-N**
OV	113	132
VO	416	5

Table 2.2: Frequency table referring to table 2.1 and Dryer (2013e).

On the one hand, the OV order is harmonic given that it can co-occur with the recessive order Rel-N, as in the upper right case in tables 2.1 and 2.2. On the other hand, the harmonic nature of N-Rel allows it to co-occur with the recessive VO order in a single language, which corresponds to the the lower left case in the two tables. It is also worth noting that the latter option is more frequent than the combination of the two dominant values, shown on the upper left in both tables 2.1 and 2.2. Such a case "in which one of the recessive alternatives is more frequent than the dominant" are predicted by Greenberg (1966–1963: 97) himself, who thus claims that "the notion of dominance is not based on its more frequent occurrence", *contra* (Croft 2003: 60-61).

Although not explicitly defined in Greenberg's original formulation (see Croft 2003: 62 fn. 1 for a critical assessment), the notion of Harmony was later rephrased in structural terms by Lehmann (1973, 1978) and Vennemann (1974, 1977), among others, with a twofold goal. The general proposal of these works is to claim that the upper left case in table 2.1 is an exceptional behavior, because the other two attested cases in tables 2.1 and 2.2 represent subcases for two opposing crosslinguistic trends, whereby the linear order of Verb and Object correlates with the one of several other pairs, regardless of the categories involved (Roberts 2017a and Moro & Roberts 2024: 617 fn. 12). Although each cited work proposes a different formulation to explain the harmonic patterns (see Dryer 1992: Sec. 1.2 for discussion), the core idea can be captured by the *Consistent Head Serialization* defined by Primus (2001: 855):

(2) For all phrasal categories X, the head of X either precedes or follows all dependents.

The definition in (2) predicts two types of languages, namely consistently head-initial and consistently head-final languages. The problem is, of course, finding criteria to determine what counts as a head or as a dependent (see Svenonius 2024: Sec. 2.1 a.o. for discussion). In Lehmann (1973, 1978) and Vennemann (1974, 1977) verbs and nouns are taken as heads, so that VO languages are also characterized by the N-Dem, N-Num, N-A orders, while OV languages tend to pattern with the Dem-N, Num-N, A-N (Primus 2001: 856).

Structural definitions of typological facts represent a natural bridging point with syntactic works, although the theoretical premises and the subsequent descriptive adequacy can still vary a lot. The *Branching Direction Theory* proposed by Dryer (1992, 2009) recognizes these problems and tries to refine the principle in (2) by resonating with the binary-branching structures posited since Kayne (1984), which will be discussed in section 2.2.3 (see also Newmeyer 1998: 344).

In particular, Dryer (1992, 2009) claims that, although the two ideal tendencies oulined by (2) have some merit, the different properties among the dependents affect their compliance with such a principle, so that Harmony does not end up providing the expected striking results (see also Primus 2001: 858). An example in this sense is represented by Dryer (2013b), who disconfirms an often-assumed correlation between the order of Adjective and Noun with respect to the one of Object and Verb. The data collected by Dryer (2013b) show that there are at least one hundred languages per each possible case, contrary to what has been seen above for Dryer (2013e).

The absence of a single zero in such a table prevents from establishing clear disharmonic relations between supposedly recessive constructions, so that even the concept of Dominance seems of little use in this case. A more recent and general claim in this direction comes from Cinque (2023: 70-71), a syntactic proposal (to be discussed in detail in section 3.2.2) where it is argued that modifiers are less compliant than complements with respect to a harmonic patterning with their selecting head, an aspect which seems to be further weakened across domains.

The reaction to these issues has always been twofold. On the one hand, scholars have tried to look at even more pairs (as can be seen on Dryer & Haspelmath 2013 and Skirgård, Haynie, Blasi, et al. 2023, Skirgård, Haynie, Hammarström, et al. 2023), while on the other hand several theories have been formulated so as to explain why Dominance and Harmony happen to work in certain cases but not in others. In particular, some attempts have tried to reach a

better explanatory adequacy by relying on processing constraints or on syntactic structures, respectively exemplified by Hawkins (2013) and Cinque (2023: Sec. 2). One line of research (see a.o. Sheehan et al. 2017 and Cinque 2023: Sec. 5.1-5.3) among the proposals of the latter kind further claims that disharmonic patterns can be explained if recessive constructions are structurally defined in ways that will be discussed in sections 2.3.2 and 2.3.3. These proposals raise in turn new questions and new problems that will be surveyed in section 2.4.

What is crucial about the two kinds of approaches that have been outlined in the previous indent is that they both rely on some specific "theory of grammar" (Greenberg 1990: 703, 706) where typological patterns are either assumed, or expected, to be ultimately made of binary relations. Although legitimate and well grounded, such a theoretical basis can be harmful (see Cinque 2013: 52 for a remark going in this direction), because it can prevent observing certain interesting correlations that will be introduced in section 3.2.3 and further discussed in chapter 4.

2.2.2 The SV/VS and OV/VO pairs

Several categorizations have been proposed by typologists with respect to the order of arguments in mono-transitive clauses. The most known is given by Greenberg (1966–1963), by which the linear order of mono-transitive clauses is divided among the six conceivable types obtainable as soon as one considers S(ubject), O(bject), and V(erb), namely SOV, OSV, SVO, OVS, VSO, and VOS.

However, since some of Greenberg's correlations depended on the contrast between OV and VO languages, it was subsequently debated whether these six types could be reduced to a two-way split exclusively determined by this difference. In this regard, Dryer (1991: 443-451) reports SVO languages whose typological properties are not consistently opposed to OV languages.

The subsequent reflection on Harmony and branch directionality leads Dryer (2013a) to argue that a more insightful typology could be stated by proposing a contingency table determined by two factors, namely the relations of linear precedence between S and V on the one hand and between O and V on the other. Starting from these premises, overlapping properties between SOV and OSV can be collapsed together (an aspect which also finds theoretical support in Haider 2000, 2013) and opposed to those between VSO and VOS, by noting that in the two pairs the verb either follows or precedes its arguments, which were captured respectively

by the SV & OV pairs on the one hand and the VS & VO pairs on the other.

The only distinction that Dryer's four-order typology neatly shares with the six-order one is between SVO (stated as SV & OV) and OVS (stated as OV & VS). Dryer (2013a: 271, 284-286) claims that this is a good move, given that the rare OVS languages attested cannot be legitimately collapsed together with the SVO ones because of the typological facts he discusses. This four-way typology has the further advantage of capturing several properties that VSO and VOS languages share.

By quoting the summary of Dryer (1997) made by Newmeyer (2005: 139), Dryer (2013a: 269) illustrates that languages can be "indeterminately VSO or VOS", that "VSO and VOS languages are now predicted (correctly) to have the same typological properties" and that "VSO and VOS languages readily change historically from one order to the other". However, treating the linear order of subjects of mono-transitive verbs as a separate typological factor may induce one to hypothesize that this observation should not vary as soon as subjects of intransitive verbs are concerned, but Dryer (2013f: Sec. 4) acknowledges that these two cases need not have overlapping linear properties within a single language.

Both Greenberg's and Dryer's typologies expose themselves to at least two kinds of problems related to the way such classifications are formulated. First, both are monocategorizing (Ivani & Bickel 2023), namely they induce the typologist to determine one and only one unmarked order for each relevant factor in a language, leading to a twofold problematic consequence. On the one hand, such approaches have little to say with respect to language-internal exceptions to a given classification (see Kashyap 2019 and Gerdes, Kahane & Chen 2021: Sec. 2). On the other hand, languages with multiple unmarked orders are likely to fall under the larger umbrella of those whose order cannot be determined, no matter if the alternatives are all the six ones posited by Greenberg or just a subset of those.

As far as the latter issue is concerned, Dryer's typology fares better than Greenberg's, because it can capture the crosslinguistic likelihood of VSO/VOS alternations (see S. Chung 2006 and references therein) as well as the OS/SO flexibility in many verb-final languages (Haider 2000, 2013), but at the same time it fosters further discussion of those languages that show alternations between SVO (or more rarely OVS, see Sailor & Mahajan 2012 for theoretical discussion of these rare cases) and one linear order where the verb either consistently precedes or consistently follows its arguments.

The syntactic proposals show mixed opinions with respect to the claim of some core under-

lying commonality shared by all verb-initial languages. On the one hand, Pleshak & Polinsky (Forthcoming: Sec. 5.1.2) highlight that also VSO and VOS languages may differ from one another in several derivational properties that give rise to the observable word order (see also Clemens & Polinsky 2017, Rouveret 2017), which is an aspect that Dryer (2013a) would fail to capture. But on the other hand, the SV/VS vs. OV/VO split proposed by Dryer has nonetheless been adapted by generative syntacticians in a structural sense. Clemens & Polinsky (2017) illustrate that this idea was grounded on a split parameterization of either one of two aspects that section 2.2.3 will introduce, namely the X-bar schema (see also Biberauer, Roberts & Sheehan 2014: 47) and the sets of functional and lexical projections.

The second major problem with Greenberg's and Dryer's typologies is related to the nature of the arguments. Setting aside the fact that even among typologists the consensus on this matter is not as solid as one would expect (see Alfieri, Arcodia & Ramat 2021: Sec. 3), several criticisms of these typologies have been made by syntacticians who pointed out that in many cases the wrong diagnostics with respect to subjects or objects have been made. One first explanatory example in this sense are the Germanic languages, which turned out to be a problem for typologists because of the verb second phenomenon (thoroughly discussed by Holmberg 2015 and Walkden & Booth 2020, a.o.) that characterizes most of them. Catasso (2024: 1) summarizes these properties for German as follows:

> Present-day German is a so-called "asymmetric" Verb-Second ([...] V2) language. Declarative main clauses exhibit obligatory V-to-C movement, and the topological position to the left of the finite verb ([...] Spec,CP) must be occupied by one XP in order to satisfy an EPP(-like) feature on the left-peripheral projection hosting the verb. In embedded clauses introduced by an overt subordinating conjunction, leftward movement into C is formally blocked by the presence of the complementizer in the same head position standardly targeted by the verb [...].

The examples in (3), adapted from Catasso (2024: Sec. 1) (to which the readers are referred for further references and discussion), illustrate what has just been quoted:

(3) a. *Tamara hat das Buch nicht gelesen*
Tamara AUX the.N.ACC book NEG read.PTCP
'Tamara has not read the book.'

b. ... *dass Tamara das Buch nicht gelesen hat*
COMP Tamara the.ACC book NEG read.PTCP AUX
'... that Tamara has not read the book.'

Although this syntactic phenomenon has been neglected for many years by typologists (see Haider 2018: Sec. 1.2 for a historical overview), it was seemingly patched afterwards by saying that languages like German are both SVO and SOV, the former classification being a byproduct of the verb-second phenomenon itself, i.e. a restriction imposed by the syntax of main clauses. In fact, Dryer (2013f: Sec. 3) claims that the subject is the constituent that most often occupies the preverbal position (as opposed to objects and adverbials).

However, S. Müller (2023a: 12-15, Sec. 6.1) convincingly argues against this latter stance, by showing not only on a structural basis, but also in terms of frequency, that the first constituent in a V2 main clause is not necessarily its subject (see Holmberg 2015: Sec. 2 for a list of categories that can be the first constitutent in V2 main clauses). Thus, the SVO characterization of German cannot be upheld, while syntactic and hierarchical considerations (such as the order of arguments in subordinate clauses and some general properties typologically related to the OV order) strongly point towards a verb-final classification of German and several other Germanic languages (see Bayer & Freitag 2020 for further discussion).

OVS languages represent another issue that has syntacticians complaining about the way typologists classify languages. Sailor & Mahajan (2012: Sec. 2) illustrate that Hixkaryana and Tuvaluan, although apparently united by an OVS word order, differ in being head-initial and head-final, respectively, and advocate the primacy of a derivational approach that could capture the difference between the two languages. This implies that the four-way typology proposed by Dryer (2013a: 271, 284-286) is not sufficient in defining the rare OVS type, because it does not behave consistently in a way that is opposite to the SVO type.

A second objection to the OVS type comes from Haider (2023), who argues that the order Patient-Verb-Agent that characterizes certain ergative languages has been misrepresented as an OVS order (see also Alfieri, Arcodia & Ramat 2021: Sec. 3 on this). According to Haider, such cases should be reinterpreted as a syntactically SVO pattern, given that the typological word order is inferred by a wrong conflation of syntactic and semantic relations. This change in the analysis of such languages would somewhat increase the number of SVO languauges bearing ergative alignments, which are so rare crosslinguistically (Svenonius 2024: Sec. 2.4) to make Mahajan (1997) propose a generalization banning them altogether. The arguments

made by Déchaîne & Manfredi (1998) against the claim in Mahajan (1997) represent an early example related to the still unsettled issues related to the adequate treatment(s) of ergativity, which would require a dedicated discussion of its own (see Deal 2015, 2016, Baker & Bobaljik 2017, Polinsky 2017, Tollan 2019, Lindenbergh & Zwart 2021, Haider 2022a, Alexiadou 2022, Sauppe et al. 2023).

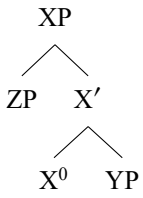
The major problems that have just been pointed out for languages that are classified as OVS also extend to the SVO ones. In fact, both word order types turn out to oscillate between patterns typical either of head initial or of head final languages, so that SVO languages are not better than the OVS ones with respect to potential misrepresentations. In fact, Haider & Szucsich (2022b) claim that, as far as SVO is concerned, at least two superficially overlapping types need to be distinguished. On the one hand, there are languages like English that are rigidly SVO in unmarked transitive clauses, while on the other hand there are others that show a freer word order in the same contexts, as many Slavic languages seem to do. The latter proposal, which will be discussed further in section 2.4.2, has however encountered several counterexamples (Haider & Szucsich 2022a, Adam & Hölzl 2024) that are symptomatic of the still ongoing debate.

To summarize, the VS/SV and OV/VO splits proposed by Dryer (2013a) represent an elaboration of earlier syntactic proposals. The major merits of this four-way typology is the identification of common traits among the languages that are exceptionlessly V-initial; ditto for the consistently V-final ones. Moreover, the somewhat peculiar status of both SVO and OVS emerges from Dryer's classification. In fact, SVO is either the only basic order or a common counterpart to some V-initial or V-final alternative. OVS is on the other hand rather rare crosslinguistically, and seems to always share the properties of other basic orders. However, Dryer's proposal has raised several objections which illustrate that neither SVO nor OVS languages form internally consistent classes at a relevant theoretical level. This problem is intertwined in turn with other unsettled debates, such as the exact formal properties of the basic orders and a crosslinguistically consistent definition of nominal arguments.

2.2.3 From the NP to the DP hypothesis

The approaches to be discussed in this section stem from the syntactic formalization known as the X-bar schema, first introduced in Chomsky (1970).[2] All the X-bar schemata consistently assume a tree structure that establishes a relation among three elements: a Head (H), a Complement (Comp) and a Specifier (Spec). Although works like Jackendoff (1977) show that there was little consensus in the beginning with respect to the actual structure of X-bar schemata (see also Punske 2014: 66 ff.), the view that prevailed posits a binary-branching structure where only Head and Comp can be sisters to each other. Spec is in a local relation with H by being the in a sisterhood relation with the node that immediately dominates H and Comp. The representation in (4) illustrates this by indicating the Head with the node X^0, while the node YP stands for the Complement.

(4)



The sisterhood relation between X^0 and YP is represented by the fact that they are on the same level and each of them is linked to the immediately dominating node X′ by means of a single branch. As the identical letter suggests, the label of this dominating node is determined by the Head, which projects its own. Fukui & Narita (2014: 10) define this process as "labeling by projection", as follows:

(5) Each phrasal constituent is a projection of a lexical item (LI, X^0) it contains.

The projection of the Head's label proceeds further to XP, which is a node that immediately dominates X′ and the Specifier ZP, so that ZP and X′ are also sisters to each other just like X^0 and YP are.

In origin, the formulation of X-bar was a set of constraints on the formal properties of Phrase Structure Rules (PSRs, see Fukui & Narita 2014: 5-9) and involved only lexical categories, so that the consequent formalization was still compliant with an Immediate-Constituent approach (IC), which Krivochen (2024: 712) summarizes as follows:

[2] See Graffi (2001: 437-438) and Corver (2013: Sec. 2) for a historical overview of its development. See instead Lichte (2021: Sec. 3.1.2) and S. Müller (2023b: Sec. 2.5) for some recent discussions.

> The crucial insight of IC 'labels' is that they specify distributional properties of the strings they exhaustively dominate (directly or transitively). Therefore, if we have a symbol N[oun] P[hrase] that dominates any other symbol or set thereof, it means that whatever set of symbols [...] is properly contained in the NP can appear within a string whose distribution and semantic properties are that of an N.

However, in later developments of generative theorizing the X-bar schema ceased to be compliant with the IC approach mentioned above, for two main reasons. On the one hand, it has been argued since Kayne (1984) that the binary-branching structure posited in the X-bar schema has to be exceptionless, so as to obtain structural uniformity and to avoid both unary and *n*-ary branching structures, where $n > 2$. On the other hand, the X-bar schema was assumed to always apply in compliance with *endocentricity*, namely the idea that any X-bar projection is determined by the Head it contains as per the defition in (5). The latter principle was enforced by positing (following Stowell 1981) that X-bar should capture not only the syntactic properties of lexical (a.k.a. contentive) heads, but also of the functional (a.k.a. noncontentive) ones (see Wiltschko 2014: Sec. 2.2, Travis 2014, Rizzi & Cinque 2016 for recent summaries about the distinction between the two).

This distinction is fully articulated in Chomsky (1986) to solve the issue related to the projection of the clause, whose root was labeled *S* (for 'sentence') even though no head was projecting this label by following (5). Given that complementizers and the inflection of finite verbs were already playing a relevant role in the syntactic analyses (see Corver 2013: Sec. 5 and Travis 2014: Sec. 2), they were both promoted as heads hosting their own X-bar projection, thus paving the way for the hierarchy of heads in the clause which, with some refinements, is currently assumed in Minimalist syntax (Chomsky 1995, 2013, Chomsky et al. 2023, a.o.):[3]

(6) $[_{CP} \ldots[_{C'}\ C\ [_{IP} \ldots[_{I'}\ I\ [_{VP} \ldots[_{V'} \ldots V\ldots]]]]]]$

As Krivochen (2024: Sec. 3) notes, the adoption of strict endocentricity made the generative enterprise depart from an IC approach and become an Array-Based one. The assumption

[3] In particular, the I head in (6) was originally assumed to bear both tense and agreement features. Such a view has been questioned since Pollock (1989), who proposes to split I into distinguished tense and agreement heads—whence the T head widely used nowadays (see also Koeneman & Zeijlstra 2014 and Krivochen 2023c: Sec. 3 for relevant discussion). With respect to the introduction of *v*P in the structure, its role and related problems, see also Travis (2014: 58 ff.).

that functional elements are heads of their own phrases (thus projecting their own labels) means that labels no longer serve the purpose of distributional consistency (as exemplified by Krivochen 2024: 712, that was quoted on page 30). As Lichte (2015: 314) further points out, the consistent application of X-bar also requires the stipulation of a null head whenever the head of the projection is apparently missing. This aspect plays a crucial role in the cartographic approach (Rizzi 1997, Rizzi & Cinque 2016) and in many aspects that will be discussed in chapter 3.

Given these premises, X-bar requirements impose a constraint on syntactic structure by evaluating the well-formedness of sentences, making PSRs unnecessary (Stowell 1981). Furthermore, the switch from the "Lexical Head Hypothesis" to the "Functional Head Hypothesis" (Corver 2013) implies that there is no upper bound to the number of functional heads that one can assume, provided that they find some empirical motivation (see Travis 2014, Putnam 2020 and Diercks, Santorini & Kroch 2025: Ch. 13 for an overview).

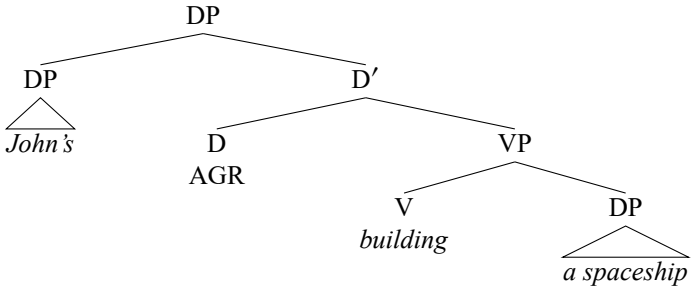
In the cartographic approach (Rizzi 1997, Rizzi & Cinque 2016, Rizzi & Bocci 2017), for instance, the basic tenet is that all the crosslinguistic syntactic variation has to lead back to one single coherent hierarchy. The latter, in turn, gives rise to "(sorted) arrays: one-dimensional ordered sequences of indexed addresses, where the source of ordering is given by Universal Grammar", as Krivochen (2024: 722) further observes. Thus, the nature and the order of functional heads is dictated by typological uniformity: no matter how morphosyntactically different languages of the world are, they can still be derived by means of syntactic movement from the sole hierarchy assumed.

Before the introduction of the binary-branching X-bar schema and of the Functional Head Hypothesis, any syntactic object modifying the noun, such as demonstratives, numerals, and adjectives were assumed to be part of a Noun Phrase (Punske 2014: 66). One of the major changes in this respect came with Abney (1987: Ch. 3), who followed the observations of Chomsky (1970) and tried to capture the different syntactic behaviors of verbal gerunds as opposed to relational nouns (which will be discussed further in section 3.1.4). Given that the former were assumed to be of verbal nature because capable of assigning case, their seemingly nominal-like distribution was problematic because the maximal NP projection was not headed in a way that complies with (5).

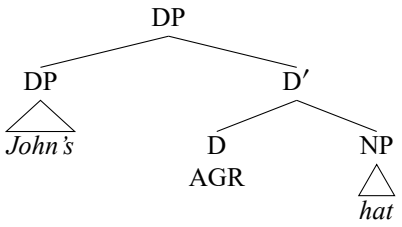
In order to capture the similarities between gerunds and other nominals and solve the issues related to proper endocentricity in the former, Abney (1987) proposes a functional projection

that dominates both the verbal projection of gerunds in (7) and the nominal projection of lexical nouns in (8). Such an additional projection is argued to be headed by a Determiner (D), which can be defined following Hsu & Syed (2020: e114) as an element contributing to the reference of the nominal expression in which it occurs (though see Giusti in press for further discussion that will be taken up in due course).[4]

(7)



(8)



The solution proposed by Abney (1987: 23 ff.) is thus similar to what Chomsky (1986) argues with respect to the clausal structure in (6), given that both works solve issues of structural endocentricity by fitting functional elements into their own independent X-bar projections. In this argumentation, the postulation of a DP is not done to fix issues related to the X-bar status of D, but because of the problematic properties of its sister lexical projection in (7).[5] The X-bar theoretic status of determiners is thus not the only basis for the DP-hypothesis (see Abney 1987: 290-297 and Salzmann 2022: 57-58), *contra* what is claimed by Bruening, Dinh & Kim (2018: 3).[6]

Abney's proposal makes emerge one parallel between the nominal and the clausal domains by claiming that in both cases the external argument (Williams 1980) occupies the specifier position of a functional head (be it D or I) so as to satisfy an agreement relation with the lexical

[4] Szabolcsi (1994: 190), Corver (2013: Sec. 8.3) and Punske (2014: 68-69) highlight that Abney (1987) does not locate *'s* in the head of D. This proposal is rather found in Abney (1986).

[5] These problems are by no means solved by the tree representation in (7), as Abney (1987: 223 ff.) himself disavows it by claiming that *-ing* is better analyzed as a nominalizing morpheme which projects an NP below DP but above VP. See Wiltschko (2014: Sec. 2.3.4) and Iordăchioaia (2020) for further discussion.

[6] By virtue of such an observation, it is even more interesting to observe the radically different discussion that Abney (1987: 288-290) and Bruening, Dinh & Kim (2018: 3) make of the very same structure that complies with X-bar theory, where DP is a complex specifier of an NP projection.

head (N or V) immediately dominated by the functional head.[7] This claim is supported by data from several languages, such as Hungarian, the Mayan language Tzutujil (spoken in Guatemala), and the Eskimo language Yupik.[8]

The examples from Hungarian are relevant for the present discussion not only as a representative case, but also because they are treated following Szabolcsi (1983: 89-90), who predates Abney (1987) in observing a parallel behavior between the two domains in such a language (see Abney 1987: 77 and Bruening 2009: 27 for further forerunners of Abney's proposal). In the examples in (9), adapted from Szabolcsi (1983: 89-90), ø indicates null morphemes.

(9) a. *az én-ø vendég-e-m*
the 1SG-NOM guest-POSS-1SG
'my guest'

b. *a te-ø vendég-e-d*
the 2SG-NOM guest-POSS-2SG
'your guest'

c. (*a*) *Mari-ø vendég-e-ø*
(the) Mary-NOM guest-POSS-3SG
'Mary's guest'

d. *Mari-ø alud-t-ø*
Mary-NOM sleep-PST-3SG
'Mary slept'

Although Abney (1987) and Szabolcsi (1983, 1994) agree that the Hungarian data in (9) suggest that there exists a parallel between the verbal and the nominal domains which involves the DP projection, their opinion diverges with respect to the exact role of the latter in the nominal structure. If the former claims that many properties of D overlap with the properties of I, the latter argues that D better parallels C (see Szabolcsi 1994: 180, 190).

---

[7] The idea that external arguments originate in VP or NP-internal positions (so that the parallel sketched by Abney is futher strengthened) is a later theoretical development known as the VP-internal subject hypothesis (Koopman & Sportiche 1991). See also Corver (2013: Sec. 8.3), Travis (2014: 57) and Wiltschko (2014: 65-66) for further discussion and references.

[8] Glottolog (Hammarström et al. 2024) shows that varieties of Yupik can be found on both sides of the Bering Strait, as well as on St. Lawrence Island. Given that Abney (1987: 39) identifies Yupik as a Central Alaskan language, he is probably referring to the variety that can be identified with the glottocode cent2127 (glottocodes will be discussed in section 4.2.1).

This is mainly due to the two different ways in which the possessor in Hungarian can be expressed. In the first one, already shown in (9c), the possessor is in nominative case and follows the optional definite article. In the second one, on the contrary, the possessor is in dative case and precedes the definite article, which cannot be omitted. As Szabolcsi (1994: 198 ff.) shows, only the latter kind of possessor can be realized discontinuously from the rest of the nominal. These facts are illustrated in (10) and (11), which are adapted from Szabolcsi (1983: 198) and Corver (2013: Sec. 8.4).

(10) a. *a Mari kalap-ja*
DET Mary.NOM hat-POSS.3SG

b. *Mari-nak a kalap-ja*
Mary-DAT DET hat-POSS.3SG
'Mari's hat'

(11) a. *[*Mari*]$_i$ *Peter látta* [$_{DP}$ *a* [t$_i$ *kalap-já-t*]]
Mary.NOM Peter saw DET hat-POSS.3SG-ACC

b. [*Mari-nak*]$_i$ *Peter látta* [$_{DP}$ t$'_i$ *a* [t$_i$ *kalap-já-t*]]
Mary-DAT Peter saw DET hat-POSS.3SG-ACC
'Peter saw Mary's hat.'

The bracketing and the traces added to (11) follow the analysis that Szabolcsi (1994) provides for these data, namely that the dative possessor can be fronted because it occupies the Spec,DP prior to the eventual extraction. The inability of the nominative possessor to extract is seen as a consequence of its impossibility to vacate the nominal domain in one fell swoop. In order to extract, it must rather raise to the Spec,DP first, where it is inevitably assigned the dative case. As opposed to Abney (1987), Szabolcsi (1994) thus proposes that the DP projection represents the escape hatch that a phrasal element within the nominal domain has to target in order to be extracted, so that it mirrors the CP in the clausal domain.

Moreover, the inflection realized in the nominal (which Abney 1987 mentions to argue for the parallel with the clausal structure) is taken by Szabolcsi (1994: 198) to be realized as an intermediate head between lexical N and the DP. This move thus paves the way for subsequent works arguing for the existence of other functional projections between N and D (see Alexiadou, Haegeman & Stavrou 2007 for extensive discussion pursing this direction).

As Bruening (2009: 27) remarks, the growth of the assumed functional projections in the

structure of both nominals and clauses and the idea that the two domains are entirely parallel were two key factors that brought to the notion of "extended projection", which will play a relevant role in section 2.3.2 (see also Riemsdijk 1998 for a related proposal and Borer 2013: Sec. 6.3 for further discussion).

Extended projections are discussed in Grimshaw (1991, 2005), where it is claimed that heads, which are assumed to project their labels in a way that is compliant with an AB approach to X-bar (see the quotation of Krivochen 2024: 712 on page 30), can be divided in two groups. The first is made of V(erbs) and N(ouns), which are the only lexical heads she posits (but see Corver 2013: Sec. 9-10 for an extension to adjectives and prepositions), while the second is made of functional heads, which are valued either as verbal or as nominal and whose syntactic organization is determined by the V and N heads they dominate. In fact, a lexical head can take as its complement any kind of head, but if it is immediately dominated by a functional head, the latter has to be of the same category of the former.

The consequence of these premises is that the clausal domain is interpreted as a verbal extended projection, because the functional heads C, T and *v* (among the others subsequently assumed—see footnote 3 and Broekhuis 2010, Ramchand & Svenonius 2014, Rizzi & Cinque 2016) form a categorially consistent projection with the lexical verb of the matrix clause by means of iterative c-selection. However, this does not entail that elements belonging to different categories cannot be interspersed in the subsequent linearization. In fact, this is allowed by the presence of specifiers, whose categorial status does not affect the definition of an extended projection, which is concerned only with c-selection between heads and their complements.[9]

2.2.4 On the necessity of functional projections above NP

The DP-hypothesis is appealing because of its capability to sketch a parallel between the nominal and the clausal domains. However, the exact nature of this parallel is something that has not been agreed upon since the very beginning of the DP hypothesis, as shown by the divergent views between Abney (1987) and Szabolcsi (1994). Moreover, the situation has not improved since, as Leu (2014: 6) and Laenzlinger (2017) illustrate.

[9] However, Grimshaw (2005: 6 fn. 5) observed that the exclusion of specifiers remains implicit, though not mandatory, as long as it is not explicitly stated. In fact, she says that "[i]t may be necessary to add the requirement that only complements, and not specifiers of L[exical]-heads or F[unctional]-heads, participate in extended projections".

The most widespread idea is that both nominal and verbal extended projections are tripartite (Grohmann 2003: 74 ff., Ramchand & Svenonius 2014: 164). The lower portion is where thematic arguments of the lexical head are merged, followed by an inflectional portion where modifiers are merged, and topped off by a "discourse related, referential and quantificational domain" (Laenzlinger 2017: 2).[10] However, the exact hierarchical relation among the various elements, the relevant features and the types of movement that can explain the facts have received contradictory analyses, as the remainder of this chapter and chapter 3 will show.

The treatment of nominals as NPs, which had never been challenged before the formulation of DP-hypothesis, has resisted among scholars who are skeptical about the strength of parallels favoring the DP-treatment, thus making the choice between the NP and the DP hypotheses a matter of ongoing debate (of which Blümel & Holler 2021, Köylü 2021, Salzmann 2022 provide good overviews).[11] This is how Punske (2014: 71) summarizes two currently opposing views on the matter:

> On one hand, we have the Cartographic approaches to the DP, with an ever-expanding array of functional heads above the noun; and, on the other, we have the Chomsky-style approach with reduced functional structure and reassertion of the noun as the central part of the nominal phrase.

The "Chomsky-style approach" is labeled as such because the DP-treatment of nominals has not only found little support from Chomsky (see Pullum & Miller 2022: Sec. 6), is but also explicitly criticized in Chomsky (2007: 26) and in Chomsky (2020). Bruening (2009), Georgi & Müller (2010), and Bruening, Dinh & Kim (2018) represent relevant works in favor of this approach, where it is argued that the nominal domain lacks functional projections altogether, while the the clausal domain can still be analyzed as the C-I-V system shown in (6) on page 30. The "Cartographic approaches" identified by Punske (2014) actually include a broader set of views which are distinguished from the opposite approach to nominals by claiming that at least *some* functional projection above N is needed to better account for the empirical facts.

[10] Wiltschko (2014: 74 ff.) proposes instead a four-layered structure for both the nominal and the clausal domain, with the additional layer being defined by aspectual properties and located immediately above the lowest (thematic) one. See instead Erlewine (2017) for the proposal of a structural portion above CP in the clausal domain.

[11] See Eynde (2020) and Eynde (2021: Sec. 2) for a treatment of this debate within the Head-Driven Phrase Structure Grammar (HPSG) framework and Osborne (2021) for the same concern with respect to Dependency Grammar (DG).

A first aspect on which the two opposing views tend to disagree is represented by head movement, namely the idea that a head can raise to the position of the head that properly governs it (i.e. the Head Movement Constraint, Travis 1984: 131 and Ritter 1991: 39) and that this operation can iterate across heads belonging to the same extended projection (see Zeijlstra 2023: 200-201). This operation plays a relevant role in the literature to account for the intra- and cross-linguistic word order variation with respect to the position of verbal elements in the clause.[12]

An often-cited example is the minimal contrast between the English and the French clauses in (12), where the different adverb placement in the two languages is taken as evidence (together with differences related to the verbal morphology) for the V-to-I raising of the main verb in French, which on the contrary is not possible in English (see Delfitto 2006: Sec. 5.3, Travis 2014: Sec. 4 and Koeneman & Zeijlstra 2014 for discussion). What is reported in (12) follows the adaptation that Broekhuis (2008: 36) makes of one of the contrasts discussed by Pollock (1989: 367).

(12) a. John <*kisses> often <kisses> Mary.
b. Jean <embrasse> souvent <*embrasse> Marie.

As section 2.2.3 anticipated, this line of reasoning has brought a twofold consequence. On the one hand, it has fostered the assumption of several functional heads in the verbal extended projection. On the other hand, the parallel behavior of the nominal extended projection that the DP hypothesis suggests has led to state that word order variation in nominals is due to the Noun moving to higher functional heads as well. Bernstein (2001, 2008), Punske (2014), Sessarego (2014: Ch. 4) and Laenzlinger (2017) provide good summaries of the functional heads proposed in this respect (see also Corver 2013: Sec. 8.9 for a broad list of references).

In the present context, the focus will be directed towards the N-to-D and the N-to-Num movement. While some arguments for the former are already discussed in Abney (1987: 285-287) with respect to English, the latter is proposed by Ritter (1991), who argues for the need of a NumP projection below DP but above NP to account for the full range of word order facts related to the genitival constructions in Modern Hebrew.

There are two ways in this language to express the genitival constructions, namely the Construct State construction or the Free State construction. In the former (i.e. CS), the head

[12] However, head movement bears many characteristics that are problematic for the theory. See Georgi & Müller (2010: Sec. 3.1) and Hall (2015: 29) for concise but detailed overviews and Dékány (2018) and Harizanov & Gribanova (2019) for broader discussion.

noun cannot be preceded by the definite article *ha-* and no bridging element can occur between the head noun and the closest argument, which in turn must be adjacent to it. This happens at the expense of adjectives (see section 3.1.4), which are positively marked for definiteness if the argument with which the head noun forms the CS is also definite (see Borer 1999: 43-44 for a more detailed list of the differences between CS and FS constructions). These facts are illustrated in (13). Except for (13e), which is adapted from Adger (2013: 161) and Bruening (2022: 5),[13] all examples are adapted from Ritter (1991: 40).

(13) a. *ha-bayit*
DET-house
'the house'

b. * *ha-beyt* *ha-mora*
DET-house.CS DET-teacher

c. * *ha-beyt* *mora*
DET-house.CS teacher

d. *beyt* *ha-mora*
house.CS DET-teacher
'the teacher's house'

e. *tmunat* *ha-xamanyot* *ha-yafa* *shel* *van* *gox*
picture.CS DET-sunflowers DEF-beautiful of Van Gogh
'Van Gogh's beautiful picture of the sunflowers'

While the behavior of the simple nominal in (13a) can be captured by the DP structure in (8) proposed by Abney (1987) with the article being an overt realization of the D head, its absence in (13d) is explained by assuming that the lexical noun raises to (the genitive-Case-assigning) D to make the latter visible (Ritter 1991: 40).

Instead of directing the discussion towards CS constructions like (13e), Ritter (1991: Sec. 2-4) focuses on the contrast between CS and FS constructions involving the same deverbal noun reported in (14). As opposed to CS, in FS constructions each argument is not bare, the head noun can be preceded by the article if it is definite and there is no adjacency re-

[13] Bruening (2022) does not mention Adger (2013) as a source for the example in (13e), but rather a p.c. from Itai Bassi. It is also worth noting that the transcription of the same Modern Hebrew examples can vary among scholars. In particular, *shel*, *šel* and *ʃel* represent alternate transcriptions of the same element. See also Bruening (2022: 2 fn. 1) in this respect.

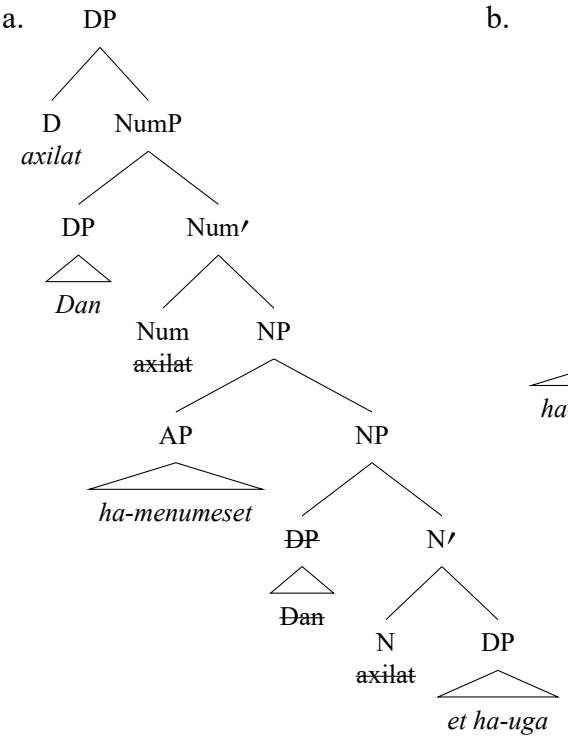
quirement between the head noun and the closest argument, so that the former is immediately followed by any adjective by which it is modified.

(14) a. *axilat Dan ha-menumeset et ha-uga*
eating.CS Dan DEF-polite ACC DET-cake

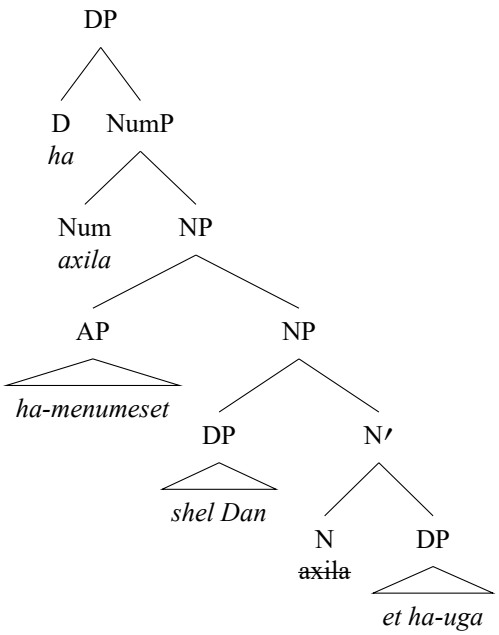
b. **ha–axila shel Dan ha-menumeset et ha-uga.*
DET-eating of Dan DEF-polite ACC DET-cake

c. *ha–axila ha-menumeset shel Dan et ha-uga.*
DET-eating DEF-polite of Dan ACC DET-cake
'Dan's polite eating of the cake'

(15) a.



b.



As the tree representations in (15) show, a head-movement account that is grounded on the assumption of an additional functional head between D and N manages to capture the contrast between the equivalent CS and FS constructions in (14). Ritter (1991) argues that such a head is the locus of number marking, whence the *Num* label (see section 3.1.2 for further discussion). In both (15a) and (15b), adjectives are instead adjuncts extending the NP node (though see section 3.1.3 for competing analyses). Finally, Ritter (1991: 43) motivates the structure of the rest of the NP on the basis of the fact that the argument in Spec,NP can bind into the complement of N (see also Bruening 2022: Sec. 3 for thorough discussion of the binding facts in Modern Hebrew nominals).

Although several alternatives have been proposed to solve some issues arising from Ritter's analysis (Borer 1999, Shlonsky 2004, Pereltsvaig 2006, Adger 2013, Borer 2013, Preminger 2020, Bruening 2022, a.o.), two problems in the structures in (15) are relevant for the present discussion. The first is that the raising of the DP in Spec,NP to Spec,NumP makes the wrong prediction with respect to the nominal in (13e) from Bruening (2022: 5).

To account for (13e), one has to assume that the DP raising to Spec,NumP is the one originating as a complement of N, which is in contrast with (15a) and represents a Relativized Minimality violation (Rizzi 1990, 2001) if *shel van gox* occupies Spec,NP. This suggests that the nominals in (13e) and (14a) have different structural properties, which leads to the second problem in Ritter's analysis, namely that the distinction between deverbal and underived nominals plays no role in her account, although it is relevant in Abney (1987), as (7) and (8) in section 2.2.3 illustrate (see also Iordăchioaia 2020, Eynde 2021: Sec. 3.2 and references therein).

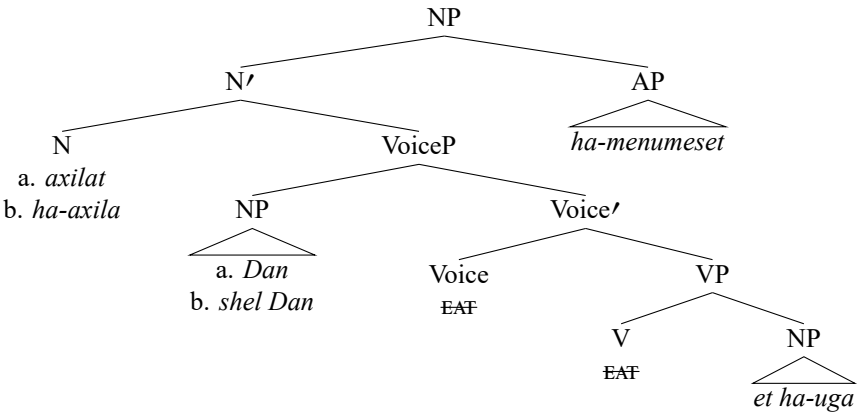
The differences between the two types of nominals in the Modern Hebrew genitival constructions are instead considered in the analysis proposed by Bruening (2022), who argues that all nominals are NP projections that are not dominated by any functional nominal material. In line with Grimshaw (1990) and Borer (2013), Bruening (2022) captures the contrasts between deverbal nominals like (14a) and underived nominals like (13e) by assuming that the NP projection of the former contains a VoiceP projection where the arguments occurring in the nominal are base-generated.[14] On the contrary, underived nominals lack VoiceP, so that arguments branch under iterated N′-levels in the projection of the head noun (see Bruening 2022: 20).

Starting from the structure in (17), Bruening (2022: 10-11, 16) captures the contrast between (14a) and (14c) by assuming that the object argument raises rightward in the former, while the entire VoiceP raises in the latter. Both raisings are due to an independent prosodic requirement that Bruening (2022: 8) assumes for CS constructions:

(16) A [+CS] N must form a prosodic word with phonological material dominated by the NP that checks its [+CS] feature.

[14] One can safely equate VoiceP with the *v*P assumed from Chomsky (1995) onwards and also labeled as such in Bruening, Dinh & Kim (2018). See also footnote 3.

(17)



It should be observed that the definite article *ha-* does not occupy a different structural position with respect to the noun head, which in turn is assumed to be in a sisterhood relation with respect to VoiceP in order to solve the problem of endocentricity in a manner that is quite different from (7) (Abney 1987: 23 ff.) but somewhat similar to Abney (1987: 223 ff.) (see footnote 5). According to Bruening (2022: Sec. 4.1), having no explanation (or a stipulative one based on [±Def] and [±CS] features)[15] for the behavior of the definite article in CS and FS constructions is as good as having an inadequate account that relies on several assumptions. A similar claim is also made in Bruening (2018a), which will be discussed in section 3.2.5.

Bruening (2022) labels his analysis where Modern Hebrew genitival constructions project NPs as an *endocentric* analysis, as opposed to the analyses that require DPs, NumPs (or any other functional projection) above NP, which he labels as *exocentric*. Although the crucial distinction between the two is related to the nature of the categories and featural makeup that are argued to be necessary to account for the facts (Baggio 2022: Sec. 1), the terminological contrast is somewhat misleading, given that both endo- and exo-centric structures comply with the labeling by projection (Fukui & Narita 2014: 10) defined in (5) on page 29. This opposition is perhaps better captured by opposing the Functional and the Lexical Head Hypotheses (Corver 2013) discussed in 2.2.3, so that Bruening (2022) applies the FHH to the clausal domain and the LHH to the nominal domain, while Abney (1987), Ritter (1991), and Szabolcsi (1994) apply the FHH to both.

Both views share a twofold crucial assumption. On the one hand, that there is one and only universally valid approach to the maximal projection of nominals. On the other hand, the relevant arguments in this respect must either support or reject the parallel between C and

[15] In (17) *shel* is placed within NP because Bruening (2022: 23) sees its insertion in (14c) as a result of Agree between the argument and the head noun bearing a [−CS] feature, whose value also determines the prosodic constraint in (16).

D proposed by Szabolcsi (1994), where both C and D are argued to behave as subordinators that enable "the clause or the nominal expression to act as arguments (i.e., theta-role bearing constituents)" (Corver 2013: Sec. 8.7).

The latter view is famously defended by Longobardi (1994) and following works (see Giusti in press: Sec. 2 for an overview and Larson 2019 for a recent paper supporting it), where it is argued that nominals need a DP projection in order to be selected as arguments, even when no overt article is present. This idea is supported in Italian and other Romance languages by word order facts that can be accounted for by claiming that nouns undergo overt N-to-D movement when no article is present (see also Bernstein 2001, 2008 for further arguments in this direction and Georgi & Müller 2010 for a counterproposal under the LHH, to be discussed in section 3.2.4).

Bruening (2009) and Bruening, Dinh & Kim (2018) predate the claims made by Bruening (2022) on the basis of observations from English, Korean and Vietnamese with respect to c-selection and local goverment selection between heads in idioms. Bruening (2009) and Bruening, Dinh & Kim (2018) claim that in both types of selections the selectee cannot vary freely and its categorial properties can be inferred from the features it does (not) show. With this in mind, they observe that in English a main Verb can c-select (a subset of) CPs, TPs or *v*Ps, as (18) and (19), adapted from Bruening (2009: 27-28) and Baggio (2022: Sec. 4.1), illustrate. However, the crucial symmetric behavior that the FHH of nominals would predict is not borne out for English. As facts from the made-up verb *glorp* (Baltin 1989: 3) show, nominal objects selected by a main verb in (20a) and (20b) never mirror the selectional patterns in (18), which means that selection for DP does not work as expected. There is also no compelling evidence of verbs that never select the bracketed determiners in (20c), as a selection for NumP would expect.

(18) a. Sue thinks [$_{CP}$ that Sam swam with dolphins]. [V c-selects CP]
b. *Sue thinks [$_{CP}$ whether Sam swam with dolphins].
c. *Sue wonders [$_{CP}$ that Sam swam with dolphins].
d. Sue wonders [$_{CP}$ whether Sam swam with dolphins].

(19) a. Sam$_i$ appears [$_{TP}$ t_i to be swimming with dolphins]. [V c-selects TP]
b. *Sam$_i$ appears [$_{TP}$ t_i is swimming with dolphins].
c. Sue saw [$_{vP}$ Sam swim with dolphins]. [V c-selects *v*P]

(20) a. John glorped the book.
b. John glorped a book.
c. John glorped (the/those) books.

Although Bruening, Dinh & Kim (2018: Sec. 4) foresee some counterarguments to their proposal, three types of objections that have subsequently been advocated in defense of the FHH to nominals. The first one does not consist in criticizing the premises of a LHH to nominals, but rather in defending the FHH by arguing that it is superior in its descriptive and explanatory adequacy. Bruening (2022) is actually a rebuttal of one of such analyses (i.e. Preminger 2020), while Bruening (2020) replies to some of such claims (based on N-to-D movement, mixed concord and selection in idioms) by sharing with Georgi & Müller (2010) and Salzmann (2020) the idea that most defenses of the FHH are not compelling, but rather based on theory-internal premises (mostly descending from X-bar theory, see section 2.2.3) that can be restated under the LHH to nominals with virtually no loss of adequacy (a fact already problematized by Stowell 1989: 234-235).

The second type of objection is related to the flawlessness of the aforementioned arguments, which are questioned by Larson (2019), Barrie et al. (2021: Sec. 4), Baggio (2022: Sec. 2-3) and Salzmann (2022: Sec. 3.2). On the one hand, the latter authors provide counterexamples where selection does not work as rigidly as expected in clausal projections, or alternatively where it applies to the functional projections that are assumed for the nominal domain (see also Pereltsvaig 2021 on the latter). On the other hand, the same authors show that the assumptions made by Bruening (2009) and Bruening, Dinh & Kim (2018) may lead to contradictory claims. The last line of counterarguments comes instead from Melchin (2019: Ch. 2) and Baggio (2022: Sec. 4.3), who argue that facts from (18)-(20) are better explained through s(emantic)-selection, rather than through c(ategorial)-selection, so that the FHH for nominals is still a valuable syntactic approach, *contra* Bruening (2009) and Bruening, Dinh & Kim (2018).

A middle ground between the two opposing views illustrated so far consists in claiming that both the FHH and the LHH are possible, although subject to crosslinguistic parameterization. Chierchia (1998) proposes in this respect a semantically-driven tripartition of the crosslinguistic variation, which is determined by the fact that a language must positively mark at least one of the two binary features that he assumes to hold for any noun, i.e. [± predicate] and [± argument] (see Giusti in press: Sec. 2 for a detailed summary of the subsequent lan-

guage types and their properties).

If on the one hand the contrasts that Longobardi (1994) highlights between Italian and English nominals have often been taken as first evidence in favor of the universality of DP, Chierchia (1998) challenges this perspective by stating on the other hand that languages of the [-pred, +arg] type completely lack the DP projection. This typological difference leads to the prediction that the NPs of these languages lack the mass/count distinction (by exclusively having a mass denotation).

These claims are however challenged by many scholars who argue that a mass/count distinction does hold in Mandarin Chinese, even though Chierchia (1998) mentions it as an example of [-pred, +arg] language (see Wu & Her 2021 and references therein for thorough discussion). And even though this proposal is reformulated in different terms by Chierchia (2010), Rothstein (2017: 147 ff.) shows that there are exceptions that Chierchia's refined typology still fails to predict.

A syntactically determined parameterization on DP vs NP languages is proposed instead by Bošković (2005) and subsequent works, where both C and D are deemed to have phasal properties (see Chomsky 2008, Citko 2014, a.o.). Thus, the presence vs. absence of a DP in a language should play a relevant role in blocking or allowing syntactic operations within and outside of nominals (see Bošković 2016a for a summary and discussion).

However, this proposal faces problems of theoretical consistency (see Bruening, Dinh & Kim 2018: Sec. 4.5) as well as difficulties in being empirically confirmed. Contrary to both Longobardi (1994) and Chierchia (1998), Bošković (2005) argues that DP is never projected in any language that lacks articles. However, some studies on articleless languages disconfirm the predicted correlations (see Giusti in press: Sec. 4 and references therein), while other works like Syed & Simpson (2017) and Köylü (2021) point out that it in certain languages the provided the diagnostics related to the presence/absence of DPs can be contradictory or impossible to evaluate.

As already pointed out, the FHH and the intermediate positions just illustrated hinge on the parallel between C and D proposed by Szabolcsi (1994), while the LHH put forth by Bruening (2009) and Bruening, Dinh & Kim (2018) crucially rejects it. However, Barrie et al. (2021) and Baggio (2022: Sec. 4.1-4.2) question the parallel which sets the basis for the whole debate and subsequently propose a revized FHH. In fact, both works observe that some functional projection dominating DP would hinder the latter to be selected in the ways that

a parallel between C and D would expect. In Barrie et al. (2021: Sec. 3), such a projection is KP, first proposed by Lamontagne & Travis (1987). KP is motivated by nominals that bear lexical case (in the sense of Woolford 2006), that are selected for a specific case by a preposition, or that show non-affixal case particles. These three arguments are exeplified in (21) following Woolford (2006: 114), in (22) based on Barrie et al. (2021: 214-215) and in (23) from Bittner & Hale (1996: 4-5), respectively.

(21) a. *bátnum hvolfdi* [Icelandic]
boat.DEF.DAT capsized
'The boat capsized.'

b. *bátinn rak á land*
boat.DEF.ACC drifted to shore
'The boat drifted to the shore.'

c. *Jóns nýtur ekki lengur við*
John.GEN enjoys NEG longer at
'John is no longer available.'

(22) a. *ich gehe mit ihm* [German]
1SG.NOM go with 3SG.M.DAT
'I go with him.'

b. *ich gehe ohne ihn*
1SG.NOM go without 3SG.M.ACC
'I go without him.'

(23) a. [*waitna ba*] [*sula ba ra*] *kaik-an* [Mískito]
man DET deer DET ACC see-PST.3SG
'The man saw the deer.'

b. *ka la yo"ii* [*ya 'u khlaa*] [Khasi]
3SG.F PST see ACC DET tiger
'She saw the tiger.'

However, showing the necessity of KP to account for certain linguistic facts does not imply that it must be assumed always and everywhere, especially if there is no direct evidence for it. In particular, one may object that KP is not necessary for the assignment of structural case (as Woolford 2006 and Butt 2006: Ch. 4 define it) in English, which is however relevant

in the discussion of (20), so that there are cases where KP and DP provide equivalent analyses differing only in the labels used (Hicks 2005: Sec. 4.4.1).[16] This aspect is not tackled by Barrie et al. (2021), but implementations towards this direction are proposed by Neeleman & Szendröi (2008), Levin (2015), and Saito (2018), while some empirical challenges are discussed by Sheehan & van der Wal (2016, 2018).

If C is not paralleled by D but rather by another functional head c-commanding D, then the FHH proposed by Barrie et al. (2021) is likely to assume that D should parallel I, as per Lamontagne & Travis (1987: 176-177) and the claim in Abney (1987) discussed in section 2.2.3. And in fact, this observation would be in line with Bittner & Hale (1996) and Wiltschko (2014: 74), who suggest that I and D are parallel to each other, as are C and K.

Baggio (2022: Sec. 4.1) also argues for the similarity between D and I, but KP plays no role in his argument. His attention is instead directed towards an often overlooked aspect of Grimshaw (1991, 2005), namely that in her proposal nominal extended projections can be topped off by prepositions (see also Giusti in press: Sec. 3). This suggests that the real parallel is not between C and K, but rather between C and P.[17] Although Baggio (2022: Sec. 4.1) tries to demonstrate the plausibility of the latter, Diercks, Santorini & Kroch (2025: Sec. 6.3.2) show that such a parallel would not be immune from problems either.

To conclude this section, one can get back to the citation given at its beginning following Punske (2014: 71), where all the FHH treatments of the nominal domain are labeled as "Cartographic approaches", even though the latter term should only refer to works that explicitly follow Rizzi (1997). The latter work proposes the splitting of CP into a series of functional, contiguous projections to account for facts like discourse-related phenomena (Rizzi 2013: Sec. 2-3) and the co-occurrence of two complementizers in the same sentence (Rizzi 2013: Sec. 6).

At first sight, an extension of this idea to the DP projection would allow to reconcile the parallel between C and D (Szabolcsi 1994) with the empirical facts supporting the need of additional functional projections in the highest portion of the nominal structure (Barrie et al. 2021, Baggio 2022). And in fact, Zamparelli (2000) is a work that splits the DP into layers that manage to account (among other things) for the syntactic distribution of a phrase like *all*

[16] This problem is further strengthened by the relational nature of case that is dictated by Agree (Chomsky 2000, 2001) by means of features that do not belong to the elements in the lexicon. The universality of KP is also challenged by Bittner & Hale (1996), who use this projection to account for ergativity as opposed to the absolutive case (see also Deal 2015: 674 ff. on this).

[17] Interestingly, the connection made in footnote 16 between ergativity and KP faces a similar problem, PP being an alternative proposed in the literature. See Deal (2015: Sec. 3.2.1) and Polinsky (2017: Sec. 4.2).

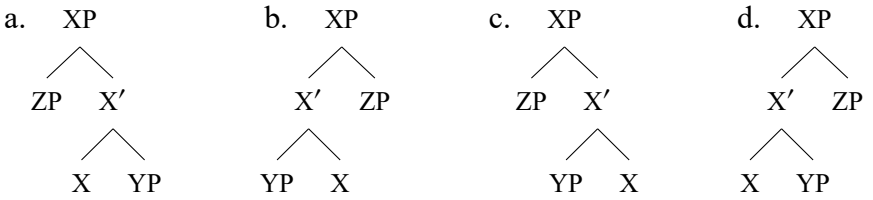
the students, which can be theoretically problematic due to the co-occurrence of a quantifier and a determiner (see Laenzlinger 2017: Sec. 2.2, Gianollo, Heusinger & Napoli 2021: 4 ff. and Giusti in press: Sec. 3.3 for references and discussion).

Still, proposals which split the DP raise problems related to the upper bound and the exact properties of the assumed functional projections, and further confirm the lack of a detailed consensus between proponents of the FHH to nominals. This represents a major issue with the latter approaches, regardless of the relevant attempts aimed at unifying these views (e.g. Wiltschko 2014, Giusti in press). In fact, the emerging inconsistencies seem to favor the opposite approach, i.e. the LHH. However, chapter 4 will show that the nominal and the verbal domains somewhat affect one another, which is exactly what the LHH predicts to be impossible.

2.3 Antisymmetric approaches

The X-bar schema introduced in (4) on page 29 does not entail that the base-generated structure should be linearized in one single manner. By following the common assumption about linearization that nodes in the tree are visited by following the pre-order traversal,[18] the sisterhood relation between a Head X and a Complement YP as in section 2.2.3 excludes the linear orders H-Spec-Comp and Comp-H-Spec from the possible options, but nothing is said about the four remaining alternatives illustrated in (24):

(24)



Thus, if X were a verbal lexical head and ZP and YP two arguments it selects (which is plausible according to VP-internal subject hypothesis, see footnote 7), the X-bar schemata can base-generate two head-medial options, as in (24a) and (24b), as well as a head-final and a head-initial alternative, respectively (24c) and (24d).

In works like Chomsky (1981) and Stowell (1981) it is assumed that the selection of one of the four templates in (24) is subject to both cross- and intra-linguistic variation, so that

[18] See Krivochen (2023b: Sec. 3 fn. 3) for definitions and references, as well as Toyoshima (2013) for an attempt to capture word order variation by assuming trees to be constant, while the way to traverse them can vary.

facts related to *dominance*, i.e. hierarchical arrangement, can be kept separate from aspects of *precedence*, i.e. linear arrangement, "although both are still regulated in syntax" (Corver 2013: Sec. 4). In particular, Biberauer, Roberts & Sheehan (2014: 47 ff.) underline that in works following Chomsky (1981) X-bar schemata are taken as an invariant principle holding universally across languages, while the choice of one of the four alternatives in (24) is regulated by two independent binary parameters (whose alternate values are capitalized in 25):

(25) a. Spec Parameter: Spec PRECEDES/FOLLOWS X′
b. Head Parameter: X PRECEDES/FOLLOWS Comp

The readers may have noted that the crosslinguistic variability of mono-transitive clauses predicted by the two options in (25) overlaps with the SV/VS and OV/VO splits proposed by Dryer (1992, 2009) for his four-way typology discussed in section 2.2.2. However, the structural properties of the X-bar schema make different predictions, namely that SVO (24a), OVS (24b), SOV (24c) and VOS (24d) are parametric variants of the same structural principle (see also Moro 2008: 208-209). This leaves unaccounted the high frequency across languages of SVO and SOV as opposed to the high rarity of OVS and the unfrequent attestations of VOS (see Dryer 2013f and section 4.2.3). Moreover, this model needs to posit derived positions so as to account for VSO and OSV orders. Although this might be understood as a factor explaining the crosslinguistic rarity of OSV, it struggles to predict that VSO is the third most common order and that it arguably shares some structural properties with VOS (as was discussed in section 2.2.2).

Bearing in mind what has been discussed up to this point, the remainder of the chapter will illustrate the alternate syntactic accounts that have been proposed to cope with the issues arising from the parameterization of the X-bar schema stated in (25).

2.3.1 Kayne (1994)

Kayne (1994) is one of the most influential amendments to (25), which raises the underspecification of "directionality parameters" (Kayne 2013) related to the X-bar template. Kayne (1994: Sec 4.1) claims that X-bar theory can be derived as a corollary from the Linear Correspondence Axiom (LCA), namely a mapping from the dominance relations on nonterminal nodes in the structure to the linear order of terminal nodes, since both are orderings charac-

terized by being transitive, (locally) total, and antisymmetric (Kayne 1994: 4 and Guimarães 2000: 94).[19] The theoretical premises of Kayne (1994) are summarized in (26) by modifying Biberauer, Holmberg & Roberts (2014: 207) for expository purposes:

(26) a. *C-command*: α c-commands β if and only if α is a category and β is contained in the sister of α.[20]

b. *Asymmetric c-command*: α asymmetrically c-command β if and only if α c-commands β and β does not c-command α.

c. *LCA*: α [is linearly ordered with respect to] β if and only if α asymmetrically c-commands β or if α is contained in γ, where γ asymmetrically c-commands β.

Although the original formulation of c-command in Reinhart (1981) does not include the distinction between segments and categories (nor does it predict that c-command can be asymmetric, as Krivochen 2023b: Sec. 1.1 observes), Kayne (1994: 11-17) uses it (by following May 1985 and Chomsky 1986) to identify categories with maximal projections. This is crucial so as to establish that ZP in (24) asymmetrically c-commmands the nodes dominated by X′ (given that X′ is an intermediate non-maximal projection, it fails to c-command the nodes contained in the specifier ZP).

More generally, given two sister nodes in the structure, (26) states that one of the two nodes must be a category while the other must not, otherwise asymmetric c-command fails and the terminal nodes cannot be properly ordered. Kayne (1994) further claims that the LCA rules out *n*-ary branching structures ($n > 2$) that Kayne (1984) already excluded, as well as the adjunction of heads to non-heads and vice versa (so that the Chain Uniformity Condition in Chomsky 1995: 252 is obeyed), the legitimacy of multiple adjunction to a single head and of multiple specifiers within a single projection. In fact, one single phrasal adjunct per phrase is allowed, and that instance of adjunction is the specifier itself (Kayne 1994: 23).

These claims are supposed to hold if one assumes along with Kayne (1994: 48-49) that the LCA in (26) must apply at all intermediate stages of the derivation. An alternate view is to assume that the LCA does not need to be respected before Spell-Out (Uriagereka 1999, 2012), i.e. the crucial step where the structure is sent to the Phonological Form (PF) and thus

[19] See also Morzycki (2016: 284-297) for related definitions and Malčić (2019) for further theoretical discussion.

[20] Biberauer, Holmberg & Roberts (2014: 207 fn. 33) observe that if *containment* is understood as an irreflexive relation (so that a node cannot contain itself) the notion of asymmetric c-command stated in (26b) becomes redundant.

needs to be linearized by following (26).

Moro (2000) and Moro & Roberts (2024) show that base-generation of a structure where two or more nodes symmetrically c-command each other is theoretically possible, but then movement must be used as a means to repair this situation prior to Spell-Out, so that a legitimate linearization can be mapped from the structure. Although the latter PF-oriented view and Kayne (1994) appear to be rather different, Krivochen (2023b: Sec. 1.4) observes that both need to restate apparent points of symmetry by adding some structure that breaks them, so that the desired asymmetric c-command relations can be established.

Another crucial aspect of (26) that Kayne (1994: 35) discusses is that the linear ordering determined by asymmetric c-command is not subject to intra- or cross-linguistic variation. It is instead a no-choice mapping parameter (Roberts 2018) that is consistently set on either precedence or subsequence (i.e. respectively "fully left-to-right or fully right-to-left" in the terms of Chomsky 2014–1995: 360 n. 108). Thus, (26) ultimately requires "that specifier and complement be on opposite sides of the head" within an X-bar projection.

Two options are then available for the syntactic hierarchical structure: Spec-H-Comp and Comp-H-Spec, respectively depending on whether the LCA consistently determines precedence or subsequence. Kayne (1994: Sec. 5) opts for the former alternative (see also Kayne 2013), a fact which also has a typological underpinning: Comp-H-Spec (i.e. 24b) would imply a universal base-generation of the Object-Verb-Subject order (OVS), which according to Greenberg (1966–1963: 76) is "excessively rare" (see also Dryer 2013f). The alternative Spec-Head-Comp base-generated structure (24a) is on the contrary mapping the S-V-O order, which Dryer (2013f) shows to be highly frequent crosslinguistically (see also section 4.2.3). The claim put forth by Kayne (1994) that all phrases are of the Spec-H-Comp type is also known as the SHCH, i.e. Specifier-Head-Complement Hypothesis (Roberts 2017a), and this must hold for the structure of any natural language.

By following Kayne (1994), any alternate linearization of S,V and O must be derived by means of movement (see also Whitman 2008: Sec. 10.6), while SVO can, but need not be (Biberauer, Holmberg & Roberts 2014: 207-208). Given that syntactic objects must move to c-commanding positions and that c-command maps precedence by being necessarily asymmetric, Kayne (1994: 47, 51) argues that LCA effectively bans rightward movement altogether while allowing its leftward counterpart. Thus, cases where the complement precedes its head, e.g. the OV order, are a consequence of leftward raising to a higher specifier po-

sition (Kayne 1994: 48). What is more, Kayne (1994: 51-54) argues that the derivation of the alternate orders can be formulated in two different ways that comply with antisymmetry, depending on whether one allows leftward head-to-head movement in the narrow syntax or not (see footnote 12 for discussion of theoretical problems linked to head movement).

With respect to the first of the two options to derive OV, Broekhuis (2010) provides a detailed word order typology that assumes independent spec-to-spec and head-to-head movements that comply with antisymmetry, as table 2.3 on the next page illustrates. This account (which will be taken up again in section 4.5.4) follows Larson (1988) and Chomsky (1995) and related works by splitting the lexical verb into a V head that selects the Object and a *v* head that takes VP as its complement and introduces the Subject in its specifier position.

The I head, already introduced in section 2.2.3, is also deemed to be universal as a consequence the Extended Projection Principle (EPP, first sketched in Chomsky 1981: Sec. 2.1). The latter is understood in the revized formulation proposed by Alexiadou & Anagnostopoulou (2001, 2007), according to whom either I or its specifier must be filled by a moved element. If CP were added on top of this structure of the clause, one would obtain the C-T-v-V system that is often adopted following Chomsky (1995). However, Broekhuis (2010) leaves CP out of the analysis (together with its cartographic expansion, Rizzi 1997) not only because it hosts elements related to the illocutionary force of the sentence, but also because this move sets aside Verb-Second phenomena that lead to misguided analyses of word order typology (as discussed in section 2.2.2).

In table 2.3, the greatest depature from the C-T-v-V system is represented by V-extP and *v*-extP, which are a reformulation of the Agreement projections proposed by Pollock (1989) and removed by Chomsky (1995: Ch. 4) due to their lack of semantic import. However, Broekhuis (2010: Sec. 4.1) agrees with G. Müller (2010: 37, 42, 2011: 128-130, 169) in claiming that it is legitimate to posit them because they affect the word order whenever they are part of the structure, i.e. they represent an effect on the output in the sense of Chomsky (2000: 109, 2001: 34, 2008: 149). In fact, V-extP and *v*-extP are absent in the derivations of table 2.3 where they are greyed out, because they are optional projections that are present only when it is necessary to derive an OV order. In all such cases, the head of the necessary ext-P(s) must be targeted by the raising verb and the respective specifier(s) must be filled by the Object.[21]

[21] The optionality of V-extP and *v*-extP also resonates with definitions (30) and (71C) in Fukui & Takano (2000: 236) and Borer (2013: 288), respectively. A similar claim is also made by Dékány (2021: 2-3) as

	Spec$_1$	**I**	**Spec$_2$**	***v*-ext**	**Subj**	***v***	**Spec$_3$**	**V-ext**	**V**	**Obj**
SVO	S	#			t_S	*v*			V	O
	S	#			t_S	V + *v*			t_V	O
	S	#			t_S	V + *v*	O	t_V	t_V	t_O
	S	V + *v* + I			t_S	t_{V+v}			t_V	O
	S	V + *v* + I			t_S	t_{V+v}	O	t_V	t_V	t_O
	S	V + *v* + I	O	t_{V+v}	t_S	t_{V+v}	t_O	t_V	t_V	t_O
SOV	S	#			t_S	*v*	O	V	t_V	t_O
	S	#	O	*v*	t_S	t_V	t_O	V	t_V	t_O
	S	#	O	V + *v*	t_S	t_{V+v}	t_O	t_V	t_V	t_O
VSO		V + *v* + I			S	t_{V+v}			t_V	O
		V + *v* + I			S	t_{V+v}	O	t_V	t_V	t_O
VOS		V + *v* + I	O	t_{V+v}	S	t_{V+v}	t_O	t_V	t_V	t_O

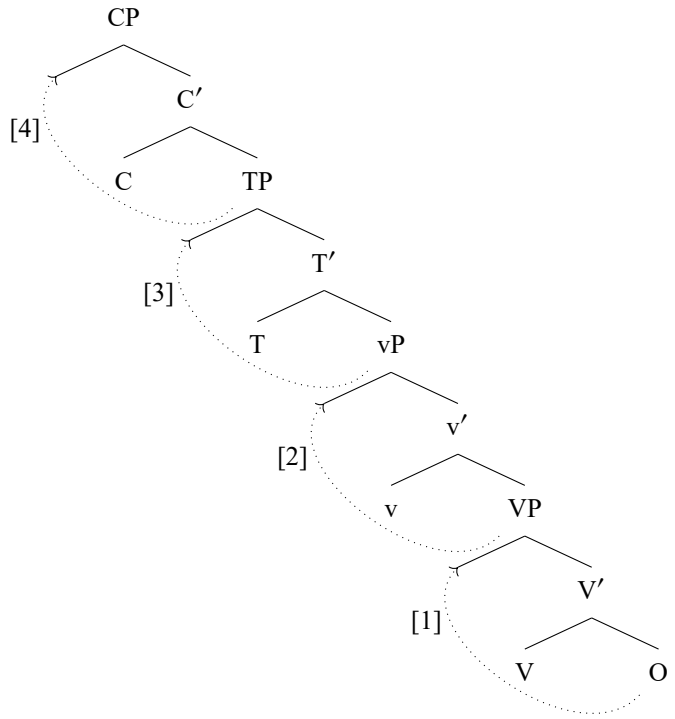
Table 2.3: Word order typology proposed by Broekhuis (2010: 17, 26).

The second possibility in which Kayne (1994: 52-54) proposes to derive OV orders is by making the complement of the verb move to the specifier of the projection in which the complement itself originates. If this kind of movement is iterated across projections, it gives rise to a consistently head-final order (see Ernst 2014: 118 ff.). Aboh (2004, 2010) calls it "snowballing" movement, while Biberauer, Holmberg & Roberts (2014) define it as "roll-up", as section 2.3.3 will further discuss.

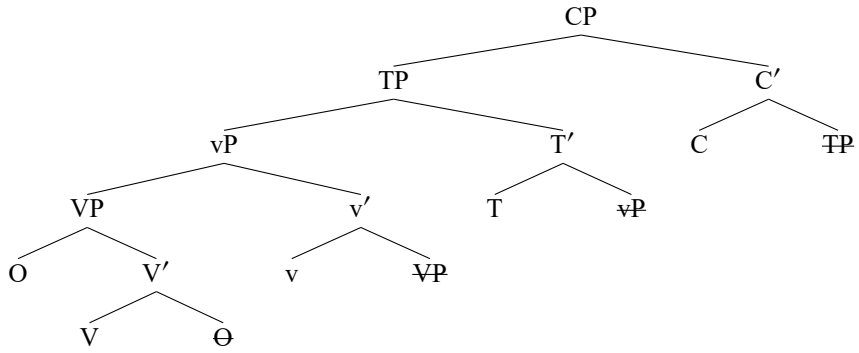
The tree in (28), which illustrates how a head-final clause can be derived, is adapted from Biberauer, Holmberg & Roberts (2014: 208), while (27) shows what the same structure would look like if the roll-up movements (numbered in compliance with the commonly adopted bottom-up derivation, Brattico & Chesi 2020: 2, Kobele 2023: Sec. 3) did not take place.

well as by Kim (2019: 128) (who follows Dimitrova-Vulchanova & Giusti 1998, Giusti 2002). Moreover, the same effect can arguably be obtained by extending to the low periphery (Belletti 2004) the *Feature Scattering Hypothesis* proposed by Giorgi & Pianesi (1997) (in this respect, see also Hsu 2017, 2021).

(27)



(28)



Although the two approaches to OV are not necessarily incompatible, head-to-head movement is not needed in the latter, because the verb is stranded as a result of phrasal movement.[22] As Kayne (1994: 53) observes, the roll-up movements in (27)-(28) provide an account for the tendency of OV languages to show agglutinative morphology. The account in (27)-(28) also predicts that head-final languages should avoid *wh*-fronting, because Spec,CP is already filled by TP (Kayne 1994: 54).

Exceptions aside (see Kayne 1994: 142-143 n. 22, Koopman 2000b and Toyoshima 2013

[22] Broekhuis (2010: 28 n. 9) actually claims that the derivation of VOS where the whole VP is fronted (as assumed in S. Chung 2006) confirms the rarity VOS as it is derived in table 2.3. Moreover, Shim (2021: 54 ff.) proposes to derive OV in Korean and Japanese by means of roll-up movements, but he follows Travis (2010) in assuming an Asp(ectual) P(rojection) whose position in the structure corresponds to V-ext in table 2.3. In this respect, it is worth noting that Travis (2010: 4) also proposes an Outer AspP, whose position would match *v*-ext in table 2.3. However, Travis (2010), Broekhuis (2010), and Shim (2021) make different assumptions with respect to both the types of movement taking place in the derivation and the features that trigger them.

for an account), the derivation in (27)-(28) presents a major problem, namely the lack of an available position for the subject of the clause. The discussion of this issue must be postponed to section 2.4.1, as it is more appropriate to first introduce the *Final-Over-Final Condition*, which attempts to account for the typological observations discussed so far by means of antisymmetry.

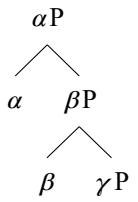
2.3.2 The Final-Over-Final Condition (FOFC)

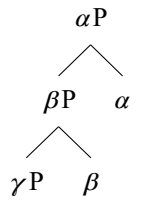
Holmberg (2017a: 1) gives two formulations in order to illustrate FOFC, the first one (29a) being the "informal" variant (see Biberauer, Holmberg & Roberts 2014: 171) and the second one (29b) its formal alternative:

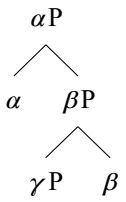
(29) a. A head-final phrase αP cannot immediately dominate a head-initial phrase βP, if α and β are members of the same extended projection.

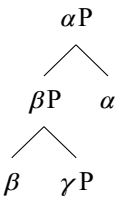
b. * [$_{\alpha P}$ [$_{\beta P}$ β γP] α], where β and γ are sisters and α and β are members of the same extended projection.

Put into a visual representation, FOFC is a syntactic principle that applies to the four structures in (30) by ruling out the structure (30d).[23]

(30) a. 

b. 

c. 

* d. 

The most recurrent argument used as an exemplification of FOFC is the rarity of the V-O-Aux order. This fact is indeed predicted by FOFC by assuming for the structures in (30) that Aux = α, V = β and O = γP, and it was first observed by Holmberg (2000: 128) with respect to the the intralingual variation of Finnish, as the examples in (31) illustrate (see Holmberg 2017b: 1 for different sentences showing the same fact):

[23] Diego G. Krivochen (p.c.) wonders why this constraint is not named after the structure it rules out, as it happens for instance for the *that-trace effect* (see Richards 2014: 4.6.1 and Dikken 2018: Sec. 5 for discussion, as well as section 5.4.4). DelBusso (2018: 129 fn. 2) takes up this objection and proposes FOIC as an alternative name, after "the uniquely banned final-over-initial structure". The focus on the final-over-final structure "abbreviates the implicational statement—final in higher only if final in lower". However, the discussion in section 2.2.1 shows that the latter formulation could also be stated by means of its contrapositive—initial in lower only if initial in higher. It should be noted that Roberts (2017a) acknowledges this fact, but rejects it on the theoretical ground that derivations and feature percolation take place from the bottom up.

(31) a. *Milloin Jussi olisi kirjoittanut romaanin?* [Aux-V-O]
when Jussi would.have written novel.INDF
b. *Milloin Jussi olisi romaanin kirjoittanut?* [Aux-O-V]
when Jussi would.have novel.INDF written
c. *Milloin Jussi romaanin kirjoittanut olisi?* [O-V-Aux]
when Jussi novel.INDF written would.have
d. **Milloin Jussi kirjoittanut romaanin olisi?* [*V-O-Aux]
when Jussi written novel.INDF would.have
'When would Jussi have written a novel?'

With respect to (30), Sheehan et al. (2017) further propose that the structure in (30c) is less likely than the other two possible syntactic configurations, i.e. (30a) and (30b). On the one hand, the better chance for the first two structures in (30) to occur is due to their structural consistency, because these syntactic configurations reflect the empirical observations that can be traced back to the notion of *harmony*, discussed in section 2.2.1.

Besides, if the head-initial phrase αP and the head-final βP assumed by the definitions of FOFC on page 54 are interpreted as dominant constructions in Greenberg's sense (as it is done by Roberts 2017a: 29-30), then FOFC can be seen as a fruitful combination of "both the Greenbergian and the Chomskyan traditions" (Biberauer, Holmberg & Roberts 2014: 170). In fact, FOFC is a single syntactic principle that is supposed to capture a broad variety of phenomena, including a ban on the possible orders related to Greenberg's Universal 20 (see section 3.2.3).

Given that FOFC applies only within a single extended projection (which was summarized in section 2.2.3 following Grimshaw 1991, 2005), FOFC-violating configurations are allowed as soon as two different extended projections are involved. Differently put, the application of FOFC is restricted to those cases where α and β in (30) are both either verbal or nominal heads, while γ can be the outer functional head of a different extended projection. Biberauer, Holmberg & Roberts (2014: 197) use some data from German to show that this prediction is borne out by O-V and PP-V orders, where the object DP and PP are both head-initial, non-verbal extended projections contained in a head-final VP:

(32) a. *Johann hat* [$_{VP}$ [$_{DP}$ *einen Mann*] *gesehen*]
Johann has INDF.M.ACC man seen

'Johann has seen a man.'

b. *Johann ist* [$_{VP}$ [$_{PP}$ *nach Berlin*] *gefahren*]
Johann is to Berlin gone
'Johann has gone to Berlin.'

This criterion can also be applied to account for V-O-Aux attestations found in three-member verbal clusters in West Flemish, first analyzed by Haegeman & Riemsdijk (1986) as evidence of Verb Projection Raising (whose definition is provided in Haegeman & Riemsdijk 1986: 418). The relevant case for the present discussion is exemplified in (33) following Biberauer, Holmberg & Roberts (2014: 203). According to Biberauer, Holmberg & Roberts (2014: Sec. 3.3), this is not a true FOFC violation because V-O is part of a nominalized projection together with the preceding modal verb, as the *infinitivus pro participio* shown by the latter arguably demonstrates (see Biberauer, Holmberg & Roberts 2014: Sec. 3.3 for further discussion).

(33) *...da Valère willen dienen boek lezen eet* [West Flemish]
COMP Valère want.INF DEM.DIST book read.INF has
'... that Valère has wanted to read that book.'

However, German also gives cases that defy the predictions above, as a complement CP and the head-final VP selecting it obey FOFC even though they belong to two different extended projections. The following illustrative examples are taken from Biberauer, Holmberg & Roberts (2014: 198):[24]

(34) a. *... *dass Johann niemals* [$_{CP}$ *dass er eigentlich ein*
COMP Johann never COMP 3SG.M.NOM actually INDF.N.ACC
angenommenes Kind sei] *besprochen hat*
adopted.N.ACC child be.SBJV discuss.PTCP AUX

b. ... *dass Johann niemals besprochen hat,* [$_{CP}$ *dass er eigentlich*
COMP Johann never discuss.PTCP AUX COMP 3SG.M.NOM actually
ein angenommenes Kind sei]
INDF.N.ACC adopted.N.ACC child be.SBJV

[24] Haider (2010: 189) argues that the sentence in (34a) is highly marked, rather than totally ungrammatical (therefore, it should be judged with a ?? rather than with a *). See section 2.4.2 for further discussion on this point.

'... that Johann has never discussed [the fact] that he is actually an adopted child.'

In order to capture such a difference with respect to FOFC-compliance, Biberauer, Holmberg & Roberts (2014: 211) propose that FOFC also applies across multiple extended projections, if they are part of a single *spine*. An extended projection EP_1 forms a spine with the extended projection EP_2 if the lexical head of EP_1 c-selects EP_2 and if EP_1 and EP_2 share the same categorial value.[25] Put differently, this amendment to the definitions of FOFC given on page 54 boils down to saying that FOFC applies under categorial identity (see Hawkins 2013: 391-392 and Biberauer, Holmberg & Roberts 2014: 197).

Although it captures the data previously presented in (32) and (34), the reformulation provided in Biberauer, Holmberg & Roberts (2014: 211) becomes problematic because Sheehan et al. (2017) go on saying that FOFC applies to extended projections, thus making the very notion ambiguous with respect to its actual breadth, which can coincide with what Grimshaw (1991, 2005) assumes (namely that there can be one single lexical head per extended projection), or span across categorially-identical extended projections.[26]

Another problem with respect to the analysis that Biberauer, Holmberg & Roberts (2014) propose for the German facts in (34) is that the *ex situ* nature of complement CPs is not independently motivated. Such evidence is however provided by Duguine (2022), who refines their account by extending the analysis to Basque. Although Basque complement clauses appear postverbally in a manner akin to the German case in (34), there is a specific subclass that appears in the *in situ* preverbal position. The two cases are exemplified by (35a) and (35b), respectively.

(35) Basque, all examples from Duguine (2022: 19).[27]

a. *Mirenek ez du galdetu* [*ea Ana etorri d-en*] [Extraposed *ea* clause]
Miren NEG AUX ask Q Ana come AUX-COMP

b. *Mirenek ez du* [*Ana etorri d-en*] *galdetu* [*In situ* indirect interr.]
Miren NEG AUX Ana come AUX-COMP ask

[25] See also the formal definitions provided by Etxepare & Haddican (2017: 138-139), but note that they do not emphasize the permissibility of multiple lexical heads in a single spine.

[26] See Philip (2013: 166, 203) and Hucklebridge (2023: 8 fn. 7) for a similar remark and Zeijlstra (2023: 184-185, 199) for a discussion that falls victim to this understated ambiguity.

[27] In the examples in (35) the auxiliary *du* precedes the main verb as it raises together with the negative marker *ez* to ΣP, a phrase first posited by Laka (1990) to account for syntactic facts of Basque that are affected by sentence polarity. See Duguine (2022: 10-11, 18) for further discussion.

c. ?? *Mirenek ez du* [*ea Ana etorri d-en*] *galdetu* [*In situ ea* clause]
Miren NEG AUX Q Ana come AUX-COMP ask
'Miren didn't ask whether Ana came.'

Duguine (2022) shows that the crucial property that distinguishes the complement clause in (35b) is that it lacks any overt realization of heads in head-initial projections of the Basque left periphery (articulated along the cartographic lines of Rizzi 1997, Rizzi & Cinque 2016, where CP is assumed to be split into a finer hierarchy of subprojections) that would lead to a FOFC violation with respect to the selecting head-final verb. Thus, what makes the *in situ* complement clause in (35c) degraded with respect to its extraposed counterpart in (35a) is the overt realization of the interrogative particle *ea*, which Duguine (2022: 13) analyzes (in line with the references she cites) as a left-peripheral illocutionary force marker.

These facts of Basque lead to a twofold observation. On the one hand, the majority of Basque complement clauses parallels the German facts observed by Biberauer, Holmberg & Roberts (2014: 198 ff.) because the linearization of a verb selecting a head-initial (subprojection of) CP as its complement would violate FOFC otherwise. On the other hand, the fact that under specific circumstances complement clauses can appear in the assumed *in situ* preverbal position provides an argument for the extraposed nature of the postnominal counterparts, which are thus not merged to the right of the main verb as a default option.

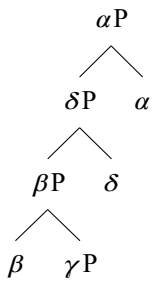
According to Biberauer (2017b) and Duguine (2022), it is the *ex situ* nature of a projection that makes it FOFC compliant, so that the cases of extraposition just discussed can be put on a par with A′-movement. As a consequence, linear orders that are ruled out by FOFC can appear after the latter type of movement takes place, because FOFC is still not violated structurally. One example can be adapted from Haider (2013: 80) in (36), where the V-O-Aux order emerges in colloquial German whenever VP is fronted in main clauses.

(36) [$_{CP}$ [$_{VP}$ *gesprochen mit ihr*] *hat*-C [$_{TP}$ *er* *nicht mehr* t_{VP}]]
spoken with 3SG.F.DAT AUX 3SG.M.NOM NEG more
'As for speaking with her, he no longer did that.'

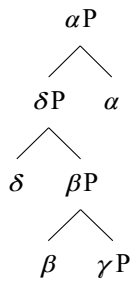
As far as amendments to the defitions of FOFC on page 54 are concerned, Duguine (2022: 21 fn. 28) further observes that the recourse to *immediate* dominance in (29a) can be dispensed with. In fact, if an additional functional δP is inserted in an intervening position between αP and βP in the FOFC-violating configuration (30d) on page 54, then δP itself will also violate

FOFC, in relation to either αP or βP (see also DelBusso 2018: Sec. 4.2 for discussion). This is further illustrated with the two structures in (37), where the trees given by Duguine (2022: 21 fn. 28) are modified so as make them consistent with (30).

(37) * a.



* b.



By witnessing what is not admitted as per the extended projections in (37), another case where FOFC has been advocated can be illustrated, namely the syntactic evolution of languages in diachrony (Biberauer 2017a: 180). If at a certain point in time t_0 an extended projection in a language L is made of consistently head-final heads α, δ, β (where $\alpha > \delta > \beta$), then the only syntactic change that can occur at a successive point in time t_1 is the head-initial configuration of αP. As the trees in (37) show, neither βP nor δP can switch in this way at t_1, because they would violate FOFC.

Hence, syntactic change in a consistently head-final extended projection needs to develop from the top down. This means that if αP becomes head-initial at t_1, then δP can do the same at t_2 and finally βP at t_3. This is exactly what Ledgeway (2012a,b) proposes for the syntactic evolution of the Latin clause structure into the subsequent Romance languages, where α = C, δ = T, and β = V. If on the contrary α, δ, and β are consistently head-initial at t_0, then the process needs to proceed from the bottom up (as Biberauer 2017a: 180 further shows), namely β switches to a head-final configuration at t_1, ditto for δ at t_2 and α at t_3.

The refinements heretofore discussed with respect to FOFC still fail to cover certain challenging counterexamples. In line with the "Pāṇinian approach" (Bickel 2015: Sec. 37.1) discussed in section 2.1, seemingly FOFC-violating cases can either prove the generalization wrong or be explained away as fringe cases that require some special treatment. On the one hand, the latter attempt leads to case-specific extentions of the principles sketched thus far, but on the other hand it exposes FOFC to the risk of being immune to falsification, as highlighted by Leivada (2020: 368).

Clem (2022: Sec. 3) summarizes the refinements that Biberauer (2017b) proposes along this vein, which can be divided into three cases. The first has already been illustrated in (36)

with respect to V-O-Aux in German (see Biberauer 2017b: 245 and Duguine 2022: 4 fn. 2). If a head-initial projection raises to an A′-position, it can be linearly adjacent and preceding another head usually dominating it in the same extended projection.

The second case occurs whenever the syntactic object that causes the FOFC violation, i.e. a potential β in (37a) or α in (37b), is demonstrably a phrase, not a head, while the third case hinges on the categorial identity that FOFC requires in order to apply (Hawkins 2013: 391-392 and Biberauer, Holmberg & Roberts 2014: 197). In fact, seemingly FOFC-violating particles merged along an extended projection might be unspecified for its category or lack formal features altogether (see Biberauer 2017b: 283). This claim is theoretically motivated by the availability of non-projecting heads in Bare Phrase Structure (Chomsky 1995) (see also Bruening 2009: 31) and empirically supported by their non-inflecting nature (Dryer 2007).

2.3.3 FOFC as a consequence of antisymmetry

A comparison of the X-bar schemata section 2.3 on page 47 with the structures captured by FOFC in section 2.3.2 shows that the former was not able to restrict the options related to the placement of a head with respect to the maximal projections it selects, while the latter assumes that the directionality of heads within an extended projection is not free. Roberts (2017a) (and Sheehan et al. 2017 more generally) claim that Kayne's antisymmetric approach discussed in section 2.3.1 not only restricts the possible phrasal templates, but it can also capture FOFC. This line of analysis dates back to Biberauer, Holmberg & Roberts (2014: 290 ff.), where an antisymmetric view of FOFC is adopted by assuming that head directionality is narrow-syntactically encoded by means of a privative ^ (*caret*) feature which can be passed on from V to an uninterrupted number of functional heads of its extended projection, as per the formal definition in Biberauer, Holmberg & Roberts (2014: 210):

(38) If a head α_i in the extended projection EP of a lexical head L, EP(L), has ^ associated with its [±V]-feature, then so does α_{i+1}, where α_{i+1} is c-selected by α_i in EP(L).

What the ^ feature does to the (portion of the) extended projection(s) it affects is triggering Comp-to-Spec "roll-up" movements of the base-generated Spec-Head-Comp order.[28] The linear precedence relation Comp ≻ Head that gives rise to the seemingly head-final projection

[28] The ^ feature is thus a pure movement feature with no semantic correlate, just like EPP and edge features. See Bidese (2023: 57).

is thus accounted for in antisymmetric terms. Crucially, the statement in (38) prevents a functional head that is part of a certain extended projection to introduce ^, which must be always detectable on the lexical head first and be passed on to dominating functional heads.

Put into the words of Etxepare & Haddican (2017: 140), (38) states that the sequence of roll-up movements "proceeds monotonically, that is, it cannot start and stop and start again". This principle allows to make a twofold account. On the one hand, it effectively rules out the FOFC-violating cases from Duguine (2022), shown in (37) on page 59, because it makes impossible for a head-initial projection, which lacks ^, to be dominated by a head-final one. In fact, the latter must have inherited ^ from the former, which must consequently be head-final as well. On the other hand, (38) allows one single "mixed" extended projection in terms of directionality (see Elordieta 2013, Etxepare & Haddican 2017, a.o.), namely an uninterrupted sequence of head-initial projections dominating a lower portion of the spine made of head-final projections.[29]

This explanation of FOFC makes contradictory predictions with respect to *spines* with mixed directionality.[30] One case in point is represented by the Basque clauses discussed by Duguine (2022) and previously reported on page 57. By following both the SHCH of Kayne (1994) and the left periphery assumed by Duguine (2022), one must infer that a Basque complement clause externally merges with the matrix verb selecting it in what superficially appears to be the extraposed position. The apparently *in situ* counterparts are actually derived via the movement triggered by ^, which makes them raise to Spec,VP. The problem, however, emerges with complement clauses that must remain to the right of the matrix verb because of the overt heads they realize in their left-peripheral head-initial projections. In order to account for this behavior, contradictory claims must be made.

Since any Basque verbal extended projection is assumed to be superficially head-final up to either TP (as in Etxepare & Haddican 2017) or FinP (as in Duguine 2022), the matrix verb must not lack ^. However, since this feature is assumed to trigger a roll-up movement whenever present, complement clauses first-merged in the desired postverbal position cannot stay there either. This contradiction can only be fixed by positing the *ad hoc* assumption that the mandatory Comp-to-Spec movement within VP is exceptionally violated if the extended

[29] The portion of the extended projection where the switch in directionality occurs boils down to the structure (30c) on page 54.

[30] This is a problem which (to the best of my knowledge) has not been explicitly addressed before. See however Haider (2010: Sec. 5.3-4) and Philip (2013: Sec. 5.1) for observations that are closely related to the issue raised in the main text.

projection that V selects is also verbal and thus forms a spine together with the main clause.

The discussion above shows that (independently of the further issues to be presented in section 2.4.1) the roll-up approach to FOFC raises several problems which make it an unsuitable explanation of the phenomenon. Two competing alternatives have been suggested in the literature. According to the first (see Etxepare & Haddican 2017, Duguine 2022, a.o.), FOFC may play no role in narrow syntax. FOFC-violating structures can thus be built in the narrow syntax and then "repaired" as an interface condition in the Phonological Form (PF). As a consequence, FOFC is seen not a hierarchical generalization, but rather as a derivational one in the senses of Whitman (2008: 234), and it is likely to be obeyed due to extragrammatical factors (see section 2.4 and references therein). This view does not necessarily imply a rejection of antisymmetry (Kayne 1994), as the requirement of the latter can be seen as interacting with FOFC in ways that partly retain the insights of both principles (as in Sheehan 2013, Etxepare & Haddican 2017).

The second alternative to the roll-up derivation of FOFC would be a narrow-syntactic restatement which gives rise to the relevant linearization in a different way. One such manner is suggested by Biberauer, Holmberg & Roberts (2014: Sec. 4.3.2), where ^ corresponds to "a PF instruction to linearize the head to the right of its sister". Although this may sound similar to the traditional formulation of the head parameter provided in (25) on page 48, it is different from it because the "direct linearization" encoded in the narrow syntax would not be relevant there, but rather satisfied at PF, where it would interact with the other strategies that Etxepare & Haddican (2017) and Duguine (2022) assume to take place at PF to repair FOFC violations.

With respect to the issue caused by roll-up movements in the Basque spines discussed on page 57, this approach would work in a less problematic manner. The head-initial complement clause would be in a sisterhood relation with the matrix verb, and this would not cause any problem until the structure is mapped at PF. At the interface, the matrix verb is realized as head-final and the complement clauses that would be FOFC-violating undergoes extraposition or A′-movement accordingly, as it is intuitively suggested by the description of the relevant the German data, which have been reported in section 2.3.2 following Biberauer, Holmberg & Roberts (2014) and Duguine (2022).

2.4 Approaches that reject antisymmetry

The consistent Specifier-Head-Complement Hypothesis proposed by Kayne (1994) is not borne out by typological observations. In fact, through a "rapid look at (a small subset of) the world's (presently existing) languages" Kayne (1994: 35) is led to assert that SVO order, being the linear order mapped by a Spec-Head-Comp base-generated structure, must be the most likely crosslinguistically. However, this contradicts "[t]he literature in language typology [...] which [...] gives little credence to the idea that [that] order predominates", as Newmeyer (1998: 360) notes. In particular, it is hard to believe SOV to be derived via syntactic movement from SVO, especially if one considers that the former is more likely than the latter on a crosslinguistic level. Antisymmetry is thus at odds with the "Rarity-Complexity Hypothesis", according to which "[t]he rarer the grammatical phenomenon, the more complex its analysis" (Newmeyer 2010: 209).

Guimarães (2008) and Abels & Neeleman (2012) show that the strong generative power (Joshi 2003: 243) of an antisymmetric approach is not as restricted as Kayne assumes on formal grounds either. This is mainly due to the fact that the LCA does not say anything with respect to the way in which labeling is determined in the structure. Thus, if projection operates in a different way from what Kayne (1994) tacitly assumes or if it is dictated by a separated mechanism (e.g., the Labeling Algorithm in Chomsky 2013, 2015), several structures that antisymmetry tries to rule out could be generated in a way which would not violate it, because they would still obey the relations of asymmetric c-command required for the correct application of the LCA. In fact, Guimarães (2008) shows that *n*-ary branching structures and multiple specifiers can be couched in a structure that is LCA-compliant if the unstated premises are reworked accordingly.

Antisymmetric approaches also face another difficulty which is pointed out by Abels (2007). The structures that Kayne (1994) tries to rule out can be derived in a LCA-compliant manner by means of the merger in the structure of silent functional projections. As will be further discussed in section 2.4.1, the latter are indeed posited by many antisymmetric analyses to capture the empirical data by means of leftward movements that reorder the constituents as desired. However, given enough projections as such, an antisymmetric analysis can end up contradicting its own theoretical premises and/or generate structures that are equivalent to structurally simpler counterparts that do not comply with the LCA (see also Ackema & Neeleman 2002: Sec. 3, Haider 2013: 216 and Bruening 2022: Sec. 2.9). This aspect will be

highly relevant in section 3.2.2, where the analysis of Greenberg's Universal 20 offered by Cinque (2023) will be discussed.

As sections 2.3.2 and 2.3.3 showed, scholars following the antisymmetric proposal formulated by Kayne (1994) do so in order to account for empirically observable asymmetries (many of which have been originally pointed out by typologists) by means of theoretically principled grounds. Because of this, many scholars argue that what Kayne (1994) proposes should not be rejected altogether, but rather reworked so as to make it in agreement with other widespread assumptions. For example, the idea that relevant facts of word order are directly mapped from the syntactic structure can be conflicting with the Minimalist claim that linear order is relevant only to the Phonological Form. However, Moro (2000) and Moro & Roberts (2024) argue that such a matter can be settled as soon as one allows theoretically circumstantiated exceptions to the LCA in the narrow syntax.

In fact, the formation of symmetric structures is not problematic if they are subsequently broken in the course of the narrow-syntactic derivation and no longer relevant at the phonological interface. As Broekhuis (2008: 75) puts it, the LCA "is essentially a PF-principle", so that it neither has to apply at every stage of the derivation nor limit the number of adjuncts per phrase to one. Although section 2.3.3 has already illustrated that there are proposals trying to take advantage of antisymmetric structures together with PF-requirements (of which FOFC is regarded as a relevant one), section 2.4.1 will illustrate works where only the latter are kept, by grounding them on extra-grammatical factors ("third factor principles" as per Chomsky 2005) such as the limitations of the human parser (Ackema & Neeleman 2002).

A rather different take on these issues is given by those works which argue that the decisive structural properties that affect the observable word order are not the ones posited by antisymmetry. Bruening (2022: Sec. 2.9), whose claims have been discussed in section 2.2.4 and will be taken up again in sections 3.1.4 and 3.2.5, agrees with Kayne (1994) in arguing that word order should be determined in the narrow syntax, but he posits many alternate base-generated structures that allow multiple adjuncts, right-branching structures and rightward movement, which are all banned by antisymmetry. Section 2.4.2 will illustrate other works that share with Kayne (1994) the idea that there is one single base-generated structure, although they counterargue that it must be endowed with different structural properties in order to make predictions that are claimed to be both broader and more adequate.

2.4.1 Symmetric base-generated structures

Within Minimalism, base-generated structures are symmetric because the fundamental syntactic operation, i.e. Merge, has been conceived since Chomsky (1995) as a set-theoretic combination of two syntactic objects, so that a new unordered one is formed. Although such a fundamental syntactic operation has undergone several revisions with respect to both its exact name and the way in which it operates (for an overview, see Milway 2022: Sec. 3.1-3.2 and Kobele 2023: Sec. 4), the idea that it can give rise to symmetric base-generated structures has remained constant throughout.[31]

Given that "the narrow syntax [...] only recognizes hierarchy" (Holmberg 2017b: 48), projections (intended as a shorthand for "proxies for sets" or for "topological graphs", Krivochen 2023d: 320) bearing the same semantic interpretation can become head-initial as well as head-final after Spell-Out (see section 2.3.1), and specifiers are equally likely to become left or right branches at PF, *contra* the assumptions made by Kayne (1994). As the discussion of Broekhuis (2010) in section 2.3.1 has shown, since Chomsky (1995) a change in word order is narrow syntactic only if it is semantically determined, whereas "meaningless movement" (Hall 2015, Cinque 2023: Sec. 5.4) is a consequence of parameterization at PF. Most of the typologically relevant left-right asymmetries concerning linear order thus need to be captured by means of some restriction that is akin to the LCA in the spirit, but that does not apply in the narrow syntax (Chomsky 2014–1995: 313).

Ackema & Neeleman (2002) propose that the limitations of the human parser represent a good candidate to explain such asymmetries with respect to the ban on rightward movement. Abels & Neeleman (2012: 68) clarify why this should be the case, as follows:

> [T]he parser needs to recover hierarchical structures from input strings that are presented to it incrementally. In other words, there is an inherent asymmetry in the parsing process: bits of the input presented to the parser earlier on are associated with a structure before substrings presented later. [...] It follows from this that leftward and rightward movement are different with respect to parsing. [...] Under certain circumstances th[e latter] is hard or impossible[.]

Based on these extra-grammatical premises, Ackema & Neeleman (2002: Sec. 9) argue

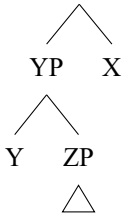
[31] An exception is represented by Pair-Merge, which creates ordered pairs in the narrow syntax. It has been proposed in the literature to account for the behavior of adjuncts, linkers and clitics. See Manzini (2021), Manzini & Pescarini (2022), Milway (2022) and references therein for discussion.

that there are two kinds of asymmetries that affect the otherwise symmetric base-generated structures. On the one hand, “head movement to the right seems to be impossible when the head crosses its own dependents”, whereas no such problem arises for the leftward counterpart. On the other hand, “XP-movement to the right is extremely local (namely phrase-bound), while for XP-movement to the left this is not necessarily the case”.

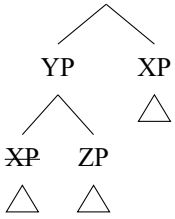
The local nature of the right extraposition of an XP corresponds to the definition that Cecchetto (2013: Sec. 5.5) eventually establishes for the Right Roof Constraint, i.e. a principle first formulated by Ross (1967) to capture the asymmetric behavior between the leftward and the rightward movement of certain phrases (mainly clausal projections, see Cecchetto 2013: Sec. 4-5 and references threin). What Cecchetto (2013) further observes is that the extra-grammatical principles that are assumed to foster this syntactic phenomenon can be easily extended to provide an account for FOFC that does not hinge on antisymmetry.

To highlight the conceptual similarity between FOFC and the Right Roof Constraint, Cecchetto (2013: Sec. 4) gives a pair of structures that are similar to those shown in (39). In (39a), the selecting category follows the selected one, whereas in (39b) the higher copy (or trace, depending on one’s assumptions) of a phrase follows the lower one. Thus, both structures in (39) are backward dependencies (i.e., selection in the former and movement in the latter) that illegally cross phrase boundaries (Clem 2022: 819).

(39) a.



b.



Based on (39a), the typological rarity of a final head dominating an initial one within the same extended projection (or spine, as per Biberauer, Holmberg & Roberts 2014: 211) depends on the fact that the former cannot select the latter directly, but must rather be in a selectional relationship with “the entire complement phrase, [...] such as when a final verb selects a head-initial DP complement” (Clem 2022: 819). Cecchetto (2013: Sec. 9) acknowledges that this principle should also hold if YP in (39a) were head-final, and defends this potential weakness of his account by arguing that the indirect selectional relationship among head-final heads is what gives rise to the often observed “agglutinative-like morphology” in verb-final

languages, which are then predicted to resort to it in order "to avoid illicit [...] inverted dependencies". In this respect, it is worth recalling something that was pointed out in section 2.3.1, namely that Kayne (1994: 53) provided the same argument in support of antisymmetry.

Still, what has been discussed so far following Ackema & Neeleman (2002), Cecchetto (2013) and related works does not provide a direct answer for another problem that a rejection of antisymmetry raises: "Under a symmetric perspective, one would expect an even balance between SV and VS languages" (Zeijlstra 2023: 206), which is not what the typological data suggest. To explain the latter asymmetry, Zeijlstra (2023: Sec. 5) starts from an account for FOFC that can find empirical overlappings with what Cecchetto (2013) proposes (see Clem 2022: 820 for an example). In his view, FOFC descends from the fact that string-adjacency is imposed to those heads that form a unit at PF, but that impossible head movements to the right limit the manners in which this condition can be satisfied.

On the one hand, no movement is needed for heads that are "string-adjacent in the first place" (Zeijlstra 2023: 197), and this is why consistently head-final (or head-initial) heads are not problematic. Although head movement can potentially fix the cases where the relevant heads are not string-adjacent, it can only take place to the left. Thus, a configuration like (39a) is doomed if X and Y are expected to form a unit at PF, while it can be an exception to FOFC if there is never a requirement for Y to end up in X (see again Zeijlstra 2023: 197).

Based on such an account for FOFC, the high frequency of SV orders as opposed to VS follows if subjects have to raise to a projection like TP. Given that such a movement is not phrase-bound, it cannot be an instance of rightward movement. Moreover, Zeijlstra (2023: 209) points out that even the base position of an external argument can be constrained in the same fashion as soon as it is the target in processes like unaccusativity and passivization. The conclusion is therefore that right subjects can only appear if the specifier hosting them is never a movement target, i.e. "there is no movement whatsoever in [that] particular specifier position" (Zeijlstra 2023: 209).

The fact that the leftward movement of a subject cannot be phrase-bound provides the cue to observe that in antisymmetric approaches this matter does not come without problems either, contrary to what one would first argue. In fact, the roll-up movements shown on page 53 following Biberauer, Holmberg & Roberts (2014: 208) have already been criticized for the fact that, if no projection is present in the structure beyond the C-T-v-V heads adopted by Chomsky (1995), then there is no available position for the external argument. More-

over, those Comp-to-Spec movements neatly exemplify one of the most debated aspects of antisymmetry. As Zeijlstra (2023: 187) puts it, "whatever the property on the head is that requires the presence of such a constituent in the specifier position is already satisfied by the complement in the sister position".

Thus, many scholars do not welcome letward Comp-to-Spec movements, as they hinge on the intuition that "movement is not allowed to be too short" (Erlewine 2020). Still, it must be pointed out that such a concept, better known as *anti-locality* (see Grohmann 2012 for an overview and Erlewine 2020: Sec. 6, Zeijlstra 2023: 183 for further references), does not appear to apply to the phrase-bound rightward movements discussed in Ackema & Neeleman (2002: Sec. 9) and Cecchetto (2013), a fact which will pointed out again in section 3.2.5.

Regardless of what their position on anti-locality is, many antisymmetric works tweak the consistency of the Comp-to-Spec movements shown on page 53 following Biberauer, Holmberg & Roberts (2014: 208) because of the need of free specifiers where phrases (like external arguments) can be first-merged (or remerged). A work to be discussed in section 3.2.3, i.e. Roberts (2017b), circumvents the problem by assuming for instance multiple specifiers per projection (see Shlonsky & Bocci 2019: Sec. 7 for related discussion), so that Comp-to-Spec is salvaged at the expense of consistency with both X-bar theory and Kayne (1994), who argues that the LCA imposes at most one specifier per phrase.

The complement YP that is a sister of a certain head X can also circumvent the aforementioned problems if it targets Spec,ZP, where ZP is a projection that dominates XP. However, the projections that are also adopted by non-antisymmetric works are often needed for mergers (or re-mergers) other than roll-up, so that it is often the case that they do not represent a good alternative in this sense. The most common solution, then, is to assume the presence of silent functional elements, whose specifiers are free to host roll-up movement. The assumption of the latter projections, which are pervasive in the analyses of U20 provided by Cinque (2005, 2023) and Steddy & Samek-Lodovici (2011) and to be discussed in sections 3.2.1 and 3.2.2, is often adopted at the cost of weakening the empirical content of antisymmetry. As shown by Abels (2007: Sec. 1), given enough (additional, functional) structure, these refinements can account for the very patterns that were to be excluded as per Kayne (1994).

2.4.2 SOV as the base-generated order

Not all works that do not adopt antisymmetry as proposed by Kayne (1994) actually support the opposing view where it is rejected. In fact, some scholars believe that the entirely post-syntactic determination of linear order is not as satisfactory as it seems. For a start, at least two theoretical considerations are deemed to disfavor it. The first is provided by Chesi (2024). Given that the goal of a theory is to generate and recognize all and only the sentences of a certain language L (Chesi 2024: 5), the structures that a formalism derives cannot be random, but rather limited to those that are "relevant descriptions for the sentences in our language L" (Chesi 2024: 9). Hence, Chesi (2024: 13-14) argues that the excessive computational efficiency provided by Merge is no longer helpful if descriptive adequacy is not met. And in fact, a crucial ingredient that allows Stabler (1997: Sec. 1.3) to formulate a testable formalization of Minimalism is the reintegration of linear order as part of narrow syntax (see again Chesi 2024: 13-14 for discussion).

The second theoretical problem with the relegation of linear order outside of narrow syntax is pointed out by Haider (2013: 211-212) and Bruening (2022: 27). Granted that syntax is a mapping algorithm, those who follow Kayne (1994) and Chomsky (1995) in arguing that linear order is mapped from syntax to PF are *de facto* arguing that there is "a second syntax after the first", doing "the job the first syntax should have done" (Bruening 2022: 27). In particular, Haider (2013: 211-212) observes that the first syntax maps structure on structure, as (sections 5.1.5 and 5.3.2 will discuss that) it needs numerations and/or workspaces to generate a series of representations, whereas the second linearizes the latter. However, the presence of two analogous devices is theoretically unappealing (Bruening 2022: 27; see also Brody 2018), and only structure-to-string mapping algorithms are empirically motivated (Haider 2013: 212).

Other reasons that cast doubt on the entirely post-syntactic nature of linear order are empirical in nature, and are given by Richards (2016: 168):

> Headedness in a projection correlates reliably with the placement and realization of phrasal stress; in particular, [...] main prominence is generally attracted to the nonhead in a head-complement pair. [... Moreover,] infants make use of prosodic cues to determine head direction [...]. In Contiguity Theory, facts like these could be embodied in a parameter that has consequences for the syntax, governing where prosodic prominence is best realized.

Contiguity Theory is the theoretical approach proposed by Richards (2016: 2) himself, where it is argued that "the building of certain aspects of phonological structure in fact begins in the narrow syntax", so that what needs to be assessed is which phonological aspects are truly relegated at PF and which are instead syntactically relevant (see also Primus 2001: 858 for discussion of related issues in this respect).

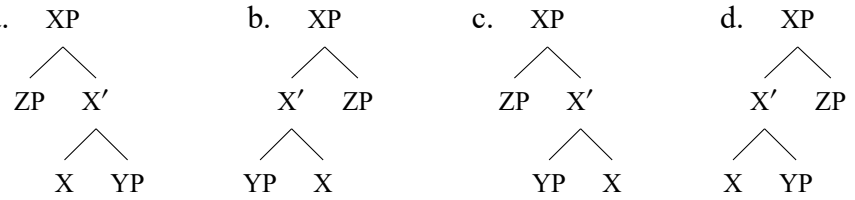
Without going into the details of this proposal, what is worth mentioning here is that Richards (2016: 6) adopts the LCA formulated by Kayne (1994) to establish the Specifier-Head-Complement order accordingly. What changes is the manner in which deviations from the LCA are accounted for. Richards (2016: 23) proposes *Untethering*, a principle which applies to a "head and its complement just after the head has been merged" by deleting all the ordering statements that would follow from the fact that the head asymmetrically c-commands the interior of its complement, so that the two sisters become eventually unordered with respect to each other.[32]

Fukui & Takano (1998, 2000) represent potential forerunners of Untethering, as they propose that linear order is obtained by undoing what the syntax builds. However, the process of structure decomposition that they assume does not take place during the ongoing narrow-syntacic derivation as in Richards (2016), but rather after Spell-Out. As Fukui & Takano (2000: 227) summarize, "pre-Spell-Out computations in syntax and post-Spell-Out computations in the phonological component form mirror images of each other". This is due to the fact that syntax builds the structure from the bottom up by means of Merge, while the closest conceptual analog of Untethering, i.e. *Demerge*, operates after Spell-Out in a top-down fashion, by applying exclusively to maximal projections like Merge does before Spell-Out (Fukui & Takano 2000: 225).

By illustrating their idea in X-bar theoretic terms, Fukui & Takano (1998, 2000) argue that their proposal is consistent with the Minimalist view where linear order plays no role in narrow syntax. In fact, the post-Spell-Out operations that they assume would lead to the same linearization for any of the four X-bar alternatives previously shown on page 47, and repeated in (40) for convenience.

[32] Thus, it becomes interesting to confront Untethering with the account for FOFC provided by Biberauer, Holmberg & Roberts (2014: Sec. 4.3.2) and discussed at the end of section 2.3.3. The latter argue that narrow syntax inserts a movement-triggering feature to those projections which will be mapped as head-final at PF, whereas Richards (2016: 23) argues that the same outcome is obtained in the narrow syntax by applying a phonologically-motivated cancellation of the asymmetric c-command relations that the LCA establishes. See Richards (2016: Sec. 5.5) for further discussion.

(40)



Starting from the root, the first maximal projection that Demerge encounters in any of the structures in (40) is always ZP, while X′ becomes a maximal projection only once Demerge decomposes one of the structures in (40) into two distinct objects, namely ZP and a structure rooted by X′. It is due to such an asymmetry between ZP and X′ that the operation that assigns linear order between the two, i.e. *Concatenate*, makes ZP precede X′ in the eventual linear order (Fukui & Takano 1998: 41). By virtue of these first steps of the linearization process, Fukui & Takano (1998: 74) argue that their proposals manages to account for the highly frequent SV order in a straighforward way.

By inspecting the content within X′, the first maximal projection that Demerge encounters is YP, while X becomes a maximal projection only once it becomes a syntactic objects of its own. In this respect, the currently discussed approach diverges from antisymmetry (Kayne 1994) in arguing that complements linearly precede their heads, so that SOV is the universally unmarked word order. By assuming a more articulate structure where both VP and *v*P are projected, SVO is obtained if V raises leftwards in the narrow syntax, targeting a *v*′ position that c-commands *v* while being c-commanded by the specifier ZP (see Fukui & Takano 1998: 44).

This theoretical apparatus adopted by Fukui & Takano (1998, 2000) allows them to make two distinct claims. On the one hand, they argue that V-raising negatively correlates with the presence of Case-particles in a language, so that an SOV language like Japanese bears them because V does not raise, while English lacks them because V must raise to generate the SVO order (Fukui & Takano 2000: 243). On the other hand, Fukui & Takano (2000: 228) further distance themselves from Kayne (1994) by observing that "if this approach is correct, then it becomes unnecessary to postulate a number of hypothetical «functional categories» triggering overt movements" in SOV languages, which further goes along with "the «traditional» view of comparative Japanese/English syntax that the «movement portion» of the Japanese grammar is very inactive, if not completely inert, compared to the English counterpart" (Fukui & Takano 2000: 228).

Fukui & Takano (1998, 2000) are not the only ones that argue *contra* Kayne (1994) that

SOV is the universal base order, whereas SVO is instead obtained by means of V-movement in the narrow syntax.[33] A similar claim can be found in fact in Barbiers (2000: 181), who assumes that "[s]yntactic structure is generated from bottom to top by uniform leftward application of merge and move, starting out with a verb". Barbiers (2000: 182) further observes that "[t]his hypothesis entails that rightward movement and right-adjunction are not available, as in Kayne 1994 and Haider 1997".

Although Barbiers puts the latter two authors on a par, Haider (1992, 1997) are works that lay the foundation for a theoretical alternative to the LCA that has been further developed in subsequent works (see a.o. Haider 2000, 2013, 2015, 2020b). As anticipated by quoting Barbiers (2000: 181), what Haider proposes hinges on the fundamental premise that "the direction of merger within a phrase is universally to the left" (Haider 2020b: 359 fn. 22), so that the projecting node always follows the non-projecting sister node (Haider 2000: 46). This basic axiom, which is addressed as the *(Basic) Branching Constraint*, rules out the X-bar templates in (40b) and (40d), so that one of the two structures predicted by the Consistent Head Serialization discussed on page 22 is utterly excluded too (Haider 2013: 2 fn. 2).

The Branching Constraint is paired up with another tenet, i.e. "the parametric direction of structural licensing" (Haider 2010: 46). The latter assumption needs to be broken down into two parts, the former being the definition of licensing: "The position of a dependent phrase P is *licensed*" if and only if "a (projection of the) phrase head h and P *minimally* and *mutually* c-command each other" (Haider 2020b: 359, italics in the original). Based on this definition, all dependents would be legitimate in any of the X-bar templates shown in (40). In fact, a head X would always license its complement YP because of the sisterhood relation holding between them, and a specifier ZP would similarly be licensed by X′, namely the sister of ZP that is projected by X. However, the definition of licensing must be put together with the second part of the aforementioned quotation from (Haider 2010: 46), which states that licensing is a directional property that is parameterized for each head in the lexicon (Haider 2013: 75, 2015: Sec. 4), just like the head parameter presented on page 48.

Bearing in mind that the Branching Constraint allows only the structures in (40a) and (40c) on the preceding page, the analysis that Haider provides in his works for clausal word order starts by following the refences mentioned in footnote 7 on page 33, where it is assumed that all arguments of the verb are first-merged VP-internally (Haider 2010: 20, 2013: 10). Like

[33] In this respect, it is worth recalling that *basic order* is not to be confused with *underlying* or *base* order, as Svenonius (2024: 8.3) points out.

any other lexical head, the verb is valued with respect to its canonical licensing directionality, which can be to the left, to the right, or underspecified (see Haider 2013: 30 and Haider 2015: Sec. 5). In the first of the three options, the phrase is head-final and "the canonical directionality is in harmony with the branching direction" (Haider 2013: 5), so that not only the head but any other node projected by it can license the dependents. Thus, in OV languages all arguments are licensed VP-internally as shown in (41). The arrows therein illustrate licensing to the left based on Haider (2015: Sec. 4, 5).

(41) $[_{VP}\ XP \leftarrow [_{V'}\ YP \leftarrow [_{V'}\ ZP \leftarrow V^0]]]$

According to Haider, (41) explains a number of properties of OV languages (like German and related languages), most notably the possibility to scramble arguments. Given that they are always (re-)merged in a position where the verb can license them, arguments (including subjects) are furthermore transparent for extraction (see also Belk & Neeleman 2017 and references in Abels 2015: 1427-1428).

Haider (2010: Sec. 5.2-5.5, 2015: Sec. 5) argues that the aforementioned observations also apply to languages belonging to the third type he assumes, namely those whose verbs have no specified directionality of licensing. The Slavic languages (Haider & Szucsich 2022b) and Hungarian (see Haider 2013: 131, based on Kiss 2003) are argued to be examples of this kind, and their distinguishing feature is twofold. On the one hand, the order of the verb and of its arguments is highly variable, while on the other hand they are likely to be analyzed as SVO languages. This is due to the fact that two of the three alternate structures that languages of this type can instantiate are *prima facie* overlapping with those of SVO languages that establish a fixed directionality of licensing, whereas the third ends up coinciding with (41). However, the three alternatives shown in (42) by following Haider (2015: Sec. 5) illustrate that these patterns are related to the fact that the V head can implement either of the two directionality values each time it is merged (or remerged).

(42) a. $[_{VP}\ XP \leftarrow [_{V'}\ YP \leftarrow [_{V'}\ ZP \leftarrow V^0]]]$

b. $[_{VP}\ XP \leftarrow [_{V'}\ YP \leftarrow [_{V'}\ V^0 \rightarrow ZP]]]$

c. $[_{VP}\ XP \leftarrow [V^0 \rightarrow [_{V'}\ YP\ [_{V'}\ \underline{\cancel{V^0}} \rightarrow ZP]]]]$

The properties described for languages that are either SOV or of the type shown in (42) are argued to be impossible for languages whose verbs license their arguments to the right

(for summaries of the contrasts, see Haider 2020b: 342, Haider 2010: 131-132, Haider 2015: Appendix, Haider 2013: Sec. 9.4). These facts, which also apply to English, follow from "the mismatch between the canonical directionality and the universal directionality of merger" (Haider 2013: 110, italics in the original), which implies that projections of the head no longer represent licensing nodes as in (41) or (42).

This implication further leads to another noteworthy point that has been proposed since Haider (1992, 1997), namely an account for the English VP-shell (Larson 1988, 2014). Given the Branching Constraint, there can be no structural counterpart of (41) if licensing is parameterized in the opposite direction. If licensing is consistently to the right, what one gets instead is the more complex structure in (43), which crucially differs from (42c) only with respect to the number and the direction of the arrows (see again Haider 2015: Sec. 4, 5). Haider (2013: 30) illustrates the rationale behind (43) as follows (italics in the original):

> Being in a foot position, a head that licenses to the right would license only the complement sister to its right [...]. The solution is the *iteration* of the basic projection scheme, namely the head–complement pattern [...]. If, as in [(43)], the basic projection scheme is reiterated x-many times, there are x-many head positions, but there is only a *single* lexical head. So, the structure will be well formed only if the head positions are co-indexed, which is equivalent to the formation of a head-movement chain. This is the *raison d'être* for the shell structures in VO clause structures.

(43) $[_{VP}\ XP\ [_{V'}\ V^0 \rightarrow [_{V'}\ YP\ [_{V'}\ \underline{\sout{V}^0} \rightarrow ZP]]]]$

The account that (43) provides for the shell structures is echoed in various works. Among others, Stabler (1997: Sec. 1.3) has been cited at the beginning of the present section, whereas Belk & Neeleman (2017) will be discussed in section 3.1.4. In addition to those, Georgi & Müller (2010: 3) explicitly mention Haider's proposal as an inspiration for their reworking of an analysis of Greenberg's Universal 20 (see section 3.2.4), namely Abels & Neeleman (2009, 2012).

The appeal of (43) is due to its capability of providing a trigger for the raising of V (Haider 2000: 62-64) within a VP projection that circumvents the theoretical difficulties raised by a *v* head (see Haider 2013: 53-55 and Travis 2014: Sec. 5.1 for discussion). Moreover, the structure in (43) becomes a predictor of what Haider (2013: 107, 142) calls *compactness*,

namely the impossibility of having an adverb break the adjacency between a verb and its object in a VO-language like English. Based on the definition of licensing given above, an adverb (or a scrambled argument) placed between a copy of V and its sister (be it YP or ZP) would disrupt the minimal c-command relation that in (43) holds between the two (Haider 2013: 110). This can be shown in (44), which adapts the data that Haider (2013: 110) provides to show the contrast between two languages that respectively instantiate the structures in (41) and (43), namely German and English.

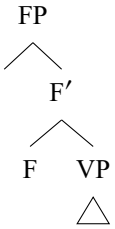
(44) a. *er hat den Studenten <freiwillig> die Lösungen*
3SG.NOM AUX DET.PL.DAT student.PL.DAT <voluntarily> DET.PL.ACC solution.PL
<freiwillig> gezeigt
<voluntarily> show.PTCP
b. he showed (*voluntarily) the students (*secretly) the solution.

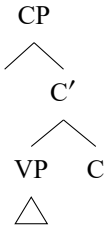
A final aspect concerned with (43) is that XP, i.e. the subject argument, is left unlicensed in its VP-internal position. According to Haider (2013: 107), the merger of XP represents in fact the derivational point from which VSO and SVO languages diverge. On the one hand, VSO languages license it as shown in (45a), namely by raising the V head once more "in a second VP-shell" (Haider 2013: 107). On the the other hand, non-VSO languages whose verbs still license their arguments to the right fix (43) by merging a functional head F above VP, as shown in (45b). In languages that adopt the latter strategy, FP is an "obligatory grammatical device" (Haider 2013: 109) and its specifier can be filled by an expletive so as to prevent the radical emptiness of the projection (Haider 2010: 36-37; see also Haider 2019 for further discussion). As (45b) shows, XP moves to Spec,FP so as to be licensed by means of Spec-Head agreement (Haider 2013: 109).

(45) a. $[_{VP}\ V^0 \rightarrow [_{V'}\ XP\ [_{V'}\ \underline{\underline{V^0}} \rightarrow [_{V'}\ YP\ [_{V'}\ \underline{\underline{V^0}} \rightarrow ZP]]]]]$
b. $[_{FP}\ XP\ [_{F'}\ F\ [_{VP}\ \underline{\underline{XP}}\ [_{V'}\ V^0 \rightarrow [_{V'}\ YP\ [_{V'}\ \underline{\underline{V^0}} \rightarrow ZP]]]]]]$

At this point it is important to stress once more that the head of FP in (45b) is assumed to be non-lexical (Haider 2010: 141). Given that the directionality of complement licensing is parameterized in the lexicon, Haider (2010: 81, 2013: 75) argues that non-lexical heads like F are forced to precede their sister, as they solely obey the Branching Constraint. Thus, the Spec-Head-Comp order that in Kayne (1994) is exceptionless also holds in Haider's approach for any non-lexical functional head. The content of the latter is in turn derived from its

interaction with the lexical (non-functional) head contained in the complement to its right, either by means of head movement or by means of feature attraction (Haider 2013: 71). Based on what has just been said, (46a) is a legitimate structure as long as VP is realized in one of the ways shown in (41)-(43), whereas (46b) is not.

(46) a. 

b. 

The interaction of a lexical category with a non-lexical (functional) head gives rise to what Haider (2013: 71) calls "*a functional extension of a lexical projection*" (italics in the original), exemplified by (46a). However, Haider (2013: 75) also assumes that some functional heads originate in the lexicon. The latter can be distinguished from the non-lexical counterparts because they correspond to overt items (e.g. a lexical complementizer or a determiner, Haider 2013: 71) and therefere cannot be targeted by head movement (Haider 2013: 68).

Moreover, the lexical nature of such functional heads allows them to have a directionality option (Haider 2013: 75), so that they can instantiate (46b). The latter structure is argued to be unproblematic on a twofold ground. On the one hand, "a projection of a *lexical* functional head does not qualify as the *functional extension* of the complement of the functional head" (Haider 2013: 71, italics in the original), while on the other hand Haider (2013: 72) observes that the Branching Constraint "predicts the absence of head movement to a functional head on the right, but it does not rule out the existence of base-generated structures with lexical functional heads on the right" as in (46b). Moreover, one can further remark following Haider (2013: 72) that in (46a) there is a single projection line and that this holds regardless of whether F is lexical or not, whereas in (46b) there are two projections lines, a scenario that is possible only if C originates in the lexicon.

The distinction between functional heads that bear some content of their own and those that do not is echoed in Cinque (2023), as section 3.2.2 will show. Moreover, one can easily acknowledge following Haider (2000: 48) that in a functional extension of a lexical projection like (46a) the Final-Over-Initial configuration that FOFC rules out is excluded as a concequence of the Branching Constraint (see again page 54). In fact, Haider's proposal not

only captures (and predates, Haider 2020a) FOFC, but also argues for the non-existence of the Final-Over-Final configuration that FOFC would allow in structures like (46a).

Still, the two accounts differ with respect to (46b), as Sheehan et al. (2017) do not distinguish functional heads depending on whether they originate in the lexicon or not. Thus, FOFC would rule out (46b) as soon as what is dominated by the CP projection is not consistently head-final, but allow Final-Over-Final configurations for both (46a) and (46b). In this respect, it would be interesting to see whether Haider's proposal fares better at capturing the exceptions to FOFC discussed in Zeijlstra (2023: Sec. 6), Clem (2022) and references therein by analyzing the problematic heads as instances of (46b).

The divergences that have just been observed also derive from the fact that Haider (2013: 339-340) does not refer to extended projections in a way that follows or elaborates on Grimshaw (1991, 2005), contrary to what Biberauer, Holmberg & Roberts (2014) and Sheehan et al. (2017) do. In fact, on page 56 it has been pointed out that a head-initial CP that is the complement of a head-final lexical verb is ruled out by FOFC, but that this is not necessarily the case in Haider's approach. According to Haider (2010: 189), examples of this kind are simply deviant in German, as already reported in footnote 24 on the same page. Such a deviance is a consequence of center embedding, which is a structural configuration that is well-known for raising parsing difficulties, as shown by the following contrast adapted from Chomsky & Miller (1963: 286), Zwart (2011: Sec. 3) and Chesi (2024: 11):

(47) a. (I saw) [a dog [that bit [a cat [that chased [a mouse that ran away]]]]]
b. *(I saw) [a mouse [that [a cat [that [a dog bit]] chased]] ran away]

Although it is "a non-trivial matter to determine whether the acceptability of a certain structure is due to the fact that such a structure is not generated by the grammar or because it imposes unmanageable processing demands" (see Frank 2004: 671 and references therein), Haider (2013: 134, 143-148, 341 ff.) argues that center-embedded head-final projections (to be further discussed in section 5.4.5) are ruled out by the grammar to prevent parsing difficulties, so that in such cases clustering becomes "the mandatory grammatical alternative" (Haider 2013: 134). However, Broekhuis (2008: Sec. 2.3) argues that an analysis along these lines is not immune to problems, mainly because some of the facts that can be observed for verbal clusters in West Germanic would require untriggered (and thus stipulative) movements within the structures that Haider assumes for the verbal complex.

The general avoidance of center embedding argued for by Haider (2013) raises comparable problems when it comes to extrapositions to the right. As discussed in section 2.3.2, many works supporting the validity of FOFC argue that extraposition is a FOFC-compliance strategy, so that OV-languages consequently prefer to have post-verbal embedded clauses whenever the latter contain a head-initial complementizer. Although Haider's utter ban of rightward movement manages to rule out some unattested structural configurations where phrases are involved (and which FOFC does not capture by being only concerned with heads; see Philip 2013: Sec. 5.1 and Feldcamp 2022: 49), it imposes a base-generation account for syntactic phenomena that are addressed as instances of extraposition (see also Philip 2013: 173 ff., 203 on this point).

Albeit Haider (2010: Ch. 5) argues that a base-generation analysis of such phenomena bears some advantages in comparison to the alternate accounts, he eventually acknowledges that it cannot be entirely correct (Haider 2010: Sec. 5.3.10, 5.5, 2013: Sec. 6.6). Moreover, the works to be discussed in section 3.1.4 will show that ruling out rightward movement altogether might be too strict of a requirement, as it either prevents or needs (convoluted and) ultimately inadequate machinery to account for the elements appearing on the right edge of the relevant phrases.

Chapter 3

Greenberg's Universal 20

This chapter consists of two sections. The first, 3.1, is broken down into four parts, where the relevant morphosyntactic properties of demonstratives, cardinal numerals, adjectives and nouns will be discussed in turn. This will allow to present the alternate structures that the generative tradition has provided for each of the four elements of U20 so as to account for the relevant facts. The second section, 3.2, will instead survey the existing analyses of U20, so as to assess the manners in which they have tried to update the empirical and theoretical content of Greenberg's original formulation.

3.1 Definitions

Greenberg (1966–1963: 87, 111) originally stated his Universal 20 as follows:

(48) "When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite".

The definition in (48) is addressed as *Greenberg's U20* in many works dealing with the typological variability of the four elements mentioned. Ever since Hawkins (1983: 118-120), the literature has been reporting more orders than the ones captured by (48) and, as a consequence, several reformulations and analyses have been proposed.[34] The variability concerning the counted attestations and the theoretical accounts proposed in the literature will

[34] Actually, Dryer (2018a: 821) observes that (48) also predicts a permutation whose attestation has subsequently been disputed in the literature on U20, i.e. Num-N-Dem-A. See sections 3.2.5 and 3.2.6 for further discussion.

be overviewed in section 3.2.

Before getting to that point, it is worth dwelling on the way Culbertson & Adger (2014: 5843), a work to be discussed in section 3.2.6, tries to popularize the consequences of (48). For a start, they follow Cinque (2005: 315) in restating the original definition of U20 along these lines:

(49) a. In prenominal position the order of demonstrative, numeral, and adjective (or any subset thereof) conforms to the order Dem-Num-Adj.
b. In postnominal position the order of the same elements (or any subset thereof) conforms to either the order Dem-Num-Adj or to the order Adj-Num-Dem.

To make the content of (49) even clearer for an audience of English non-experts on the topic, Culbertson & Adger (2014: 5843) show that the well-formed English phrase *these two red cars* represents an instance of the Dem-Num-A-N order predicted by (48) and (49a), while **red two these cars* is ruled out by being an instance of A-Num-Dem-N. They further observe that, as long as all the modifiers appear on the same side with respect to the noun, (48) and (49b) also predict *cars red two these* and *cars these two red* (i.e. N-A-Num-Dem and N-Dem-Num-A, respectively).

Although the illustration just given represents a useful glimpse on some of the orders that (48) and (49) capture, it misses an important difficulty that any crosslinguistic comparison of this phenomenon has to face, namely the fact that the elements of interest for U20 bear distributional (see a.o. Wiltschko 2014: Ch. 1-3 for discussion) and morphosyntactic properties (e.g. with respect to patterns of nominal concord, M. C. Baker 2008b, Norris 2017a,b a.o.) that vary across languages. Except from Rijkhoff (2002) and Dryer (2018a), most analyses of U20 that section 3.2 will summarize do not discuss these aspects in depth, even though it is useful in order to explore the variety of constructions in which the various elements of U20 can appear, and subsequently set aside part of these constructions on principled grounds.

The remainder of the present section will provide a brief survey on these matters by focusing in turn on each of the four elements that U20 deals with. The upcoming discussion will benefit from Gil (2001: 1275), who highlights that the propertis of a given linguistic category can be presented along two lines. On the one hand, one can focus on the *internal* syntactic structure and the basic semantic properties of such a category. On the other hand, one can be more concerned with the *external* structural relations that a certain category entartains

with the rest of the expression it appears in. Some of the issues that sections 3.1.1 to 3.1.4 leave open will be taken up in section 5.5.2, where it will be shown that the formalism to be presented and discussed in chapter 5 can shed novel insights that could be worth exploring further in the future.

3.1.1 On demonstratives

Demonstratives have repeatedly been claimed to represent a universal category in the languages of the world (see e.g. Diessel 1999: 2, Lyons 1999: 107, Alexiadou, Haegeman & Stavrou 2007: 95 and Giusti 2018: 12). In his description of the origin, function, and typology of demonstratives in the world's languages, Diessel (1999) proposes four different subcategories, i.e. demonstrative pronouns, demonstrative adverbs, demonstrative identifiers and demonstrative determiners (see Wiltschko 2014: Sec. 6.3.2 and Wamsley 2019: Sec. 1-2 for further references and discussion). Given that U20 is concerned with adnominal modification, only the last of these types is relevant to the present discussion.

Demonstratives are largely associated with two formal properties, namely *definiteness* and *person* (Guardiano 2012: 107). Definiteness can be unpacked further into several subcomponents (see a.o. Lyons 1999, Arkoh & Matthewson 2013, Schwarz 2013, Hsu & Syed 2020, Bernstein, Ordóñez & Roca 2021: Sec. 3), and Guardiano (2012: 107) defines the crucial one to consider with respect to demonstratives as *locality* (italics in the original):[35]

> Here, by locality I refer to the property of signaling the referent as identifiable on the basis of contextually related information, i.e., of referring to elements located in the external context (*deixis* in a proper sense), already mentioned in the linguistic context (*anaphora* in a broad sense), or available in the shared knowledge of the participants and marked as relevant in such a 'shared universe' (*topicality*):

The clearest (and universal, Diessel 1999: 2) instance of locality in demonstratives is represented by the fact that "[a]ll languages encode **at least** one distinction between so-called proximal and distal demonstratives, either by explicit lexical forms, such as in the contrast

[35] It is worth pointing out that according to Guardiano (2012: 107-108, italics in the original) demonstratives "contrast with *default* definiteness operators (i.e., definite articles)" because the former "encode a *subset* of the properties" linked to definiteness, to the exclusion of *familiarity*, *uniqueness* and *existence*. This claim is in contrast with Roberts (2017b: 172) and Dékány (2021: 99), where it is argued that the opposite holds, i.e. the features of demonstratives form a superset that properly includes those of definite articles.

between *this* and *that* in English, or via reinforcing elements, such as in the contrast between the Danish expressions «den *her*» and «den *der*»" (Rocca, Tylén & Wallentin 2019: 2, italics in the original, boldface added). The latter example from Danish allows to illustrate how the realization of locality can vary crosslinguistically. For a start, there can be demonstratives that do not convey a specific degree of proximity (see Lyons 1999: 112, 115-116, Alexiadou, Haegeman & Stavrou 2007: 103 and Wamsley 2019: Sec. 3.2 for examples and discussion), but this does not imply that they do not bear a locality feature, as they are still used to signal the contextual availability of the modified referent (Lyons 1999: 21, Rocca, Tylén & Wallentin 2019: 2).

The aforementioned example from Danish also shows that the degree of proximity can be encoded in a separate morphosyntactic element that is typically addressed as a *reinforcer*, whose categorial nature varies crosslinguistically. It can be an adverbial like *there* in the example in (50a) provided by Jenks & Konate (2022: 14), a larger PP (Giusti 2018: 24) as shown in (50b) based on Brugè (1996: 20-21), or an element that reduplicates the form of the demonstrative with which it co-occurs. Such a reduplication can be partial, as reported in (50c) following Dékány (2021: 115), or complete, as illustrated by means of (50d), taken from Paul (2009: Sec. 3).[36] What is more, adverbial reinforcers are also attested at the left edge of the nominal domain, as shown in (50e) following Roehrs (2010: 227).

(50) a. them there fellows come through here, [...]. [Southern US English]

b. *el libro este de aquí* / *Este libro de aquí*
DET.M.SG book.M.SG DEM.PROX of ADV.PROX / DEM.PROX book.M.SG of ADV.PROX
'The book this of here' / 'This book of here' [Spanish]

c. *ez a könyv e* [Hungarian]
DEM.PROX DET book PROX
'This (here) book'

[36] Although Dékány (2021: 114-115) provides three arguments in favor of the reinforcer nature of the postnominal element in (50c), Dékány (2021: 116) claims that it cannot be a genuine example because "the consensus view is that reinforcers are merged within the DP projected by the demonstrative pronoun". However, it will be shown shortly that Rauth & Speyer (2021) argue that postnominal reinforcers are right-adjoined PPs, which would be in line with the postnominal nature of Hungarian PPs (Dékány 2021: 116). With respect to (50d), it is worth pointing out that Malagasy distinguishes demonstratives into six different degrees of proximity, so that *io* indicates a referent that is really close to the speaker (Paul 2009: Sec. 3).

d. *io olona ratsy fanahy io* [Malagasy]
DEM.PROX person bad spirit DEM.PROX
'This mean person'

e. *ot der guter Mann* [Yiddish]
ADV DET good man
'This good man'

The nature of said reinforcers also illustrates that the locality property is not exclusive to demonstratives, but is also typical of other categories (see also Lyons 1999: 18). The same holds true for the second property of demonstratives that has been mentioned following Guardiano (2012: 107), i.e. person. The assumption of the latter can be motivated by observing that locality makes demonstratives be "intrinsically carriers of 3rd Person, as they point to an object which is distinguished from the speaker and the hearer" (Giusti 2018: 24). This is confirmed by typological data from WALS (Bhat 2013), where it is highlighted that in a relevant number of languages demonstratives play the role of third person pronouns (see a.o. Basque, Hualde & Ortiz de Urbina 2003: 151). What is more, the feature makeup of demonstratives and pronouns can overlap even in languages where the former are not a suppletive form of the latter. One such case is represented by Turkish (see Leu 2014: Sec. 3.7.2 and references therein):

(51) a. *bu(n)* 'this one' (close to speaker and hearer)
ben 'I' (nominative)
b. *şu(n)* 'that one' (further away from speaker and hearer)
sen 'You' (nominative)
c. *o(n)* 'that one' (far away from speaker and hearer)
o 'He/she/it' (nominative)

Many of the works cited so far in this section (e.g. Royer 2022: 405 and Jenks & Konate 2022: 21) follow Postal (1969) in observing that pronouns can be analyzed as definite determiners with null or elided NPs, an approach that is further supported by the crosslinguistic frequence of adnominal pronoun constructions of the *we linguists* type (see Höhn 2017 for extensive discussion). Consequently, many authors assume that demonstratives encode a third property in addition to definiteness and person, which Giusti (in press: Sec. 6) calls *referentiality*. In formal terms, such a property corresponds to the interpretation of an individual

index (Giusti 2018: 24, Royer 2022: 400 fn. 17) that is arguably syntactically represented in adnominal pronoun constructions, demonstratives and anaphoric definites (Jenks & Konate 2022). Anaphoric definites are determiners that are used to refer to a previously mentioned linguistic expression, as the example taken from Schwarz (2009: 240) illustrates:

(52) *Hans hat schon wieder angerufen. Ich will # vom / von*
Hans AUX already again call.PTCP 1SG.NOM want of.DAT.ART$_{weak}$ / of
dem Idioten nichts mehr hören.
DAT.ART$_{strong}$ idiot.DAT nothing more hear
'Hans has called again. I don't want to hear anything anymore from that idiot.'

As the subscripts in the glosses show, the fact that the article and the preposition cannot be fused felicitously is taken as evidence in favor of two types of definite articles in German. Schwarz (2013: 540) further observes that the English translation "quite naturally uses *that* in this case" (italics in the original), which highlights that the article in (52) bears locality as per Guardiano (2012: 107). And in fact, it is not only true that German anaphoric definites can be used as demonstratives (see Schwarz 2009: 34-37), but there is also evidence that demonstratives can work as anaphoric definites in languages that lack articles altogether. Jenks (2018: 501) argues for example that in Mandarin Chinese "unique definites are realized with a bare noun, and anaphoric definites are realized with a demonstrative, except in subject position", as Jenks (2018: 501) shows by means of (53).

(53) *wo zuotian yudao #(na ge) nansheng*
1SG yesterday meet DEM.DIST CLF boy
'I met the #(aforementioned) boy yesterday.'

Although demonstratives are not translational equivalents of definite articles in languages that lack the latter (Jenks 2015: 105), the data in (52) and (53) show that the two bear overlapping uses and properties, both intra- and cross-linguistically (see Stroh-Wollin 2016: 116, a.o.). Moreover, there is ample evidence that demonstratives often grammaticalize into articles over time (see Diessel 1999: 128 ff. and Gelderen 2011: Ch. 6).

It is for these reasons that a syntactic analysis of demonstratives represents a major interest among supporters of the Functional Head Hypothesis to the nominal domain. In fact, the DP projection is not only argued to formally encode features like definitness and person (see a.o.

Guardiano 2012: 107-108 and Bernstein, Ordóñez & Roca forthcoming: Sec. 1), but as soon as one takes "D (or its specifier) to be the locus of interpretation of the referential index of the whole nominal expression, the universality of DP is straightforwardly supported by" the universality of demonstratives (Giusti in press: Sec. 1).

An important starting point for the FHH formalization of the aforementioned facts is to assume that demonstratives are phrasal elements that eventually occupy Spec,DP at the end of the derivation (see a.o. Hsu & Syed 2020: e117, Jenks & Konate 2022: 15 and references therein).[37] Two first arguments in favor of the phrasal nature of demonstratives come from the fact that they can stand alone by functioning as pronouns, and can co-occur with the definite article (Bernstein 1997: 92-93, Jenks & Konate 2022: 13-14). The latter fact has been implicitly presented on page 82 with the examples in (50b) and (50c) from Spanish and Hungarian, respectively. Jenks & Konate (2022: 15) argue that the argument becomes even stronger once one looks at examples like (54) from Alexiadou, Haegeman & Stavrou (2007: 103), where demonstrative and reinforer are adjacent and precede the co-occuring article.

(54) a. *afto-edho* *to* *praghma* [Modern Greek]
DEM-ADV.PROX DET thing
'This thing'

b. *afto-eci* *to* *praghma*
DEM-ADV.PROX DET thing
'That thing'

According to Jenks & Konate (2022: 13-15), the adjacency of demonstrative and reinforcer in (54) is compatible with a view where they form a single phrasal constituent. However, the latter claim represents a first point on which opinions diverge, as Rauth & Speyer (2021: Sec. 4) review three proposals that support such a constituency (i.e. Brugè 1996, Roehrs 2010, Leu 2014) and conclude that each of them suffers from problems of inconsistency or of empirical inadequacy.

Another debated aspect is related to the structural position where demonstratives are first-merged. By focusing on "data from languages where a non enclitic definite article is in D and D is DP-initial", Guardiano (2012: 103) argues that six different types are attested with

[37] There are cases where demonstratives bear head-like properties, but they are always explained away by the authors who report them (e.g. Hsu & Syed 2020: e118, Dékány 2021: Sec. 3.2.1 and Jenks & Konate 2022: 15-16). In general, such counterexamples are not seen as something that undermines the claims reported in the main text, but rather details that need further discussion or slight refinements of the proposal.

respect to the position of the demonstrative and its co-occurrence with the definite article in the nominal extended projection. However, only the first two of them are consistent with a view in which the demonstrative occurs in the left specifier of the DP, thus being the leftmost edger of the nominal domain (Giusti 2018: Sec. 2.2). These two types correspond to the cases where the co-occurrence with the article is either prohibited as in English or mandatory as in Hungarian (Guardiano 2012: 104).[38]

(55) a. <this> (*the) man <*this>
b. *<ez>* **(a)* *ferfi* *<*ez>* [Hungarian]
DEM.PROX DET man DEM.PROX
'This man'

The four additional types that Guardiano (2012: 103-106) finds are at odds with the base-generation of a demonstrative in Spec,DP because they are cases where the demonstrative is postnominal, as in the Spanish case *El libro este* shown in (50b) on page 82. According to Guardiano (2012: 106), Spanish is of a particular type mainly because its demonstratives "never co-occur with the article when they are DP-initial; they also occur postnominally, and in this case a definite article is compulsory in D". Greek is similar to Spanish, but already exemplifies a distinct type. This is due to the fact that Greek differs from Spanish in imposing in all cases a mandatory co-occurrence of the demonstrative with the definite article, as shown in (56) by adapting data from Guardiano (2012: 105).

(56) *<to spíti>* *ekeíno* *<to spíti>*
<DET.N house.N> DEM.DIST.N <DET.N house.N>
'That house'

Both Spanish and Greek differ from the remaining two types found by Guardiano (2012: 104), because in the latter the demonstrative can only occur postnominally. Such types, respectively exemplified by Welsh and Hebrew, differ from one another because in the former the definite article is DP-initial, whereas in the latter it "either spreads on all the items, included the demonstrative, or is not visible at all" (Guardiano 2012: 104). In (57), the example

[38] The premises discussed in Guardiano (2012: Sec. 3) predict three further types that are not attested. The first is the case where demonstratives can only be leftmost edgers as in (55), but with an optional article. The other two options correspond to languages lacking articles and whose demonstratives are never or optionally leftmost edgers of the nominal domain.

in Welsh is taken from Willis (2006: Sec. 4.4), whereas the example from Hebrew is adapted from Guardiano (2012: 104).

(57) a. *y cwpan gwyrdd arall 'ma / hwn* [Welsh]
DET cup green other DEM.PROX / DEM.PROX
'This other green cup'

b. *more ze / ha-more ha-ze* [Modern Hebrew]
teacher DEM.PROX / DEF-teacher DEF-DEM.PROX
'This teacher'

Once the crosslinguistic variation just presented is agreed to be the consequence "of parameters on the pronunciation of functional projections" (Hsu & Syed 2020: e117), what remains to clarify is the base structure that leads to the observed variation. A widespread assumption in this respect is that the demonstrative must be first-merged "in the specifier of a functional projection between DP and the lexical NP" (Rauth & Speyer 2021: 12-13). Given that adjectives are also assumed to have their base position in that portion of the nominal extended projection, the argument is often further supported by enhancing the formal properties that demonstratives share with adjectives, e.g. the possibility to bear inflection (Bernstein 1997: 93, Leu 2014, Giusti 2018: Sec. 2.1) and to occupy comparable syntactic positions when not realized in Spec,DP (Guardiano 2012: 105-106, Giusti 2018: Sec. 1.1).

The literature debates on three possibilities to explain the just mentioned facts. The first account is provided in works like Brugè (1996), Giusti (2002) and Guardiano (2012), where it is argued that the demonstrative is base-generated low in the structure, above the lexical NP but "presumably not higher than the structured adjectives" (Guardiano 2012: 108). This view is also supported in the analysis of U20 given by Roberts (2017b) and will be discussed further in section 3.2.3. For the time being, it is worth observing that such a perspective struggles to account for cases where the demonstrative precedes other modifiers of U20, but follows the article and/or the noun (see references in Dékány 2021: 94). The Cangin language Noon exemplifies this, as Soukka (1999: 179) shows for (58):[39]

[39] Although in (58) the spread article bears the same form and the same locality feature of the demonstrative, there are good reasons to treat the latter as a separate category. First, a demonstrative cannot occur without the (spread) definite article, but its presence in the nominal is not mandatory (Soukka 1999: 178-179). Moreover, elements like genitive NPs interpose between the noun and the demonstrative (Soukka 1999: 184), and demonstratives have a four-way contrast, while articles bear only three. This fourth type of demonstrative denotes very distant referents (Soukka 1999: 427) and bears a form that diverges from that of other demonstratives and articles (see Merrill & Apel 2021: 354 for further data and discussion). It

(58) *enoh-cii* *cii* *ci-yaanaaw-cii* *ci-daaŋkah-cii* [Noon]
cow-DET.PROX DEM.PROX AGR-white-DEF.PROX AGR-ten-DEF.PROX
'These ten white cows'

The example in (58) is more adequately captured by the second perspective on demonstratives not base-generated in Spec,DP. Roehrs (2010) and Baggio (2019) and Dékány (2021: 148) argue (among others) that the demonstrative is high in the structure, in a position where it is dominated by DP but still c-commands the other elements of Greenberg's U20. Dékány (2021: 148) calls this projection DeixP and argues that it is also more plausible in semantic terms than a lower counterpart (Dékány 2018: 3, 130). Similarly, Baggio (2019) supports this view based on diagnostics applied to Italian and Spanish, like coordination, ellipsis and scope. With respect to the latter, Baggio (2019: Sec. 3.4.3) observes that in (59) the demonstrative must scope over the element meaning 'other', and not vice versa.

(59) *las* *otras* *personas* *estas* [Spanish]
DET.F.PL other.F.PL person.F.PL DEM.F.PL
'These other people'

Although the two accounts just discussed mutually exclude each other, they share the idea that the demonstrative should be able to undergo leftward raising independently of other elements of Greenberg's U20. As sections 3.2.1 and 3.2.2 will show, Cinque (1996: 454 fn. 16) follows Brugè (1996) in acknowledging that demonstratives could be base-generated lower than Spec,DP, but he does not take up this possibility in his subsequent analyses of U20 (Cinque 2005, 2023) that rely on antysimmetry (Kayne 1994). In any case, neither of the two alternatives implies a rejection of antisimmetry. On the one hand, Roberts (2017b) argues that the low first-merged position of the demonstrative can be reconciled with antisymmetry (though see section 3.2.3 for discussion), while Baggio (2019: Sec. 6) argues on the other hand that the alternative with a demonstrative immediately below DP is compatible with the roll-up derivations discussed in section 2.3. However, the latter alternative arguably requires a DP split into multiple projections, so that there are at least two specifiers that can host the raised Dem and a raised remnant containing N, respectively (see also Giusti 2018 for related discussion).

is also worth recalling that a nominal expression like (58) is excluded by the survey in Guardiano (2012: 103) because it bears an enclitic article.

Willis (2006: Sec. 5.1) uses the Welsh example in (57a) on page 87 to argue that both alternatives suffer from a conceptual problem. In (57a), the precedence relation among elements is as shown in (60a). By following a.o. Cinque (2005: 327, 2014: Sec. 3), Kayne (2008: Sec. 2.1), and Velde (2022: 910), elements meaning 'other' can be claimed to scope over other adjectives and have a peculiar syntactic position in the structure. Thus, it can be argued that the two aforementioned perspectives on the base position of the demonstrative would derive (60a) either from (60b) or from (60c). Both analyses must then make a further twofold assumption to correctly account for (60a). On the one hand, an additional and independent leftward raising of *arall* 'other' is required, while the remnant containing both Adj and N must raise past both Dem and 'other' on the other hand. The option in (60b) further competes with the theoretical alternative in (60d), where all modifiers stay *in situ* and only N raises (see Willis 2006 for further discussion).

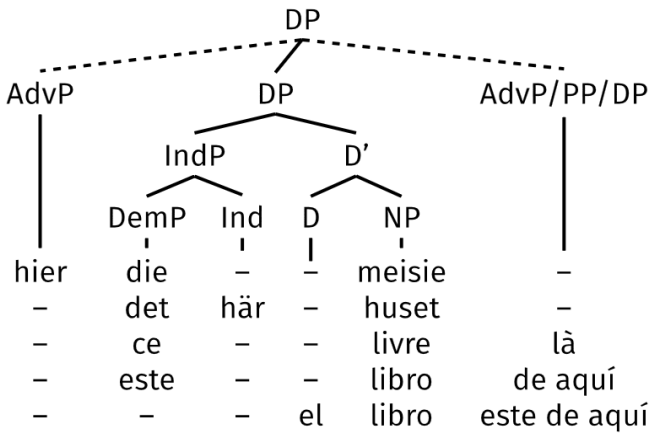
(60) a. D ≻ N ≻ Adj ≻ 'other' ≻ Dem
b. D > 'other' > Adj > Dem > N
c. D > Dem > 'other' > Adj > N
d. D > Adj > 'other' > Dem > N

At this point, Willis (2006: 1822) observes that (60d) implies that both 'other' and Dem are in their base position in Welsh, but not in English. One can then point out that the opposite holds for (60c), while the base order in (60b) corresponds to neither the Welsh nor the English basic order. Thus, any theoretical alternative proposed in (60) raises a relevant problem that Willis (2006: 1822) discusses for (60d), as follows (italics in the original):

> Although this produces the correct outcome for Welsh, we are left to wonder about the crosslinguistic implications of such an analysis. We are forced to say that, in an English example [.. there must be] a landing site for *other*, a new functional projection will need to be posited, for instance OtherP_2 [...] This is an unsatisfactory analysis, not least because it reintroduces crosslinguistic variation in the merged position of adjectives by the back door. The two functional projections, OtherP_1 and OtherP_2, merely restate the fact that, in some languages, 'other' appears very late in the adjective cluster, whereas in others it appears very early. Even worse, the theory is too permissive: the mirror-image nature of the variation becomes an arbitrary result [...].

Given the aforementioned issues, one could follow Willis (2006: Sec. 8) and Adger (2013) in arguing that demonstratives can be right branches in the nominal structure.[40] One first option is thus to reconcile this claim with the high demonstrative below DP supported by Roehrs (2010), Baggio (2019), and Dékány (2021), so as not to loose consistency with facts of scope. Alternatively (or in addition to that), one could follow Rauth & Speyer (2021: 17 ff.), who argue that reinforcers undergo adjunction to the outer DP-layer, with right adjuncts being capable of containing a reinforced demonstrative like the Spanish *este de aquí* discussed in (50b) on page 82. The representation in (61) shows what Rauth & Speyer (2021: 17) propose to account for the relevant facts related to the Germanic and the Romance languages. Although they posit that demonstratives are either part of the right adjunct or first-merged in Spec,DP, one could add a base-generation of IndP in a position immediately below DP to account for the data from Noon in (58).

(61)



Rauth & Speyer (2021: 17 ff.) propose the representation in (61) to account for the grammaticalization path that demonstratives have undergone in the German dialects of Rhine and Moselle Franconian. A crucial assumption is that demonstratives can start their process of grammaticalization from a DP-external position. This claim is in line not only with what Bernstein, Ordóñez & Roca (2021: 14) similarly argue for the demonstratives in Latin, but also with the theoretical proposal that Velde (2022) formulates to account for the variation that is found among the Bantoid languages in the nominal domain.

A potential problem to the claims just mentioned is however raised by the Spanish data discussed in Guardiano (2012: 106), where PPs not only appear postnominally, but also

[40] This would also be in line with a LFG analysis of the nominal domain in Welsh, i.e. Mittendorf & Sadler (2005).

follow the right demonstrative:

(62) a. *el profesor este del chico*
DET.M.SG professor.M.SG DEM.PROX.M.SG of.DEF.M.SG boy.M.SG
'This teacher of the boy'

b. *la casa esta con dos jardines*
DET.F.SG house.F.SG DEM.PROX.F.SG with two garden.PL
'This house with two gardens'

The data in (62) represent a problem because PPs are argued to be merged within the projection of the lexical NP (see a.o. Laenzlinger 2017) in a view where what is further right in the string is lower down in the structure (Baggio 2019: Sec. 2.2). However, section 3.1.4 will show that these assumptions related to the structural position of PPs are not immune to problems, so that it may well be the case that future rethinking will eventually show that the data in (62) are not as problematic as they first seem in relation to what is proposed in (61) by Rauth & Speyer (2021).

3.1.2 On cardinal numerals

Greenberg's Universal 20 can be included among the "formal linguistic approaches to the syntax and semantics of numeral-containing expressions", where numerals are understood as "linguistic objects that include the notion of cardinality in their denotation" (Ionin & Matushansky 2018: 1). The number of members that a single language includes in the category of numerals is subject to broad crosslinguistic variation, as there are languages where any natural number can be expressed, whereas others have a very impoverished means of counting (see a.o. Hammarström 2007, Epps et al. 2012 and Doetjes 2021b: 73). The language-specific syntactic properties of cardinal numerals also vary greatly, as they can be argued to belong not only to a separate category that may also include other (types of) quantifiers, but also to the categories of nouns, or adjectives, or (rarely) verbs, or a mix thereof (see a.o. Rutkowski 2006 for the language-internal variation in Polish and Ionin & Matushansky 2018: Ch. 6-7 for data from multiple languages).

When it comes to U20, the focus on numerals is narrowed down by excluding ordinal and modified numerals (e.g. *third* and *at least three*). The direct modification holding between the numeral and the noun further leads to the exclusion of partitives (like *one of my friends*,

Ionin & Matushansky 2018: Ch. 10) and pseudopartitives (see Klockmann 2020: Sec. 2.1 for a list of examples and references), whereas the interest towards the basic order that each language adopts for U20 leads to the exclusion of non-canonical and pragmatically marked phrases like *an amazing sixteen performances* (which are discussed in Ionin & Matushansky 2018: Ch. 8, Klockmann 2020 and Keizer 2022, a.o.).

Crosslinguistically, the cardinal numeral 'one' stands out for being often intertwined with the grammaticalization of indefinite articles (see a.o. Gil 2001: 1277 and Gelderen 2011: 202). However, this and related facts should not lead to its aprioristic exclusion when making a crosslinguistic survey on U20. In fact, the co-occurrence in unmarked contexts of this numeral with a demonstrative is attested, although (63a) shows that there are languages like Italian where this option is ungrammatical. The examples in (63b) and (63c) are respectively taken from Wamsley (2019: 11) for the Kuki-Chin language Hakha Chin and from Royer (2022: 401) for the Mayan language Chuj.[41]

(63) a. *questo* (**uno*) *gatto* [Italian]
DEM.PROX.M.SG (*one.M.SG) cat.M.SG
'This cat'

b. [*hi arti pum khat hi*] *ka ei* [Hakha Chin]
DEM egg CLF one DEM 1SG eat
'I eat this one egg.'

c. *yam* [*jun tz'i' chi'*] [Chuj]
grab one dog DEM
'Grab that (one) dog.'

As discussed in Ionin & Matushansky (2018: Sec. 3.1) and references therein, in the generative tradition there is no consensus on the phrasal vs. head status of cardinal numerals. In this respect, Danon (2012: 1282) further observes that many works on the matter assume that "the structural relation between the numeral and the noun that it combines with is uniform, both cross-linguistically and language-internally" (see Ionin & Matushansky 2018: 3 for a list of references). The appeal of this theoretical view consists in the fact that it allows to argue that simplex and complex cardinals (e.g. *four* and *four hundred*, respectively) can be captured

[41] Royer (2022: 400-401) puts (63c) in contrast with an alternate sentence where *jun* 'one' is replaced by the classifier *nok'*. No matter if the context is such that there is one single dog or more than one, (63c) is always felicitous. See the referenced work for further discussion.

by means of the same syntactic representation (see Witkoś & Dziubała-Szrejbrowska 2018 and references therein for formal accounts along these lines), even though complex cardinals necessarily require richer structures due to the addition and/or multiplication of some numerical unit with a base, as in *twenty-five* (see Comrie 2013, Hammarström 2007 for definitions and further discussion). To the exception of Cinque (2023), this potentially problematic assumption is implicitly endorsed in all the generative approaches to U20 that will be discussed in section 3.2.

Danon (2012) argues that there are good reasons to claim that cardinals can be realized as heads as well as phrasal specifiers, and that this variation can be found within a single language. Most of his arguments are based on the morphosyntactic behavior of pluralized or coordinated numerals (e.g. *hundreds of books* and *two or three books*, respectively) or on partitive constructions. In addition to those, Danon (2012: 1286-1287) backs up his proposal by showing that in Welsh numerals up to 10 can be realized indistinctly with one of the two alternatives reported in (64), with no semantic difference:

(64) a. *naw ddyn*
nine man.SG
b. *naw o ddynion*
nine of man.PL
'Nine men'

Further language-internal evidence supporting the idea that cardinals are not consistently phrasal is provided by Artiagoitia (2006, 2008) with respect to Basque and by Syed & Simpson (2017) and Simpson & Syed (2020) with respect to Bangla. The latter argue that in Bangla numerals from 2 to 5 and higher numerals are respectively heads and specifiers of the same functional projection. As will be discussed in greater detail in section 4.5.2, the leftward raising of an adjective and/or a noun is possible in Bangla only with lower numerals. Thus, the head status of the latter arguably follows from the fact that the specifier within their projection is left free to work as an escape hatch for structurally lower nominal elements that need to raise. On the other hand, the fact that such raisings are bled by higher numerals is argued to follow from the phrasal status of the latter.

As discussed in section 2.3.2, Duguine (2022) posits that the clausal domain in Basque is consistently head-final (but with left specifiers), with the exception of (part of) its left

periphery. By assuming that the nominal domain is also consistently head-final, Artiagoitia (2006) argues that *bat* 'one' and *bi* 'two' are heads of a head-final functional projection due to their postnominal nature, whereas any other numeral occupies the specifier of the same projection by being prenominal. Artiagoitia (2006: 124 and 120) illustrates this constrast in the standard variety of Basque by means of (65a) and (65b).[42] The interested readers are referred to Artiagoitia (2006, 2008) for the full list of arguments backing up these theoretical assumptions.

(65) a. *mutil eder bi hau-ek* [N-Adj-Num-Dem]
boy beautiful two DEM.PROX-DET.PL
'These two beautiful boys'

b. *lau sagar eder hau-ek* [Num-N-Adj-Dem]
four apple beautiful DEM.PROX-DET.PL
'These four beautiful apples'

It is worth noting at this point that the facts just presented are not necessarily regarded as compelling by other works. For instance, it is interesting to compare what (65) illustrates for Basque with the analysis provided by Shlonsky (2020) with respect to numerals in the Semitic language Mehri. In Mehri, numerals for 'one' and 'two' behave in a different way as opposed to higher numerals, which is something similar to what (65) illustrates for Basque. However, according to Shlonsky (2020) all numerals are phrasal and occupy the same structural position. Shlonsky (2020: Sec. 2) argues that the different word orders in (66) follow from the leftward movement of N, which either raises all the way past Dem or stops above Adj, depending on the numerals involved in the derivation.

(66) a. *a-k'əló:mət əlyó:məh θró:h ʿá:fər*
DET.pen.PL DEM.PROX.PL two red.PL
'those two red pens' [N ≻ Dem ≻ 'one'/'two' ≻ Adj]

b. *a-k'əló:mət rəbó:t ʁəyú:g łyéx*
DEM.PROX.PL four man.PL old
'those four old men' [Dem ≻ 'three'/'four'... ≻ N ≻ Adj]

[42] As Artiagoitia (2008: Sec. 3) shows, Western Basque features prenominal demonstratives, so that in such a variety the counterparts of (65a) and (65b) realize the order Dem-N-Adj-Num and Dem-Num-N-Adj, respectively.

From a theoretical point of view, Ionin & Matushansky (2018: Sec. 3.1) agree with Danon (2012) in arguing that two structures are possible where a numeral modifies a noun directly. However, Ionin & Matushansky (2018: Sec. 3.3) show that the cardinal-as-specifier analysis proposed by Danon (2012) is not compatible with their approach, as in the latter cardinal numerals are either phrasal adjuncts or heads selecting NP as their complement (with both alternatives applying recursively to form complex numerals). Moreover, Ionin & Matushansky (2018: Sec. 7.3-7.4) argue that their analysis manages to capture the facts discussed by Danon (2012), even though the cardinal-as-specifier approach is dispensed with.

In contrast with both Danon (2012) and Ionin & Matushansky (2018), Witkoś & Dziubała-Szrejbrowska (2018) suggest a third way, namely a generalized cardinal-as-specifier approach. If on the one hand Danon argues that the cardinal-as-the-head and the cardinal-as-specifier analyses lead to two distinct mechanisms of case assignment within the nominal domain, Witkoś & Dziubała-Szrejbrowska (2018) argue on the other hand that the same asymmetries can be captured by assuming that the numeral merged in the specifier of a functional projection either bears its own independent case feature or does not, thus conditioning the subsequent morphosyntactic behavior of the whole construction.

At this point it is worth discussing the nature of the functional projection that is argued for cardinal numerals in the aforementioned works (but Ionin & Matushansky 2018). Artiagoitia (2006, 2008), Syed & Simpson (2017), Simpson & Syed (2020), and Witkoś & Dziubała-Szrejbrowska (2018) label this projection as QP. This choice is linked to the fact that cardinal numerals bear overlapping properties with quantifiers like *some* and *many*, so that the latter and cardinal numerals are often claimed to belong to the same functional projection, just as has been argued for demonstratives and definite articles in section 3.1.1 (for discussion, see a.o. Zamparelli 2000: 222, 250 ff., Gil 2001: 1278, Roehrs 2020: 547-550, Dékány 2021: 72 ff., Doetjes 2021b: 74). However, Danon (2012) differs from such works by arguing that the functional head that can have its specifier filled by a cardinal numeral is the same Num head first assumed by Ritter (1991) (see section 2.2.4).

Although reasonable at first sight, it is problematic to follow Danon (2012: Sec. 4) in assuming that cardinal numerals and (plural) number marking are so tightly related. In fact, the discussion in Wiltschko (2021a: Sec. 8.5.2) shows that there is evidence of the realization of number marking in heads other than Num, e.g. D and the lexical N. Moreover, Landau (2016: Sec. 5) questions Danon's assumption by proposing an analysis of concord that hinges on the

Num head being merged amidst structured adjectives, hence lower than the projection that can host the cardinal numeral.[43] Landau (2016: 993 fn. 12) further observes that the analyses of Greenberg's U20 tend to call the latter projection NumP, albeit it does not correspond to the functional projection first assumed by Ritter (1991). In fact, sections 3.2.2 and 3.2.3 will show that the heads that realize number marking are not labeled as NumP either in Cinque (2023) or in Roberts (2017b).

Given these premises, it can now be mentioned that languages can be divided into a somewhat dichotomous opposition, based on which numeral-noun constructions are expected to display either number marking or *sortal classifiers*. As Tang & Her (2019: 297-298) clarify, "in a plural-marking language such as English, count nouns are quantified by numerals directly" (as in *five books*), whereas (67a) follows Tang & Her (2019: 298) in illustrating that in a classifier language like Mandarin Chinese the analogous process of quantification between a numeral and a noun requires the intervention of a word or morpheme belonging to a specific class, i.e. a sortal classifier.

(67) a. *wu ben shu*
five $\text{CLF}_{\text{volume}}$ book
'Five books'

b. *wu xiang shu*
five MENS_{box} book
'Five boxes of books'

c. *wu ping shui*
five $\text{MENS}_{\text{bottle}}$ water
'Five bottles of water'

The examples in (67b) and (67c) further follow Tang & Her (2019: 298) in showing that classifier languages distinguish between sortal classifiers and *mensural* classfiers, so that the latter are used in (pseudo-)partitives where both a mass noun and a count noun can be involved, e.g. *five boxes of books* and *five bottles of water*, respectively. Although the partitive constructions in (67b) and (67c) are semantic equivalents of the English examples just men-

[43] In particular, Landau (2016: Sec. 5) establishes this specific position for Num so as to provide an adequate analysis of mixed concord, which is found in nominal domains "(i) where every element expresses the same features but not necessarily the same feature values and (ii) where elements expressing concord do not all express the same features" (Norris 2017a: Sec. 3). See both Norris (2017a) and Landau (2016) for further references and discussion.

tioned, both Tang & Her (2019: 298) and Doetjes (2021a: 221-222) observe that the measure terms that one must use in English bear different syntactic properties as opposed to the mensural classifiers of Mandarin Chinese. In fact, Wu & Her (2021: 40) highlight that sortal and mensural classifiers "appear in exactly the same syntactic position and are in complementary distribution". Furthermore, the semantically similar measure phrases that are used in English "are, from a grammatical point of view, more similar to ordinary nouns", as they are not "permitted in positions where ordinary nouns are not allowed" (Doetjes 2021a: 221), contrary to what happens with sortal and mensural classifiers.

The aforementioned exclusion of (pseudo-)partitives from an analysis oriented towards Greenberg's U20 leads to focus on the anticipated dichotomy between number marking and sortal classifiers. In this respect, the literature has recently been focusing on a strong typological tendency known as the Greenberg-Sanches-Slobin Generalization (GSSG), defined as such because the first formulations date back to Greenberg (1990–1972) and Sanches & Slobin (1973). The refined version given by Tang & Her (2019: 325) states that "(a) a language is unlikely to have both sortal classifiers and morphosyntactic plural markers, and (b) if a language does have both, then their use is in complementary distribution". As Tang & Her (2019: 302) observe, the GSSG just reported is empirically correct in not excluding languages where neither classifiers nor plural markers are present. What is more, this generalization (whose potential counterexamples are discussed in Kim & Melchin 2018, Tang & Her 2019, Doetjes 2021a) lays the foundation for nominal extended projection provided by Borer (2005: Ch. 4), which is structured as follows (see Borer 2005: 96 ff.):

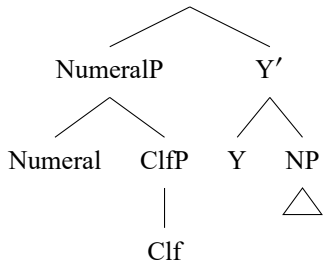
(68) $[_{DP}\ D\ [_{\#P}\ \#\ [_{DivP}\ Div\ [NP]]]]$

In (68), NP is taken to have a mass denotation, and this does not hold for a subset of languages as in Chierchia (1998), but rather for any language. The (optional, Borer 2005: 96) Div head bears a dividing function and, as soon as it is realized by means of either CL (i.e. a sortal classifier) or PL (i.e. number marking), the noun becomes countable and can interact with the counting system in #P (which is also an optional projection). In (68), the GSSG follows from the fact that number marking and sortal classifiers compete for the same position, and a language is more likely to select either of them than to have them both, although such an option is not excluded (see the discussion w.r.t. Armenian in Borer 2005: 94-95).

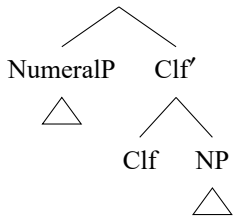
Similarly to what has been discussed for numerals, there is no established consensus on

how to adequately capture the behavior of sortal classifiers syntactically. In this respect, Little, Moroney & Royer (2022) provide a middle ground among competing proposals by arguing that two options are attested crosslinguistically. Such alternatives are given in (69a) and (69b) by generalizing over the language-specific structures discussed in Little, Moroney & Royer (2022: 16), and correspond to what Little, Moroney & Royer (2022: Sec. 4.5) address as the *classifier-for-numerals* and *classifier-for-noun* constructions, respectively.

(69) a.



b.



As summarized by Little, Moroney & Royer (2022: sec. 4.5), the classifier-for-numerals structure in (69a) accounts for facts that they discuss with respect to the Mayan language Ch'ol. Since the sortal classifier is selected by the numeral, (69a) correctly predicts that in Ch'ol there are two subsets of numerals, whose members either obligatorily require or never admit the presence of a sortal classifier when involved in a nominal-noun-construction. This fact is also argued to correlate with the necessary co-occurrence of numerals with a sortal classifier whenever the former are cited in mathematical operations and the like.

The two properties that (69a) captures are argued to be utterly excluded by the counterpart in (69b), and the latter vice versa predicts two patterns that are found in the classifier-for-nouns language Shan (Kra-Dai family, spoken in Myanmar) but not in Ch'ol. In languages like Shan, the presence of the classifier is not conditioned by the numeral that is involved, but rather depends on the noun (i.e. there is a subset of nouns that require the presence of classifiers to combine with numerals, whereas another subset excludes it). In a related manner, classifers co-occur with the nouns that require them also if no numeral is present in the nominal structure.

Aside from the patterns just discussed, neither of the two structures in (69) is a good compromise that manages to capture most of the generalizations holding crosslinguistically with respect to sortal classifiers. On the one hand, even though Little, Moroney & Royer (2022: Sec. 3.4) observe that (69b) is at odds with Borer's (2005) analysis of the GSSG, they argue that it represents a good result, as it manages to account for the fact that number marking and

sortal classifiers can co-occur in Ch'ol. However, an argument frequently made with respect to the GSSG is that it better holds in languages where number marking or sortal classifiers are mandatory (see Kim & Melchin 2018, Tang & Her 2019, Doetjes 2021a), and, crucially, the plural marker that Little, Moroney & Royer (2022: 19) discuss for Ch'ol is optional. On the other hand, (69a) is in line with the claims made by Her (2017), Allassonnière-Tang & Her (2020), and Cinque (2023) as a consequence of the following facts, which are summarized based on Cinque (2023: 14-16) (see also Tatsumi 2020, 2021 for related discussion):

(70)	a.	Num CLF N	multiplier base N	Card Ord N
	b.	N CLF Num	N base multiplier	N Ord Card
	c.	N Num CLF	N multiplier base	N Card Ord
	d.	CLF Num N	base multiplier N	Ord Card N
	e.	*CLF N Num	*base N multiplier	*Ord N Card
	f.	*Num N CLF	*multiplier N base	*Card N Ord

The patterns in (70) are concerned with cardinal numerals and sortal classifiers, with multipliers and bases (e.g. *four hundred*) and with cardinal and ordinal numerals (e.g. *the first two days*). In all these constructions, the lexical noun is always found at either one of the edges, so that a tendency towards some sort of harmonization across patterns might be at play. In fact, Allassonnière-Tang & Her (2020) claim to have obtained statistically significant data indicating that classifiers, multipliers and bases harmonize (see Allassonnière-Tang & Her 2020: 515-516):

(71)	a.	[CLF [base multiplier] N]	[[multiplier base] CLF N]
	b.	[N CLF [base multiplier]]	[N [multiplier base] CLF]

The facts observed above are neatly captured as soon as one assumes that a cardinal numeral and a sortal classifier that co-occur form a constituent to the exclusion of the noun, as in the classifier-for-numeral structure in (69a). Still, as a concluding remark it is worth mentioning that none of the patterns in (70) and (71) are concerned with additive complex numerals, which have in turn the potential to further complicate the landscape related to numeral constructions. In fact, in section 3.1.1 it has been shown that a demonstrative and a reinforcer can co-occur on opposite sides with respect to the noun, and the same thing can happen with additive complex numerals. The example in (72) follows Willis (2006: 1828)

in showing that in Welsh the postnominal part of the numeral is a PP, as the literal translation of *tair ar hugain* is 'three on twenty'.

(72) *y tair merch ddawnus ar hugain hyn*
DET three girl talented on twenty DEM.PROX
'These twenty-three talented girls'

3.1.3 On adjectives

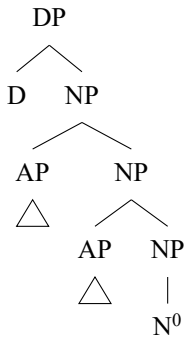
Given that Greenberg's U20 focuses on one single descriptive adjective modifying the noun, many of the concerns that are debated with respect to the syntax of adjectives do not represent a central issue for the present work. In fact, U20 shifts the focus away from adjectives in their predicative use (Kotowski 2016: 23 ff.) and in their gradable and comparative forms (Morzycki 2016: Ch. 3-4, Kotowski 2016: 61 ff.). Moreover, since one single adjective is considered, the importance of the interaction of multiple adjectives with one another is diluted (Scontras, Degen & Goodman 2017, Dyer et al. 2023, a.o.). However, given that there is little consensus on the formal structures that better capture the facts linked to adjectives, the discussion to follow will necessarily touch upon some of these aspects.

Giusti (2018: 3) and Roehrs (2020: 545) illustrate (among others) that three main structural analyses have been proposed for adjectives. They can be analyzed as adjoined to the nominal domain (Jackendoff 1977), or as heads selecting nouns as their complements (Abney 1987), or as phrases that are base-generated in the specifier of a functional projection (Cinque 1996, 2010). Moreover, some scholars assume that certain adjectives have a head status while others have a phrasal status (see Alexiadou 2014: 91-92, 95 and references therein), similarly to what has been discussed in section 3.1.2 following Danon (2012).

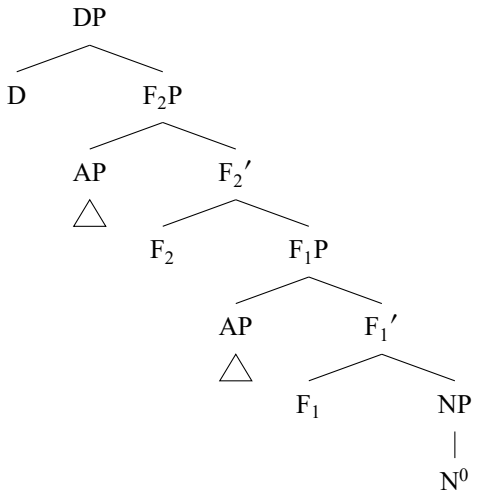
At first sight, the adjunction analysis in (73a) is rather appealing because it accounts for the optionality of adjectives and their stackability (see Svenonius 1994, a.o.). However, the main problem of such an account is that it predicts a high variability with respect to the order among multiple adjectives modifying the same noun, contrary to fact. As Sproat & Shih (1988: 465) observe, a phrase like *a nice round plate* is perfectly natural in English, whereas *a round nice plate* is cumbersome. Leivada (2022: 11 fn. 1) points out that this phenomenon is addressed as Adjective Order Preferences (AOPs) or as Adjective Order Restrictions (AORs), depending on one's stance with respect to how deviant the departures from the unmarked

order are.

(73) a.



b.



Given that adjectives are first merged as specifiers of distinct functional projections, the alternative in (73b) seems better suited to account for a non-arbitrary order of adjectives. Moreover, Cinque (1994: 94) argues that (73b) is superior to (73a) because it further allows to set up "a clear limit on the number of non-coordinated APs". As Cinque (2010: 48) points out, (73b) correctly expects that any two adjectives belonging to the same class should not co-occur in non-coordinated contexts, because they are supposed to target the same Spec,FP. However, Kotowski (2016: 38-39, 144-148) objects that both the exact number of fuctional projections allowed and the semantic classes aimed at restricting the co-occurrence among adjectives must be assumed in an arbitrary and *ad hoc* fashion.

Still, one can follow Alexiadou (2014: 92) in observing that the constrast between English and French in (74a) is easy to capture as soon as one assumes that in French N^0 undergoes head movement as per Longobardi (1994). However, although "[o]ne empirical prediction of the N-movement analysis is that the order of DP-internal adjectives should remain constant cross-linguistically" (Alexiadou 2014: 101), the data in (74b) provided by Laenzlinger (2017: 15) falsify this prediction. In fact, postnominal French adjectives like those in (74b) are in the mirror-image order with respect to their English counterparts.

(74) a. *le* *livre* *rouge*
DET.M.SG book.M.SG red
'The red book'

b. *la* *voiture* *américaine* *rouge* [N ≻ Adj_2 ≻ Adj_1]
DET.F.SG car.F.SG American.F.SG red
'The red American car' [Adj_1 ≻ Adj_2 ≻ N]

In order to account for cases like (74b) in a manner that complies with antisymmetry (Kayne 1994), Cinque (1996, 2010) adopts a strategy that has been anticipated in section 2.4.1 and that will be central to his analyses of U20 (to be discussed in sections 3.2.1 and 3.2.2). He assumes the presence of an additional silent functional projection above each adjective. As a consequence, the structure is endowed with additional specifiers that can be targeted by maximal projections whenever the latter undergo the roll-up movement discussed in section 2.3.1. As (75) shows, an additional depart from (73b) is represented by the fact that the head raising of N is no longer assumed, as it is replaced by the phrasal movement of NP (see Cinque 2010: Sec. 3.6 for further discussion). In order to obtain what (74b) shows for French, the NP must first target Spec,F_2P in (75), and then the whole structure dominated by F_2P further raises to Spec,F_1P.

(75) $[_{DP}$ **D** $[_{F_1P}$ $\overset{\uparrow}{[}$ F_1 $[_{XP}$ **A_1P** $[X$ $[_{F_2P}$ $\overset{\uparrow}{[}$ F_2 $[_{YP}$ **A_2P** $[Y$ $[$ **NP**]]]]]]]]]]

$[_{DP}$ **D** $[_{F_1P}$ $[_{F_2P}$ **NP** $[$ F_2 $[_{YP}$ **A_2P** $[$ Y ~~NP~~]]]] $[$ F_1 $[_{XP}$ **A_1P** $[$ X ~~F_2P~~]]]]]

As sections 3.2.1 and 3.2.2 will show, the premises leading to the derivation in (75) are essentially the same that allow Cinque (2005, 2023) to derive the attested permutations of U20. However, in section 2.4.1 it has been pointed out that according to Abels (2007) the assumption of additional functional projections weakens the restrictive power of antisymmetry. Similar concerns are also caused by the shift from N^0- to NP-raising, as Dékány (2018: Sec. 3.4.2) observes that phrasal movement is not subject to the locality conditions that are traditionally imposed on heads by the Head Movement Constraint (Travis 1984). In addition to these issues (and to those discussed a.o. by Pereltsvaig 2006, Willis 2006), Norris (2017b: 3) observes that nominal concord among the noun and its modifiers is usually assumed to motivate the fact that (74b) is mapped from (75). However, this claim is theoretically troublesome for reasons that also apply to the N^0-raising posited in (73b) on the preceding page. Norris (2017b: 3) states the problem as follows:

> Let us assume for the moment that this [N^0-]movement puts the head into the

> requisite syntactic relationship with the adjective to allow for agreement to occur. So, for example, when N^0 moves to $F_1{}^0$ [in (73b)], the adjective in Spec, F_1P can agree with it, and so on. [...] However, N^0 does not move all the way to the top of the nominal extended projection in every language. Thus, something additional must be said to account for concord on adjectives (and other elements) that are not specifiers of one of N^0's landing sites [...]. Cinque (1994) suggests that this agreement can be registered by movement of the N^0 and subsequent feature-checking at Logical Form (LF). However, this solution is assumed rather than argued for, and it is not clear how it could be supported empirically. Furthermore, the agreement relation necessary for this N^0-movement-based account is, in any case, unclear. The relation is akin to the Spec/Head configuration, which has been argued to underlie agreement (especially subject-verb agreement), but it is actually reversed. In a Spec/Head analysis of subject-verb agreement, the head (e.g., I^0) acquires ϕ-features from the subject in its specifier. In the Spec/Head analysis required for this kind of analysis of concord, the adjective in specifier position (e.g., of F_1P) would acquire its ϕ-features from the functional head [...].

In addition to the aforementioned problems, the assumption of a hierarchy of functional projections that orders adjectives by dividing them into rigid classes as per Cinque (1994) and Scott (2002) is at odds with empirical evidence discussed in Scontras (2022, 2023) and Dyer et al. (2023), which suggests, on the contrary, that "there is no single source of ordering preferences" (Scontras 2023: 372). This does not necessarily imply that the hierarchies discussed above are not pointing in the right direction, but still "some other factor(s) must be at play in delivering the specific orderings observed" (Scontras 2023: 370).

By using technologies and datasets from natural language processing, Dyer et al. (2023) focus on triples consisting of two adjectives modifying a head noun across 32 distinct languages. The authors classify such triples "into three templates: noun-final (AAN), noun-medial (ANA), and noun-initial (NAA)" (Dyer et al. 2023: 1186), and find confirmation of the existence of AOPs crosslinguistically (Dyer et al. 2023: Sec. 5.1). However, the single best predictor of the observed patterns changes across the three templates, so that postnominal adjectives do not seem to be linearized in a sequence that is either consistent with, or a mirror-image of, prenominal adjectives. In AAN languages, the best predictor turns out to be subjectivity (a result in line with Scontras, Degen & Goodman 2017), with less subjective

adjectives appearing closer to the noun (see Dyer et al. 2023: 1188 for further discussion). In the ANA template, the prenominal adjective tends to be a shorter word than the postnominal counterpart, so that word length emerges as the decisive factor in this case (Dyer et al. 2023: 1187, 1194). The results for the remaining template lead Dyer et al. (2023: 1194) to state that "it may be the case that adjectives in NAA languages are less likely to be used for the purpose of establishing reference". This supposition stems from the fact that the single best predictor in the latter case is what Dyer et al. (2023: 1188) define as *pointwise mutual information* (PMI), which "quantifies the information that adjectives and nouns have in common on the basis of the extent to which they occur together [...]; adjectives with higher PMI with the modified noun should appear closer to the noun".

In any case, it must be pointed out that the syntactic model proposed by Cinque (2010) does not impose one unique order for *any* two or more adjectives of a single language. What is more, Cinque (2010: Sec. 3.5.1) also expects that some languages may lack AORs altogether. These possibilities follow from the fact that he adopts a dual source for adjectives (see Cinque 2010: Sec. 3.1, Alexiadou 2014: Sec. 2, 4.1 and Kotowski 2016: 25 ff. for related discussion).[44] The first source corresponds to adjectives that are merged close to the noun in the way shown in (75) on page 102. In this structural context, they are assumed to comply with the AORs by being instances of what Sproat & Shih (1988) call *direct* modification.

Reduced relative clauses represent instead the second source for adjectives assumed by Cinque (2010: 63), and they are merged in the structure farther away from the noun. Their assumption is expected to lead to what Sproat & Shih (1988) call *indirect* modification, where adjectives can be ordered way more flexibly. Sproat & Shih (1988: 477-478) argue that in cases like (76) this follows from the fact that there is no hierarchy at play among adjectives, but rather parallel structures "oftentimes accompanied by comma intonation and/or overt conjunctions" (Kotowski 2016: 126). However, Cinque (2010: 31) argues that the alternations in (76) do not follow from flat structures, but rather from the fact that each adjective is contained in a reduced relative clause.[45]

[44] As observed in Scontras (2023: 363), the dual source of adjectives was already proposed in Cinque (1994: 29), but with a different goal in mind. Given that Cinque (1994) assumed a structure like the one in (73a) on page 101, the problem to circumvent was the mirror-image order that certain postnominal adjectives display in Romance as opposed to the prenominal counterparts in English. Before adopting the antisymmetric analysis summarized in (75) by following Cinque (2010), the dual source of adjectives was thus used to argue that a predication structure could account for the mirror-image order of postnominal adjectives.

[45] Focal stress on the first adjective (discussed in Cinque 2010: 58-59 and Kotowski 2016: 129, a.o.) is somewhat related to parallel modification, as it also allows to override the AOPs. However, Cinque (2010:

(76) a. She loves all those Oriental, orange, wonderful ivories.
b. She loves all those wonderful, orange, Oriental ivories.
c. She loves all those orange, Oriental, wonderful ivories.
d. She loves all those wonderful, Oriental, orange ivories.

Although the proposal in Cinque (2010) provides interesting insights, there is a growing body of evidence which suggests that the assumptions made there should be reworked (at least partially). In particular, a claim that turns out to be problematic is that direct modification should be associated with structured adjectives (that comply with AORs) while indirect modification should correlate with (reduced) relative clauses. For a start, these dichotomies lead Cinque (2010: 99 ff.) to make certain predictions with respect to adjectives in Bosnian/Serbo/Croatian that according to Talić (2017) find little or no confirmation. Similarly, Cinque expects Greek prenominal adjectives to be rigidly ordered when not involved in polydefinite constructions (see Cinque 2010: 105 ff. and references therein for discussion), but Leivada & Westergaard (2019) and Leivada (2022) put this claim to question as well.

It is also relevant to observe that languages lacking AORs are expected to only have adjectives that are reduced relative clauses (Cinque 2010: Sec. 3.5.1), but Brown (2024) challenges this prediction by providing data from the Papuan language Naasioi, where direct modifiers paradoxically bear no AOR. Following Cinque, the lack of AORs predicts adjectives in Naasioi to be reduced relative clauses, but the empirical facts do not suggest that. Beyond occuring on their own, adjectives in Naasioi can be included in reduced or non-reduced relative clauses that are overtly signaled by the presence of a determiner or a relativizer, respectively (Brown 2024: 430-431). Moreover, the order between adjectives per se and (reduced) relative clauses is not free, which is something that is expected only if the former were direct modifiers (Brown 2024: 432).

There is one last objection that Brown (2024: 432) raises. Cinque (2010: Sec. 3.4) benefits from the opposition between direct and indirect modification to argue that the two correlate with opposing values of other semantic and syntactic dichotomies, as already seen for example with respect to stacked adjectives as opposed to (reduced) relative clauses. Among other splits that will not be discussed here, Cinque (2010) follows Sproat & Shih (1988) in claiming that indirect modifiers can only bear intersective readings (see Alexiadou 2014: 98

77) argues that in at least some of these cases a FocusP is involved, so that they can receive a special treatment in the syntax.

for further references). As Morzycki (2016: 14-16) clarifies, intersectivity is the set-theoretic notion involved in the interpretation of a phrase like *famous Canadian surgeons*, so that people fitting this description are at once included into three distinct (but, crucially, intersecting) sets, whose members are respectively famous, Canadian and surgeons.[46] A rather clear example of non-intersectivity is represented instead by an adjective like *alleged*, as *an alleged murderer* does not necessarily belong to the set of murderers, and the assumption of a set of 'alleged individuals' is problematic as well (Morzycki 2016: Sec. 2.2.4). Given that Brown (2024: 432) reports the existence of nonintersective adjectives in Naasioi, the latter must be direct modifiers, *contra* what Cinque (2010) predicts.

The adjectives that Cinque (2010) classifies as nonintersective can be understood as members of the complement set of intersective adjectives. However, Morzycki (2016: 18) highlights that the notion of intersectivity hinges on "a single, testable hypothesis about what adjective–noun combinations mean", but the same does not hold for non-intersectivity. In fact, Morzycki (2016: Ch. 2) and Kotowski (2016: 45) observe that the class of adjectives with a nonintersective reading is heterogeneous, and problems do not stop by further distinguishing the readings of such adjectives according to notions like privativity or (non-)subsectivity (see Alexiadou 2014: Sec. 3.5, Kotowski 2016: 42 ff. and references therein).[47] In English and German (among other languages), the former defines a closed class with peculiar syntactic and semantic behaviors (Kotowski 2016: 46, Morzycki 2016: Sec. 2.2.5, 2.4.4), while the latter does not help define a non-heterogeneous class of nonintersective adjectives (Kotowski 2016: 46-52, 55-61, Morzycki 2016: Sec. 2.3, 2.4).

The notion of intersectivity does not come without problems, either. On the one hand, this is due to the fact that there exist adjectives like *big* and *small* which are at first sight intersective but bear some typically subsective properties (see Morzycki 2016: Sec. 2.2.3, 2.4.4). On the other hand, Svenonius (1994: 450-451) observes that two adjectives modifying a noun receive different readings depending on their order. Although Kotowski (2016: 129-131) argues that this phenomenon is orthogonal to the study of AORs in the narrow sense, it is worth reporting in (77) one notable contrast that Svenonius provides in this respect (and which is further discussed in Pereltsvaig 2006: Sec. 2, Kotowski 2016: 129-130, Morzycki

[46] See however Kotowski (2016: 59-60) for discussion of the problems raised by a classification of ethnic adjectives as intersectives, and Morzycki (2016: 41-42) for related discussion.

[47] Morzycki (2016: 18) remarks that the term *subsective* is used in this context to mean *subsective but not intersective*. This clarification becomes necessary once one observes that "all intersective modification is subsective as well", given that "the intersection of two sets is a subset of both of them" (Morzycki 2016: 18).

2016: 52-53, a.o.).

(77) a. A dead dangerous animal (i.e. a stuffed lion)
b. A dangerous dead animal (i.e. an infected sheep)

When *dangerous* precedes *dead* as in (77a), an intersective reading is preferred, so that the denoted entity must be dangerous independently of its dead state. Therefore, (77a) is felicitous when referred to a lion. When the order between the two adjective is instead reversed as in (77b), the danger is linked to the current dead state of the animal, so that (77b) felicitously applies even to otherwise harmless entites (e.g. a sheep). The contrast in (77) will be discussed further in section 5.2.1, as relevant insights in this respect can be provided by Tree Adjoining Grammars.

The beginning of the present section has followed Giusti (2018: 3) and Roehrs (2020: 545) in observing that three main structural analyses have been proposed for adjectives, but only two of them have been discussed so far. The last one consists in analyzing adjectives as heads selecting the noun as their complements. As observed by Cinque (2010: Sec. 4.1.2) and Alexiadou (2014: 91-92), this claim has often been supported among works arguing that prenominal adjectives are heads whereas the postnominal counterparts are phrases. Abney (1987) proposes that all adjectives in English are heads, as they are by and large prenominal (though see Cinque 2010: Ch. 5) and Abney (1987: 126) and Kotowski (2016: 117) show that they are constrained as in (78) when they select a complement:

(78) a. the proud (*of his son) man
b. the man proud *(of his son)
c. a <*of bananas> full <*of bananas> truck
d. a truck full *(of bananas)

Alexiadou (2014: 96) follows Svenonius (1994) in observing that treating all (prenominal) adjectives modifying a noun as heads is problematic, both language internally and crosslinguistically. In *some barely hot black coffee*, *barely* modifies the adjective *hot*, but not *black* (as the coffee can be barely hot, but completely black). This is unexpected if both adjectives were heads, as the degree adverb should scope over both. This does not necessarily mean, however, that no adjective can be a head (as admitted by Cinque 2010: Sec. 4.1.2 and Alexiadou 2014: 96).

Alexeyenko & Zeijlstra (2021b: Sec. 3) and Haider (2022b: Sec. 2) provide good summaries of the many proposals that have been put forward so as to capture (78). For a start, one can observe that in German such restrictions do not apply, as shown in (79) by following Kotowski (2016: 117). In view of what will be discussed later, it is worth highlighting that attributive adjectives in German agree in case, gender and number with the noun they modifiy. Moreover, the same asymmetry between English and German that (79) illustrates also emerges with comparative adjectives (see Haider 2022b: 202-203 for examples) and, interestingly, adverbials. The example in (80) follows Haider (2022b: 215) in reporting a contrast between English and German with respect to the latter.

(79) a. *ein mit Bananen voll-er Laster*
INDF with banana.DAT.PL full-NOM.M.SG truck
'a truck full of bananas'

b. *ein auf seinen Sohn stolz-er Vater*
INDF on his.ACC son proud-NOM.M.SG father
'a father proud of his son'

(80) a. he has more often (*than anyone else) scored.

b. *er hat viel öfter (als jeder andere) gepunktet*
3SG.M.NOM AUX much more.often (than anyone else) score.PTCP
'He has scored much more often than anyone else.'

Haider (2013: 13-16, 34-37, 2022b) explains the patterns in (78), (79) and (80) as a corollary of his syntactic model, previously discussed in section 2.4.2. He argues that if adjectives and adverbials in such examples were phrasal specifiers of functional heads, similar restrictions should also apply to verbal and nominal arguments in specifier positions, contrary to fact (Haider 2022b: 225). Hence, adjectives and adverbials in (78), (79) and (80) are argued to be phrases adjoned to the left, because the (Basic) Branching Constraint bans any right counterpart of this operation.

The asymmetries between English and German are then argued to follow from the canonical directionality in which lexical heads (i.e. adjectives, nouns, verbs, adverbs) license their dependents. English is a uniformly head-initial language (Haider 2022b: 203), so that what is observed in (78) and (80a) follows from the fact that left adjuncts (i.e. adjectives and adverbials) "are outside of the structural licensing domain of the head of the phrase and therefore

they are in need of an alternative way of getting structurally licensed" (Haider 2022b: 199). Such an alternate manner is represented by *proper attachment*, which according to Haider (2022b: 226) states that there must be no intervening projection in the linearization between the head of the adjunct and the nominal (or verbal) projection, so that the two end up being mininally distant from one another. Hence, *of his son* and *than anyone else* are respectively ungrammatical in (78) and (80a) because they would constitute intervening projections between the nominal or verbal projection and the head of the adjoined phrase.

Haider (2022b) calls what has just been discussed for the English examples in (78) and (80a) as the *Left-Left Constraint* (LLC), while in former works it is called as the *Edge effect* (e.g. Haider 2013: 13-16, 34-37). This kind of syntactic restriction is expected to extend to other uniformly head-initial languages like Romance and North-Germanic (Haider 2022b: 203), and in fact Haider (2022b: Sec. 1, 2.5) provides data in support of this claim. In the German examples in (79) and (80b), no problem arises because adjuncts are head-final, so that no projection can intervene between their heads and the nominal or verbal projections (see Haider 2022b: 202-203 for examples supporting this claim). Moreover, Haider (2022b: 200) points out the following (italics in the original):

> In German and Dutch [...] and all other Germanic OV languages, the LLC applies only to adjuncts of NPs but not to adjuncts of VPs or APs. This is a predictable fact since NPs are head-initial while VPs and APs are head-final phrases in these languages.[...] Since the LLC is a constraint on left-adjunction to *head-initial* phrases, VPs and APs are not in the scope of this constraint. Consequently, in uniformly head-final languages – Japanese, for example – there is no context at all for the LLC to apply.

Thus, the LLC is not expected to apply to languages of the third type either, as they do not impose a fixed canonical directionality on their lexical heads. And in fact, Haider (2022b: 205-206) and Haider & Szucsich (2022b: Sec. 2.4, 2022a: Sec. 6) argue that this prediction is confirmed as well. As a result, the LLC constrains only the syntax of head-initial phrases adjoining to head-initial projections, because their heads bear a canonical directionality which is not "in harmony with the branching direction" assumed for the whole syntactic structure (Haider 2013: 5). To effectively realize the complex attributes that are constrained by the LLC, languages affected by the latter can either realize them postnominally, or extrapose the phrase that would cause the fatal intervention (Haider 2022b: 203; see also Sheehan 2017:

123-125, 131-135, Alexeyenko & Zeijlstra 2021b: Sec. 1.2 for related remarks). These two rescue strategies are respectively illustrated in the examples in (81) and (82), taken from Haider (2022b: 204):

(81) a. *une fière* (* *de soi*) *femme*
INDF.F.SG proud.F.SG (* of REFL) woman.F.SG
b. *une femme fière de soi*
INDF.F.SG woman.F.SG proud.F.SG of REFL
'A woman proud of herself.' [French]

(82) a. *ein so seltener* (* *wie der Apollofalter*) *Schmetterling ist der*
INDF as rare.AGR (* as DET Apollo.butterfly) butterfly COP DET
Segelfalter
scarce.swallowtail
b. *ein so seltener Schmetterling* (*wie der Apollofalter*) *ist der*
INDF as rare.AGR butterfly (as DET Apollo.butterfly) COP DET
Segelfalter
scarce.swallowtail
'A butterfly as rare as the Apollo butterfly is the scarce swallowtail.' [German]

However, the LLC proposed by Haider (2022b) is seriously challenged by the data from Basque. The discussion in Artiagoitia (2010) suggests that the nominal domain is head-final in Basque, as the noun licenses its arguments "in harmony with the branching direction" (Haider 2013: 5). This is shown by means of the example in (83a), taken from Artiagoitia (2010: 5) (see also Artiagoitia 2008: 89). However, Artiagoitia (2006: 115, 2008: 87) illustrates that attributive adjectives are postnominal but "cannot take lexical complements", as reported in (83b). Thus, the latter restriction (also pointed out in Alexeyenko & Zeijlstra 2021b: Sec. 2.2 and Alexeyenko & Zeijlstra 2021a: 52 by means of another example) is similar to what the LLC captures, but crucially involves an attributive projection which is neither prenominal nor combined to a head-initial nominal projection. As Hualde & Ortiz de Urbina (2003: 138) point out, the only grammatical counterpart of (83b) requires the use of a relative clause, as shown in (83c).

(83) a. *Velazquez-en Felipe erregear-en bost erretratu-ak*
Velazquez-GEN Felipe king-GEN five portrait.ART
'The five portraits of King Felipe by Velazquez'

b. *emakume* (* *bere lan-az*) *harro-a*
woman (* her work-INS) proud-ART
'The/a woman proud of her work'

c. *bere lan-az harro-a dagoen emakume-a*
her work-INS proud-ART is.COMP woman-ART
'A woman who is proud of her work'

In order to provide an account for the facts discussed so far with respect to adjectives modifying nouns, Alexeyenko & Zeijlstra (2021b,a) do not focus on potentially similar restrictions holding for adverbials, but rather bring the attention to the form that complex adjectival phrases have in their predicative use. Thus, they argue that the facts discussed above can be captured by means of what they call MAG, i.e. the Modifier-noun Adjacency Generalization (Alexeyenko & Zeijlstra 2021b: Sec. 4, italics in the original):

(84) The linear order A-XP-N (or N-XP-A, if adjectives are postnominal) is possible iff:

a. the linear order A-XP (or XP-A) is available in the predicative position, *and*

b. the adjective modifying the noun exhibits:
(i) an overt attributive marker, which is a morpho-phonologically independent clitic/free word form, *or*
(ii) an agreement marker that is specified for all φ and case (κ) features of N *and* is also present on predicative adjectives.

Although both Haider (2022b) and Alexeyenko & Zeijlstra (2021b,a) try to account for overlapping phenomena, they diverge in using adverbials and predicative adjectives as supporting evidence for their claims. This difference is in turn due to the fact that agreement features are pivotal in the definition of MAG in (84b), while they play no role in the definition of the LLC. In fact, the LLC is exclusively grounded on structural terms because adverbials are inert with respect to agreement (Haider 2022b: 225).

The requirement in (84a) establishes that the predicative use of a complex adjective can be taken as a litmus test to observe whether in a language a phrasal projection has the potential

to interpose between the adjective and the noun it modifies. The linear adjacency between an adjective and a modifying noun is then expected whenever the premise in (84a) does not hold. However, English, German and Basque satisfy (84a), as they do show phrasal projections in predicative position that would beak the adjacency between the adjective and the modified noun in the corresponding nominal constructions, as shown in (85). By means of (85b) Alexeyenko & Zeijlstra (2021a: 56) thus highlight that in German an adjective in predicative position can alternatively precede or follow the XP it selects.

(85) a. the father is proud of his son.

b. *er ist <auf seinen Sohn> stolz*(**-er*) *<auf seinen Sohn>*
3SG.M COP <on his son> proud-AGR <on his son>
'He is proud of his son.' [German]

c. *bere lan-az harro-a dago*
her work-INS proud-ART COP
'She is proud of her work.' [Basque]

Given that (84a) is proven to hold for English, German and Basque, one can further discuss why the linear adjacency between an adjective and a modifying noun applies to these languages. For English and Basque, this is a consequence of (84b-i), as the assumed attributive marker is not "morpho-phonologically independent" (Alexeyenko & Zeijlstra 2021b: 25), i.e. it is either an affix or a null morpheme (Alexeyenko & Zeijlstra 2021b: Sec. 6.2-6.3). Alexeyenko & Zeijlstra (2021b: 29) explicitly argue in favor of a null morpheme for English, and the same should also apply to Basque (although they do not discuss it).[48]

The situation is different with respect to German, since (84b-ii) becomes relevant due to the presence of "fully marked attributive adjectives" in the language (Alexeyenko & Zeijlstra 2021b: 18), as previously seen in (79) and (82). In German, the adjacency between an adjective and the modified noun follows from the fact that the second conjunct in (84b-ii) is not satisfied, as (85b) shows that predicative adjectives in German must be bare. In this respect, Alexeyenko & Zeijlstra (2021b: 8, 18) put German in constrast with (Modern) Greek,

[48] In Basque, "any number-marking or case-marking which is present appears only once, at the end of the noun phrase" (Hualde & Ortiz de Urbina 2003: 113; see also Hualde & Ortiz de Urbina 2003: 171 ff. and Artiagoitia 2008: 81 for further discussion). Hence, adjectives bear no agreement markers, as "there is no NP-internal agreement" (Hualde & Ortiz de Urbina 2003: 113). Sex marking on adjectives is an exceptional feature of Spanish loanwords, and it is only attested in the Western varieties (Hualde & Ortiz de Urbina 2003: 137).

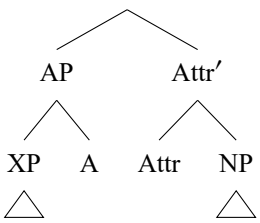
because the latter fully satisfies (84a) and (84b-ii), so that the adjacency between a noun and an attributive adjective is not required.

(86) a. *o* *perifan**(*-os*) *pateras*
DETM.SG.NOM proud-M.SG.NOM father.M.SG.NOM
'the proud father'

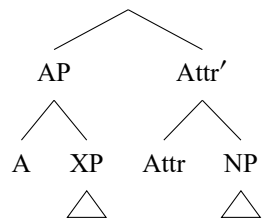
b. *afto* *to* *simantiko* *gia* *tis* *arxes*
DEM.PROX.N.SG.NOM DET.N.SG.NOM important.N.SG.NOM for DET authorities
zitima
issue.N.SG.NOM
'This important issue for the authorities'

Given the empirical differences treated so far, it is now possible to highlight that the definition of MAG in (84b) accounts for a single syntactic phenomenon by means of two different sets of structures, which are reported in (87) and (88) by following Alexeyenko & Zeijlstra (2021b: 23-24, 2021a: Sec. 3.2). The representations in (87) correspond to (84b-i), and the null morphemes that Alexeyenko & Zeijlstra (2021b: Sec. 6.3, 2021a: Sec. 4.2) assume for English (and Basque) are instances of the Attr head shown there. However, neither (87a) nor (87b) are valid structures for English attributive adjectives, and the same applies to the mirror image of such structures when it comes to Basque. English rules out (87a) because in predicative position no instance of XP-A can be observed, and the same would hold for Basque with respect to the mirror image of (87a), which would map N-A-XP. At the same time, English cannot instantiate (87b), because the null morpheme requires to be linearly adjacent with A; ditto for Basque with respect to the mirror image of (87b).

(87) a.



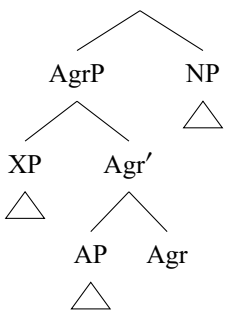
b.



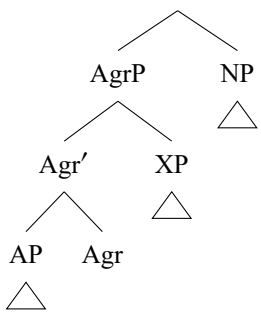
Alexeyenko & Zeijlstra (2021a: Sec. 4.1) argue that what has been discussed for German indicates that the language can realize (87a), but not (87b). In fact, the attributive marker cannot be part of the phrasal projection in which the adjective is included, as this would predict

the presence of such a marker also in predicative contexts, contrary to what has been observed in (85b). The contrast with the Greek facts in (86) follows instead by arguing that attributive adjectives in Greek instantiate (88b), wheras (88a) is not an alternate option because of the lack of language-internal evidence required by MAG in (84a). Due to theory-internal reasons, Alexeyenko & Zeijlstra (2021b: 23, 2021a: 55) argue that in (88) the adjacency between the A and Agr is trivially satisfied because the latter is the transitive grammatical category that relates A with XP (Wiltschko 2014: 311). The latter theoretical assumption also justifies the fact that, as opposed to (84b-i), the MAG in (84b-ii) is not concerned with the morphosyntactic status of the Agr head.

(88) a.



b.



To conclude this section, it is worth highlighting once more that the two sets of structures in (87) and (88) are deemed to complement one another so as to adequately account for a single syntactic phenomenon, i.e. adjacency restrictions between attributive adjectives and nouns. In this respect, the distinct Attr and Agr labels can be somewhat misleading, because this notational difference between the two sets of trees is simply aimed at capturing the fact that Attr is not part of the complex phrase that includes the adjective, and thus cannot occur in predicative contexts (Alexeyenko & Zeijlstra 2021b: 23). However, Attr "may in fact be categorially identical with Agr" (Alexeyenko & Zeijlstra 2021a: 55), or it can even be the only functional head to be assumed in the analysis, as the features that Agr bears "might even be introduced on the lexical A-head itself" (Alexeyenko & Zeijlstra 2021b: 22).

These theoretical premises strikingly parallel the account that Little, Moroney & Royer (2022) provide for the crosslinguistic variation related to sortal classifiers. Similarly to what has just been seen, in section 3.1.2 two sets of structures have been discussed to capture the morphosyntactic behavior of a single nominal modifier in relation to a unique functional head. In section 5.5.2, it will be argued that only one set of structures is needed to account

for these phenomena, but that this possibility crucially becomes possible only by adopting (a version of) Tree Adjoining Grammars.

3.1.4 On nouns

The focus of U20 is mainly directed towards non-proper, non-unique count nouns that do not take argument/adjunct PPs. However, the low merge site of complements to the nominal assumed by the Functional Head Hypothesis (see again section 2.2.4) can be problematic to those proposals that account for U20 by raising either N^0 or NP, as discussed in sections 3.1.1 and 3.1.3. Given that this issue is however very briefly (if ever) touched upon in such works (which will be presented in section 3.2), the present section aims to provide a richer discussion than it is done for instance in Cinque (2005: 327 fn.34), Abels & Neeleman (2012: 65), Roberts (2017b: 164) and Cinque (2023: 123 n. 34).

As section 2.2.4 showed, scholars defending a FHH that sketches a parallel between the nominal and the clausal domains tend to divide the extended projections into three distinct areas, where the lowest is dedicated to the assignment of thematic roles (Longobardi 1994, Grohmann 2003, Laenzlinger 2017, Kim 2019). This position has however been challenged by works like Adger (2013) and Baggio (2021), where it is argued that noun modifiers can be structurally closer to the the noun itself than complement PPs. Two main reasons have been advocated in defence of this claim, namely language-specific scopal readings and a crosslinguistic word order tendency that Adger (2013: 93) defines as *PP peripherality*:

(89) When AP modifiers and PP complements both occur to one side of N inside a noun phrase, the PP is separated from the N by the AP.

According to Adger (2013), nominals bearing arguments are not limited to complex event nominals, whose arguments are likely to be introduced in the structure by some inner verbal projection (see discussion in sections 2.2.3 and 2.2.4), but also encompass relational nouns. In particular, Adger (2013: 61, 96) puts to test the PP-peripherality with examples of the latter type, namely *father* in *the father of the girl*, as well as depiction nominals like *picture* in *the picture of the tree*, property nominals like *color* in *the color of the room*, and nouns denoting part-whole relations like *side* in *the side of the table* or nouns of inalienable possession like body-part nominals (e.g. *head* and *hand*).

The statement in (89) seems confirmed by a first inspection of both head-initial languages,

in examples (90)-(92), and head-final ones in (93).[49]

(90) a. *makemake au i ka lihi lōʻihi o ke pākaukau* [Hawaiian]
like 1SG ACC DET edge long of DET table
'I like the long edge of the table.'

b. *makemake au i ka lihi o ke pākaukau lōʻihi*
like 1SGACC DET edge of DET table lōʻihi
* 'I like the long edge of the table.'
OK 'I like the edge of the long table.'

(91) a. **an dealbh de Mhàiri mòr brèagha* [Scottish Gaelic]
DET picture of Màiri big beautiful

b. *an dealbh mòr brèagha de Mhàiri*
DET picture big beautiful of Màiri
'the big beautiful picture of Màiri'

(92) a. *ani ohev et ha-katze ha-arox shel ha-shulxan* [Modern Hebrew]
1SG like ACC DET-edge DEF-long of DET-table

b. *aʃna nḥib l-ḥaʃiya t-twila l-t-tawla* [Lybian Arabic]
1SG like.1SG DET-edge.F DEF-long.F for-the-table.F
'I like the long edge of the table.'

(93) a. *ami ṭebil-ṭa-r lɔmba dhar-ṭa pɔchondo kori* [Bangla]
1SG table-CLF-GEN long edge-CLF like do

b. *eni-kkə tabl-inte valiya attam isṭam aaṇə* [Malayam]
1SG-DAT table-GEN long edge like COP

c. *masa-nın uzun kenar-ın-ı begen-iyor-um* [Turkish]
table-3SG.GEN long edge-3.POSS-ACC like-IPFV-1SG
'I like the long edge of the table.'

Adger (2013) takes the aforementioned examples as the empirical basis for a revised structure of the nominal domain where PPs that are usually assumed to be merged as complements to the noun are actually merged in a step of the derivation that follows the merger of adjectives. However, it should be noted that two alternate views have been proposed in the liter-

[49] (90) is adapted from Adger (2013: 98), (91) from Adger (2013: 105), (92) from Adger (2013: 102-103), (93) from Adger (2013: 163).

ature so as to rescue the widely assumed structure of the nominal domain while accounting for the word order facts. The first is that the PP appears on the right because it is stranded after some leftward movement involving the selecting nominal, while the second is that PPs modifying a nominal are extraposed by means of rightward movement. The former line of analysis has been proposed in Cinque (1994) and Laenzlinger (2005) and is the most likely line of analysis for those who ban rightward movement altogether (see section 2.4), while the latter has been proposed on independent grounds by Willis (2006), Belk & Neeleman (2017), and Bruening (2022).

Willis (2006: 1812-1815) focuses on the Welsh nominal domain and observes that the PP complement of the head noun has to follow not only adjectives, as per the PP-peripherality in (89), but also the possessor whenever present, as in the examples below from Willis (2006: 1812):

(94) a. *disgrifiad* (*manwl*) *y gyrrwr o 'r ddamwain* [Welsh]
description (detailed) DET driver of DET accident

b. **disgrifiad o 'r ddamwain y gyrrwr*
description of DET accident DET driver
'the driver's (detailed) description of the accident'

Willis (2006: 1815) shows that some mandatory rightward extraposition of the PP complement would manage to account for these facts, but he rejects it because he assumes that extraposition is generally optional, as per the observations in Cinque (1994: 85-87).[50] However, the readers may recall from section 2.3.2 that Biberauer, Holmberg & Roberts (2014: 198 ff.) analyze the postverbal complement CPs of German as a consequence of a mandatory extraposition that is motivated as a FOFC-compliance strategy, and that the Basque data discussed by Duguine (2022) provide evidence for the *ex situ* nature of postverbal complement CPs in both languages. The claim for the existence of mandatory rightward extrapositions is also independently made by Edmiston & Potsdam (2017) with respect to object CPs in the VOS language Malagasy (see also Aissen 1992 on Tzotzil), as the following example adapted from Edmiston & Potsdam (2017: Sec. 4.2) shows:[51]

[50] Willis (2006: Sec. 8) thus concludes that the best alternative to capture the Welsh facts is a structure in the nominal domain where no movement is involved and rightward-projecting positions are allowed instead. Willis (2006) claims that the most natural assumption is to treat such positions as specifiers (a view independently argued for Basque by Artiagoitia 2006, 2008), but that viewing them as adjuncts would be a less radical alternative.

[51] In constrast with Biberauer, Holmberg & Roberts (2014: 198 ff.) and Duguine (2022), the facts discussed

(95) a. *manantena Rabe* [*fa hamono ny akoho Ranaivo*]
hope Rabe COMP FUT.kill DET chicken Ranaivo
b. **manantena* [*fa hamono ny akoho Ranaivo*] *Rabe*
hope COMP FUT.kill DET chicken Ranaivo Rabe
'Rabe hopes that Ranaivo will kill the chicken.'

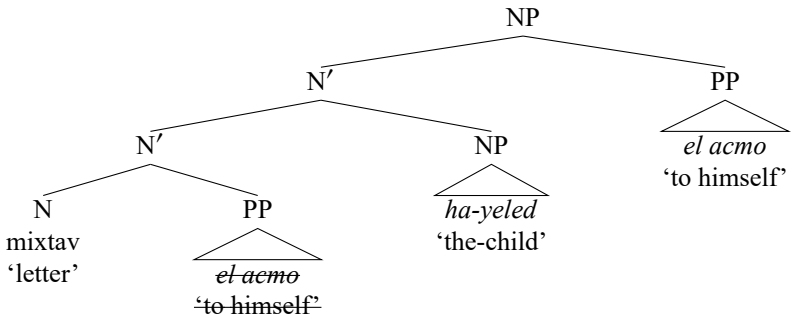
If it is legitimate to claim that the movement operations that are possible in the clause should be paralleled in the nominal domain, as per Laenzlinger (2017), then the aforementioned facts open the door to extraposition phenomena in the nominal domain, such as the rightward movement of complement PPs suggested by Willis (2006: 1815). Although he never mentions Willis (2006), Bruening (2022) also argues for the right extraposition of PPs while discussing the genitival constructions in Modern Hebrew, presented in section 2.2.4 and problematized with respect to derived nominals. In the case of underived nominals, Bruening (2022: Sec. 3) posits that a Goal expressed by a PP is base-generated as a complement of the underived possessum. However, if the latter enters in a Construct State with the possessor, the mentioned PP must undergo rightward raising so as to linearly follow the members of the CS.[52] This way, the possessor can satisfy the prosodic requirement in (96a), already presented on page 40, and effectively form a CS construction, as in (96b). The fact that the possessor can bind into the PP is accounted for by claiming that this operation must hold after reconstruction of the latter to its base position, as the structure from Bruening (2022: 29) in (96c) shows:

(96) a. A [+CS] N must form a prosodic word with phonological material dominated by the NP that checks its [+CS] feature.
b. *mixtav ha-yeled el acmo* [Modern Hebrew]
letter DET-child to REFL
'the child's letter to himself.'

by Edmiston & Potsdam (2017) do not seem *prima facie* to be treatable as a case of FOFC-compliance, because the VP in Malagasy is generally assumed to be head-initial.

52 Here *possessor* and *possessum* are intended in a coarse-grained sense. These terms can be used in ambiguous manners, as in Bruening 2022: 4: "With non-derived nouns that take arguments, like 'picture', only what might be thought of as the internal argument can be the *possessor* in the construct state, an external argument (or "author") or *possessor* may not be" (italics added).

c.



Although Bruening (2022: 28) rejects a view that tries to make parallels between the clausal and the nominal domains (see also section 2.2.4), the fact that both complement CPs and complement PPs are independently accounted by means of extraposition represents a similarity worthy of further research (see also Biberauer 2017a: Sec. 3.2 and Roberts 2017b: 164 on this), especially given the parallel between C and P assumed by Grimshaw (1991, 2005) and discussed in section 2.2.4 following Baggio (2022: Sec. 4.1) and Giusti (in press: Sec. 3).[53]

The CS constructions discussed by Bruening (2022) are also relevant to Adger (2013: 103-105, 129-132, 161-166), because they constitute a language-internal alternative to the examples in (92), in which the Free State construction (see section 2.2.4) is used in its stead.[54] Furthermore, the example in (97), repeated here from page 38 for convenience, is potentially problematic for Adger's claim that adjectives are merged closer to the head nominal than PP complements.

(97) *tmunat ha-xamanyot ha-yafa shel van gox* [Modern Hebrew]
picture.CS DET-sunflowers DEF-beautiful of Van Gogh
'Van Gogh's beautiful picture of the sunflowers'

In (97), the non-prepositional restrictor which is closer to the head nominal forms a CS with it, while the other is still introduced by a preposition, as is typical of FS constructions.

[53] The readers who know works such as Baltin (2006: 264 ff.) or Sportiche (2017) may be tempted to infer that the reconstruction for Condition A required for the PP complement in (96c) represents a further parallel with complement clauses selected by a main verb, which (as opposed to adjunct clauses) need to reconstruct for Condition C. However, Bruening (2018b) argues that extraposed CPs do not reconstruct (as already observed by Haider 2010: 202), because the data that he extensively discusses are better captured by means of an alternative principle to c-command thorougly discussed in Bruening (2014), i.e. *precede-and-command*. Given these two claims, the little space that Bruening (2022: 29 ff.) dedicates to argue for reconstruction in (96c) seems not satisfactory (Haider 2010: 203) and therefore worthy of further scrutiny, especially in light of the criticisms formultated by Witkoś (2021: Sec. 4).

[54] Given that these two constructions can be found in many Semitic languages, Adger (2013) prefers to call the CS as the "Semitic Construct Construction".

With respect to the former, Adger (2013) concludes that PP-peripherality is not at play because it is not realized as a PP, while it operates as expected in relation to the latter, which thus follows the adjectives modifying any of the nouns within the CS.

The analyses of the CS and FS constructions proposed by Adger (2013) and Bruening (2022) share the assumption that the base-generated structure is not optimal for mapping the former. For this reason, the positive feature(s) that the CS bears trigger some movement(s) by means of which the genitival relation between the head of the nominal and its non-prepositional restrictor can be checked under adjacency. However, the two proposals remain incompatible because of the different structures that they assume. Bruening (2022) claims that PPs move to the right to no longer intervene in the CS in (96b) and (97), as (96c) illustrates, while adjectives are not problematic in this respect by being right-adjoined APs (Bruening 2022: 38 ff.). On the contrary, Adger (2013: 161) claims that APs are always interveners while PPs need not be, and that the linear adjacency giving rise to the CS construction is obtained by having the head of the nominal and its non-prepositional restrictor independently raised leftwards, so that they eventually enter the desired local configuration.

In line with both views, Baggio (2021) presents data from Italian in favor of a high structural position of PPs that are *semantic* arguments of an underived noun (Baggio 2021: 393-394). The data from Baggio (2021: Sec. 3) in (98) highlight that the relevant scopal relations cannot be explained if the PP argument is structurally lower than both the privative adjective and the PP modifier, as (98) shows:[55]

(98) *un presunto gioiello di ambra della regina* [Italian]
INDF.M.SG alleged.M.SG jewel.M.SG of amber of.DEF.F.SG queen.F.SG
'an alleged/putative amber jewel belonging to the queen'

a.	?* of amber, of.the queen > alleged > jewel	?*[[[*alleged* N] PP_{MOD}] PP_{ARG}]
b.	of.the queen > alleged > jewel, of amber	[[*alleged* [N PP_{MOD}]] PP_{ARG}]
c.	alleged > jewel, of amber, of.the queen	[*alleged* [[N PP_{MOD}] PP_{ARG}]]
d.	* of amber > alleged > jewel, of.the queen	*[[*alleged* [N PP_{ARG}]] PP_{MOD}]

However, the observations made by Baggio (2021) do not represent a decisive argument in favor of Adger (2013) over Bruening (2022), nor the other way around. The data in (98)

[55] Baggio (2021) further defends his view with data from coordination, ellipsis and from pseudo-partitive constructions. The interested readers are redirected to the original work.

can be reconciled with the structure in (96c) by claiming that in the latter binding applies after reconstruction of the PP to is original low position, while linear order and scopal interpretations are computed with respect to the highest member of the chain. The most problematic aspect for an account of (98) that follows Bruening (2022) is that the movement of PP_{ARG} must be dictated by a requirement which is different from the [±CS] feature bearing a pivotal role in (96a) and in the subsequent contrast between CS and FS constructions.

Although this last fact may seem a point in favor of the account provided by Adger (2013), it might be the case that the high position of PP complements in the structure predicted by the PP-peripherality in (89) be actually a consequence of an adjacency requirement between a noun and the adjective(s) modifying it, so that intervening PPs are ruled out. An attempt in this direction is made by Belk & Neeleman (2017), whose focus is however directed towards the crosslinguistic variation in the linear order of PP modifiers and APs. In particular, they observe that a principle akin to PP-peripherality applies to PP modifiers if both the AP and the PP are postnominal. Given that this does not hold when the same modifiers are prenominal, an asymmetry between the two cases arises, as (99) shows. Belk & Neeleman (2017: 25) then try to capture this observation by means of the "JOIN-First Constraint", whose definition is reported in (100):

(99) PP_{MOD}-AP-N N-AP-PP_{MOD}
AP-PP_{MOD}-N *N-PP_{MOD}-AP

(100) a. The JOIN domain of an AP modifier consists of that AP and any XP intervening between it and the noun.
b. No category lacking JOIN can precede an AP in its JOIN domain.

The statement in (100) leads to two observations. First, (100a) refers to XPs because Belk & Neeleman (2017) claim that the constraint they propose applies to any kind of PP, as well as to nominals in the genitive case. Thus, Belk & Neeleman (2017: 3-5) explicitly argue that PP-peripherality is too strong, if formulated in the symmetric manner given in (89) following Adger (2013: 93). However, the only case where they discuss PPs that meet the criteria chosen by Baggio (2021: 393-394) appears in an unproblematic case from Spanish (Belk & Neeleman 2017: 10), while the N-final languages in (93) that confirm Adger's generalization are overlooked. Therefore, it can be argued that it is better not to conflate PP complements together with other PPs, because keeping them separate has the advantage

of not losing the valuable while addressing the problem.[56] Although this observation is potentially at odds with works like Bruening & Al Khalaf (2019), where it is argued that the distinction between arguments and adjuncts should be dispensed with within nominals (see also Adger 2013: 57 ff.), it should be noted that the latter proposal is not immune to problems either (see Witkoś 2021: Sec. 4 for discussion).[57]

The second aspect that can be pointed out for the "JOIN-First Constraint" in (100) is that linear order is required to be relevant in the syntax, as is proposed by Bruening (2014) and Bruening (2022: Sec. 2.9). In fact, the structure plays a decisive role in ruling out the unattested order in (99) as well as in determining the available scopal readings in a specific nominal (see also Ackema & Neeleman 2002: Sec. 2). As a result, the structural height of APs and PP modifiers can vary freely in N-medial and N-final languages, and this affects both word order and the scopal interpretation, homomorphically in the N-medial case in (52a), isomorphically in the N-final one in (52b).[58] On the contrary, the N-initial case in (52c) is mapped from a right-branching structure (Belk & Neeleman 2017: 28), by means of rightward extraposition (Belk & Neeleman 2017: Sec. 5), as in Willis (2006: 1815) and Bruening (2022: 29), or through a shell formation which explicitly mirrors the same mechanism discussed for the clausal domain in section 2.4.2 (Belk & Neeleman 2017: Sec. 3, 4.1).[59]

(101)	a.	[PP_{MOD} [N AP]]	[[AP N] PP_{MOD}]	[N-medial]
		[AP [N PP_{MOD}]]	[[PP_{MOD} N] AP]	
	b.	[PP_{MOD} [AP N]]	[AP [PP_{MOD} N]]	[N-final]
	c.	[[N AP] PP_{MOD}]		[N-initial]
		[[[N $t_{PP_{MOD}}$] AP] PP_{MOD}]	[N [AP [t_N PP_{MOD}]]]	

[56] This view has the advantage of not making the proposal in Baggio (2021: Sec. 4) necessarily at odds with Belk & Neeleman (2017), *contra* the (arguably overgeneralizing) claims made by both works. See also Neeleman (2017: Sec. 5, 8) and references therein for discussion of multiple PPs in Dutch and English.

[57] Moreover, Manning (2003: Sec. 3.1), McInnerney (2022) and references therein challenge the adequacy of the argument/adjunct opposition in the clausal domain as well, a conclusion which would hinder the non-parallelism between the verbal and the clausal domains proposed in Bruening & Al Khalaf (2019).

[58] "Strictly speaking, an isomorphism is defined by the ability to reverse the mapping and obtain the same structure. [...] Crucially, though, the hierarchical relations are never reversed in homomorphic linearisations." (Martin et al. 2020: 4 fn. 2).

[59] A caveat is worth mentioning. In order to account for certain scopal readings in Spanish, Belk & Neeleman (2017) claim that if an adjective A_1 precedes another adjcective A_2, [N [A_1P [PP [t_N A_2P]]]] can be a legitimate structure. However, in such a case the rightmost adjective obeys (100) by making "reference to the nominal trace" (Belk & Neeleman 2017: 36). This latter claim threatens their whole proposal, because by the same line of reasoning [N [PP [t_N AP]]] would also be legitimate, thus contradicting the empirical premise in (99).

In conclusion, the ideas that PPs that are semantic arguments (Baggio 2021: 393-394) are necessarily generated low in the structure and that they must be moved leftwards together with the noun heading the nominal are both empirically challenged and needy of further investigation in the future. Although the aspects that have been discussed throughout this section will no longer be central in the remainder of the present work, the readers should bear in mind that the proposals that try to account for the variation of U20 (to be discussed in the next section) by banning rightward movement altogether are hard to reconcile with the facts observed and analyzed by Willis (2006), Adger (2013), Belk & Neeleman (2017), Baggio (2021), and Bruening (2022). This problem holds especially if the movement of NP together with its complement PP is subsumed. However, leftward N- or NP-movement may not be necessarily off track either. Rather, one may try complement them with what has been discussed throughout this section, by determining whether and how right branches or rightward movement can be integrated to improve analyses that must exclude them because of antisymmetry. For further insights in this respect, see Haider (2010: 231-236).

3.2 Previous analyses

For the empirical part related to Greenberg's Universal 20, Cinque (2005) presents a convenience sample (by following the terminology proposed by Körtvélyessy 2017: 127) which has been highly influencial for the subsequent publications on the same matter. By gathering all the available data that he was able to find with respect to U20, Cinque (2005: 319-320) lists the twenty-four logically possible alternate orderings by clustering them into six groups of four elements each. Within each group, the linear order of the modifiers is kept constant, while the position of the noun varies, as follows:

(102)	a.	Dem	Num	A	**N**	[very many languages]
	b.	Dem	Num	**N**	A	[many languages]
	c.	Dem	**N**	Num	A	[very few languages]
	d.	**N**	Dem	Num	A	[few languages]
	e. *	Num	Dem	A	**N**	[not attested]
	f. *	Num	Dem	**N**	A	[not attested]
	g. *	Num	**N**	Dem	A	[not attested]
	h. *	**N**	Num	Dem	A	[not attested]

i.	*	A	Dem	Num	**N**	[not attested]
j.	*	A	Dem	**N**	Num	[not attested]
k.		A	**N**	Dem	Num	[very few languages]
l.		**N**	A	Dem	Num	[few languages]
m.	*	Dem	A	Num	**N**	[not attested]
n.		Dem	A	**N**	Num	[very few languages]
o.		Dem	**N**	A	Num	[many languages]
p.		**N**	Dem	A	Num	[very few languages]
q.	*	Num	A	Dem	**N**	[not attested]
r.		Num	A	**N**	Dem	[very few languages]
s.		Num	**N**	A	Dem	[few languages]
t.		**N**	Num	A	Dem	[few languages]
u.	*	A	Num	Dem	**N**	[not attested]
v.	*	A	Num	**N**	Dem	[not attested]
w.		A	**N**	Num	Dem	[very few languages]
x.		**N**	A	Num	Dem	[very many languages]

As (102) shows, fourteen orders are attested with different frequencies, while ten are not found as the only unmarked order of a language. From this observation, Cinque develops two related models (Cinque 2005, 2023) aimed at constraining the possible derivations in such a way that the unattested orders cannot be derived (modulo pragmatically marked contexts where single modifiers can be subject to topicalization or focalization, see Cinque 2005: 321, 327). Cinque's major works on Greenberg's Universal 20 try to capture these facts by complying with all the main points discussed in section 2.3.1 with respect to the anti-symmetric approach (Kayne 1994). Therefore, Cinque adopts what could be referred to an Array-Based approach (in the terms of Krivochen 2024) that relies on endocentricity, binary branching and widespread syntactic movements that take place within a single base-generated structure. The latter complies in turn with the Specifier-Head-Complement-Hypothesis (see sections 2.2.3 and 2.3.1).

Cinque (2005) has been highly influential to formal analyses that have been published afterwards. In fact, most of them take for granted the fourteen attested orders listed in (102) and try to explain them by means of different theoretical assumptions, where the distinction

between base-generated and derived orders is a relevant matter of debate. After discussing Cinque (2005) in section 3.2.1, from section 3.2.3 to section 3.2.5 analyses of this kind will be surveyed, i.e. Cinque (2023), Roberts (2017b), Abels & Neeleman (2009, 2012), Georgi & Müller (2010), and Medeiros (2018).[60]

An alternate view which will be discussed in sections 3.2.5 and 3.2.6 is represented by those works that claim that the number of attested permutations of U20 reported by Cinque (2005) is not correct, being it either smaller (Rijkhoff 2002, 2015) or broader (Nchare 2012, Ouwayda & Shlonsky 2017, Dryer 2018a, Bruening 2018a, Steedman 2020). Given that the works reporting fewer orders are somehow encompassed in the analyses provided by Abels & Neeleman (2009, 2012) and Georgi & Müller (2010), it is the group of works relying on a broader range of linear orders that constitutes a more challenging case for any analysis stemming from Cinque (2005). In fact, any account of the attested orders of U20 makes the complementary claim of explaining why the remaining are still unattested. Any counterexample that shows that the proposed model undergenerates (namely, that the model fails to derive an attested unmarked permutation of Greenberg's Universal 20) represents an issue (as Bruening 2018a argues) that has to be dealt with, either by demonstrating its spurious nature or by making an argument that explains its exceptional behavior (Bickel 2015: 903).

An underexplored third alternative consists in devising a model that does not ban the exceptions altogether, but rather claims that they are more costly in some relevant respect (as argued by Hucklebridge 2023 with respect to the FOFC patterns in the clausal domain) and thus so rare that they are unattested even in convenience samples. Under this view, which will be further pursued in section 5.5.1, the collected data would fail to find positive evidence for the the ultimately "Zipfian" distribution of U20 (see Merlo & Ouwayda 2018: Sec. 4.2, 7.1 and Croft 2003: Sec. 3.1), characterized by a long tail of highly rare alternatives for which one could claim that "absence of evidence is not evidence of absence" (S. Müller 2023a: 161).[61]

[60] The range of works that provide an account for the data in (102) is broader, but they will be set aside either because they suffer from problems that the main text discusses for other accounts, or because their premises are too distant from the theoretical perspective pursued in this work. An non-exhaustive list can encompass Hawkins (1983: 119-120), C.-W. Chung (2003), Artiagoitia (2006: Sec. 4, 2008: Sec. 1), Cysouw (2010), Steddy & Samek-Lodovici (2011), Manova (2021), Chesi, Nieto & Zamparelli (2022), Bernardi (2022) and Ferrer-i-Cancho (2024).

[61] By following the last alternative proposed in the main text, the distinction between frequency and grammaticality is thinned out to the point where it disappears by becoming a continuum. See Steddy & Samek-Lodovici (2011: 460 fn. 9) and Abeillé (2024: 7-10) for related discussion.

3.2.1 Cinque (2005)

The analysis of U20 proposed by Cinque (2005) is an application of antisymmetry (Kayne 1994) which inherits strengths and weaknesses from it. In fact, Cinque tries to describe the crosslinguistic quantitative and qualitative variation of the attested arrangements by assuming that any attested permutation can be derived from a [Dem [Num [A [N]]]] base exclusively via phrasal movement. This can be done by assuming that each element of U20 is a maximal projection and, for the sake of the Chain Uniformity Condition (Chomsky 1995: 253), a free specifier position is targeted whenever a movement is supposed to take place. Hence, Cinque (2005) assumes that each modifer of the noun is first merged as a specifier of a functional projection, which is in turn dominated by a dedicated Agreement Projection (AgrP), whose specifier is the landing site for phrasal movements of phrases c-commanded by the related Agr head.

In order to account for the data on U20, Cinque (2005: 317) assumes that the Noun can raise (leftwards) by itself within the nominal domain or pied-pipe the modifiers whose functional projections properly contain it at base-generation.[62] On the contrary, Dem(onstrative), Num(eral), and A(djective) cannot raise autonomously, nor can they be subject to remnant movement once the Noun has raised past them (Cinque 2005: 326-327). If NP first raises to a position where it c-commands the modifiers it subsequently pied-pipes, it gives rise to the "whose picture[s]" type of pied-piping, i.e. [NP [XP]] (Cinque 2005: 321 fn. 21). If instead the noun pied-pipes some modifier(s) without first raising to a c-commanding position, this gives rise to the "picture[s] of who[m]" movement [XP [NP]] (Cinque 2005: 321 fn. 22).[63]

The structure in (103) is now illustrated for three purposes. First, it serves as an adaptation of the base structure shown in Cinque (2005: 318), which can be inferred by ignoring the elements moved to Spec,Agr$_y$P and Spec,Agr$_x$P).[64] Second, the arrow labeled with [1] shows how the Dem-Num-N-A order is derived under this account by raising NP to the specifier of

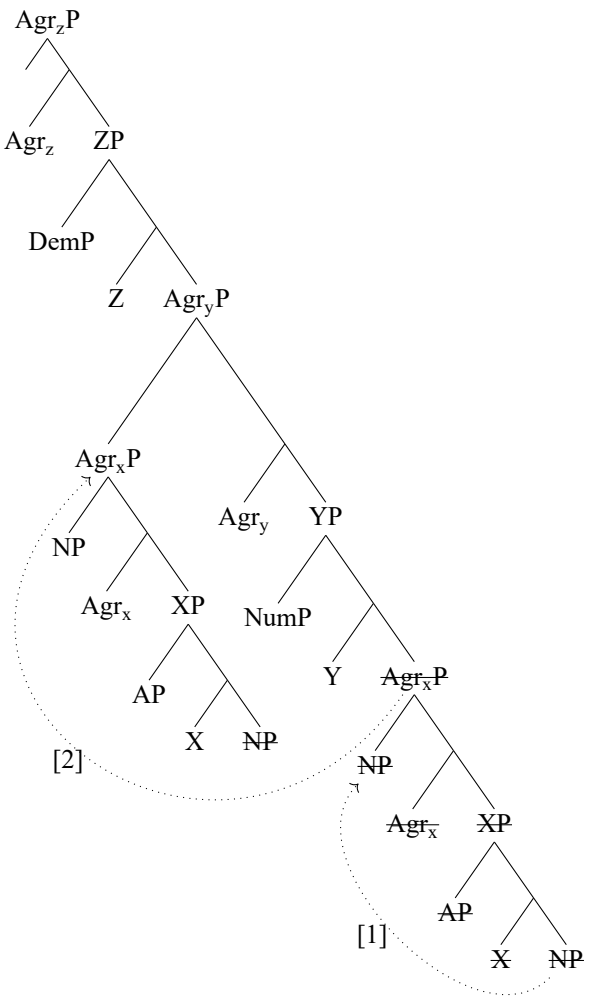
[62] Ross (1967) first introduced this concept for discussing the *wh*-movement of phrases containing a *wh*-operator, just like the pied-piper of Hamel(i)n used to enthrall out of town the rats—and then the children. See also Horvath (2006), Cinque (2009: 169 ff.) and Cinque (2017b: 530 fn. 9) for further discussion.

[63] Cinque (2005) talks about "whose picture" and "picture of who" types of movement, while in later works (e.g. Cinque 2023) the former is addressed as "whose pictures" and the latter as "pictures of whom". It is also worth observing that Cinque never discusses how case is assigned in the *wh*-constructions that give the name to such movements, nor does he state whether this aspect should be paralleled in the derived orders of U20. See Stowell (1989: Sec. 2.2) for an early discussion of these issues.

[64] Cinque (2005: 318) actually labels the needed functional heads W, X, Y by placing W and Y as the farthest and the closest from NP, respectively. Here, the order of such heads is reversed so as to highlight the bottom-up nature of the derivation (see Fukui & Takano 1998, Uriagereka 1999, a.o.).

the closest AgrP. And finally, the arrow labeled with [2] illustrates the *whose pictures* pied-piping of [NP [AP]] to the specifier of Agr_yP (which subsumes the NP-raising showed by [1]).

(103)



The three types of movement mentioned above, where movement of NP is always involved, not only fulfill the goal of generating the attested orders, they also fail to derive the unattested ones.[65] Although wrong with respect to the range of possible linearizations, the original formulation of U20 proposed by Greenberg (1966–1963: 87, 111) can therefore be seen as correctly capturing three distinct cases that have to be restated in derivational terms (see also Crisma 1991 for an ealy attempt). Moreover, the quantitative variation among the orders that the movements assumed by Cinque (2005) can generate seems to correlate with the types of movement involved in their derivation. Put differently, if one order is less attested

[65] See Svenonius (2024: 13 fn. 4) for a list of works that predate the antisymmetric account developed by Cinque (2005) or influence the subsequent refinements proposed in Cinque (2023), to be discussed in section 3.2.2.

than another crosslinguistically, it is because the steps required to derive the former are more marked than the latter. By following these assumptions, Cinque (2005: 321) shows that the *whose pictures* type of movement is the least marked, while the *pictures of whom* is the most marked, because they are respectively the most and the least likely crosslinguistically based on the frequencies in (102). Easy empirical evidence for the unmarkedness of the *whose pictures* movement, which is also maintained in Cinque (2023), it that the N-A-Num-Dem order is both the most widely attested and the only one which consistently applies this type of movement cyclically across all the projections within the nominal domain (Cinque 2005: 324).

Cinque (2005: 325) admits however that the correlations between the frequency of a certain order and the types of movement that derive it have no clear independent motivation, aside from the degree of markedness of different *wh*-phrases in English. This parallelism is also reported in Cinque (2023: 5), where the examples in (104) are used to sketch an analogy between *wh*-movements and movements that derived the attested orders of U20. The *whose pictures* and *pictures of whom* pied-pipings are illustrated by (104b) and (104c), respectively, while the autonomous raising of NP can be put on a par with the *wh*-element in (104a).

(104) a. Aylan, [who] you have certainly seen [pictures of [__],..
b. Aylan, [[whose] pictures] you have certainly seen [__],..
c. Aylan, [pictures of [whom]] you have certainly seen [__],..

Although the problems linked to the parallel between the nominal and the clausal domains will be postponed to the discussion of Cinque (2023) in section 3.2.2, it should be observed that linking the frequency of a certain order in (102) to the nature of the derivational steps required to derive it may cause an "overfitting" of the data (Merlo & Ouwayda 2018: 27). The base-generated structure in (103) represents a related issue with respect to the antisymmetric assumptions made by Cinque (2005). If the Specifier-Head-Complement-Hypothesis leads Kayne (1994) to assume that the universal base is a structure that leads to the SVO order, then the Dem-Num-A-N base should positively correlate with it.

Although Zamparelli (2000: 21-22) observes that this claim resonates with Crisma (1991), it is confirmed neither by previous typological observations (see e.g. Primus 2001: 856 and the discussion in section 2.2.1) nor by syntactic works like Szabolcsi (1994: 180). In agreement with the latter works, the data in appendix A.3 will disconfirm a clear correlation be-

tween SVO and Dem-Num-A-N (see also section 4.3).

Another issue worth discussing is the derivation of N-Dem-A-Num, which is peculiar for being the only one that Cinque (2005) can derive in two different ways. Although this characteristic may be interpreted as a wrong predictor of a higher crosslinguistic likelihood, Cinque (2005: 323 fn. 27) argues that this order might be spurious because it is rarely attested and often found as an alternative to other orders. This would be captured by the fact both derivations require a highly marked combination of movements. The first option starts with the *whose pictures* pied-piping to Spec,Agr$_y$P above NumP depicted in (103), while the other option begins with the *pictures of whom* pied-piping of [AP [NP]] to the same target position (Cinque 2005: 323). The problematic step that the two derivations share is the raising of NP above DemP, i.e. out of Agr$_x$P to Spec,Agr$_z$P. N-Dem-A-Num is thus the only order where pied-piped material must be stranded in an intermediate position, so that its exclusion from the unmarked order of U20 would make the claim for intermediate stranding unnecessary.

However, Abels & Neeleman (2012: 61-62) argue that Pitjantjatjara and Nkore-Kiga (which are a Pama-Nyungan language spoken in Australia and a Bandoid language spoken in Uganda, respectively) represent genuine attestations of N-Dem-A-Num, and Cinque does not contest the validity of this order of U20 in later works, as Cinque (2023) and the data collected by Cinque (2021) in appendix A.13 show. Still, Abels & Neeleman (2012: Sec. 5.2.3) claim that the stranding of pied-piped material in an intermediate position is theoretically problematic. In fact, they argue that it finds no parallel in other domains, as the patterns of preposition stranding in English, already discussed in Postal (1972), indicate. The examples in (105) and (106), adapted from Abels & Neeleman (2012: 60), show that pied-piped material can be stranded only in their *in situ* position in cases of *wh*-movement and of movement under relative clause formation, respectively.

(105) a. [$_{PP}$ With which friend] did you say t_{PP} that Mary went to the movies t_{PP}?
b. [$_{DP}$ Which friend] did you say t_{DP} that Mary went to the movies with t_{DP}?
c. *[$_{DP}$ Which friend] did you say [$_{PP}$ with t_{DP}] that Mary went to the movies t_{PP}?

(106) a. The famous rock [$_{DP}$ pictures of which] I think t_{DP} that Bill must have seen t_{DP} before
b. The famous rock [$_{DP}$ which] I think t_{DP} that Bill must have seen pictures of t_{DP} before
c. *The famous rock [$_{DP}$ which] I think [$_{DP}$ pictures of t_{DP}] that Bill must have seen

t_{DP} before

What can be observed in English is not necessarily representative of a ban holding crosslinguistically. In fact, Davis (2020: Ch. 2) summarizes the cases that have been reported in favor of the possibility of intermediate stranding as a consequence of pied-piping. Among the example discussed there, (107) adapts a case from West Ulster English (McCloskey 2000: 61), while (108) adapts data from Afrikaans first provided by Plessis (1977). What such cases have in common is that the material that is assumed to be stranded linearly follows the "pied-piper" (Horvath 2006: Sec. 3.2, Abels & Neeleman 2012: 60) in some derivational step preceding the stranding itself, whereas the material that fails to be stranded in (105) and (106) linearly precedes its pied-piper (Davis 2020: 30).[66]

(107) What <all> did he say <all> (that) he wanted <all>? [West Ulster English]

(108) *Waar<voor> dink julle <voor> werk ons <voor>?* [Afrikaans]
where<for> think 2SG <for> work 1PL <for>?
'For what do you think that we work?'

One may thus use the examples in (107) and (108) as independent evidence for the fact that the account of N-Dem-A-Num in Cinque (2005) should opt for the stranding of AP as a result of the *whose pictures* pied-piping. However, traces and brackets have been omitted from (107) and (108) because Rackowski & Richards (2005), Dikken (2009), Besten (2010), and Bošković (2016b) argue that such patterns are not the result of intermediate stranding, but rather of different mechanisms acting on input sentences that do not necessarily coincide with what intermediate stranding accounts would assume.[67] As section 3.2.4 will illustrate, Abels & Neeleman (2009, 2012) propose an alternate account of U20 which is deemed to be superior to Cinque (2005) because (among other things) it accounts for N-Dem-A-Num by means of a derivation where stranding of pied-piped material in intermediate positions is dispensed with.

66 The discussion in the main text overlooks the flip side of the coin, i.e. the phrase-internal position of the pied-piper, discussed a.o. by Horvath (2006: Sec. 3.1) and Richards (2014: Sec. 4.2). The facts emerging from this complementary inquiry do not contradict however what is reported following Davis (2020: 30), as soon as one acknowledges with Horvath (2006: 582) that the raising to a Spec position might well be what makes the pied-piper linearly precede the rest of the phrase.

67 See Beermann & Hellan (2011) for related arguments with respect to preposition stranding in German.

3.2.2 Cinque (2023)

Bernardi (2022) criticizes the quantitative predictions of the Dem-Num-A-N order made by Cinque (2005). Since it is base-generated, it should *ipso facto* either outnumber any other permutation which would necessarily require some additional movement, or be banned by a generalized version of the *in-situ* generalization proposed by Alexiadou & Anagnostopoulou (2001, 2007), as in Roberts (2017b: 182).

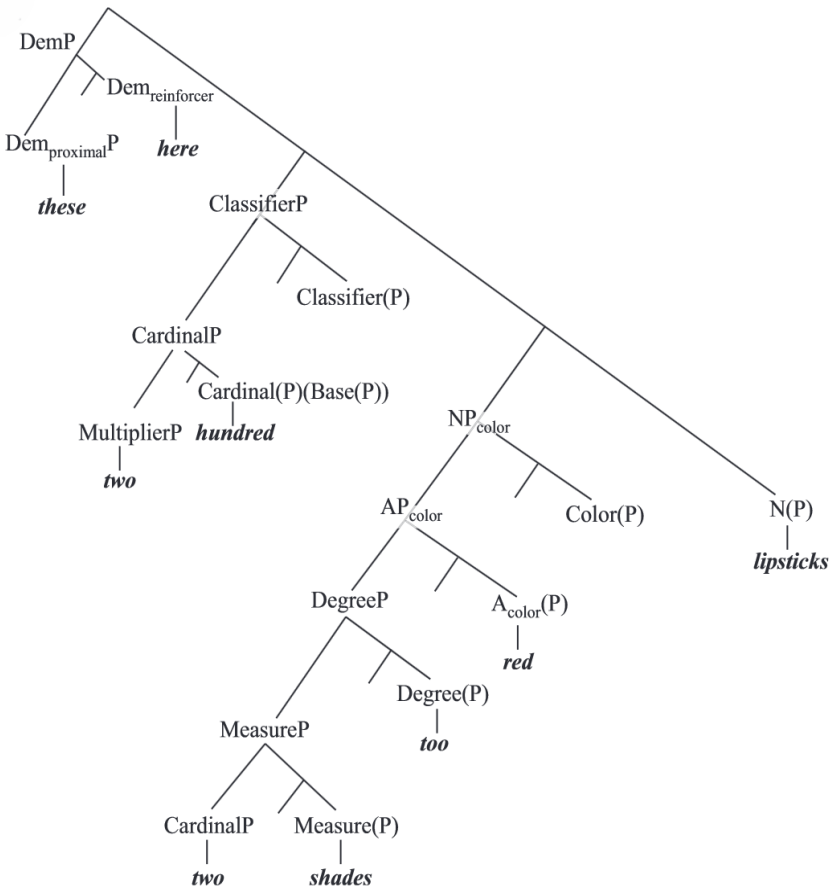
This argument is however contradicted by Merlo & Ouwayda (2018), where a computational modeling of Cinque (2005) is proposed. The total absence of movement emerges as the least costly operation among the ones assumed by Cinque only if one excludes from the picture the total cyclic *whose pictures* movement that gives rise to N-A-Num-Dem. Merlo & Ouwayda (2018: 13) tentatively attribute the desirability of a complete *whose pictures* movement (as opposed to none) by appealing to agreement as a key factor feeding movement, as in Shlonsky (2012). However, the quotation from Norris (2017b: 3) already discussed in page 102 has shown that such an explanation is hardly tenable.

Cinque (2023: 11) takes a rather different stance by claiming that Dem-Num-A-N is actually derived by means of the *pictures of whom* pied-piping applying across all projections, rather than by absence of movement. Cinque (2023: 23) achieves this change by modifying the base-generated structure in the nominal domain, which is no longer "a unique downward/rightward spine" as it was in Cinque (2005), but rather "a more highly articulated constituent". This is illustrated in (109) on the next page with a structure which is taken directly from Cinque (2023: 23).

One first problematic consequence of the representation in (109) is that it labels the rightmost node as N(P) "to underline the undetermined head (X°)/phrasal (XP) status of the core element of the projection (something that projects but also moves to a specifier position)" (Cinque 2023: 11). Although this assumption complies with a Bare Phrase Structure approach (Chomsky 1993), Cinque (2023) does not discuss how the subsequently assumed N(P)-movements must be qualified with respect to restrictions related to structure preservation and chain uniformity (see Emonds 2012: 335-336, Harizanov 2019: Sec. 2.1 and Krivochen 2023b: Sec. 1.2 for discussion).

DemP, ClassifierP and NP_{color} in (109) are subprojections of the nominal. Each of those represents a local domain which contains one of the modifiers of U20. In fact, Cinque (2023: Ch. 2-3) lists a series of empirical observations that can be captured by assuming that each

(109)



subprojection can be reordered internally by means of the three types of movement already posited in Cinque (2005), but whose linear and structural compactness cannot be affected by elements belonging to other subprojections. These premises imply that Cinque (2023) favors a classifier-for-numeral analysis (see 3.1.2) where "copious movevements are needed" (Her 2017: 292), while linear discontinuities between determiners and demonstratives (discussed in section 3.1.1) are problematic for the analysis. In fact, neither Cinque (2005) nor Cinque (2023) do discuss data suggesting a possible low origin of demonstratives (even though it is discussed in Cinque 1996: 454 fn. 16).

Although the representation in (109) is explicitly designed to hint "that the entire nominal projection is headed just by N(P)" (Cinque 2023: 40), it is a simplification of a far more complex structure that complies with the FHH (see section 2.2.4) and, in particular, with the view that the maximal projection of nominals is KP, as argued in Barrie et al. (2021). In fact, KP is not shown in (109), where the root of the structure should actually be labeled DeicticP, i.e. a "modifier projection" (Cinque 2023: 42) whose specifier is filled by DemP. In a more

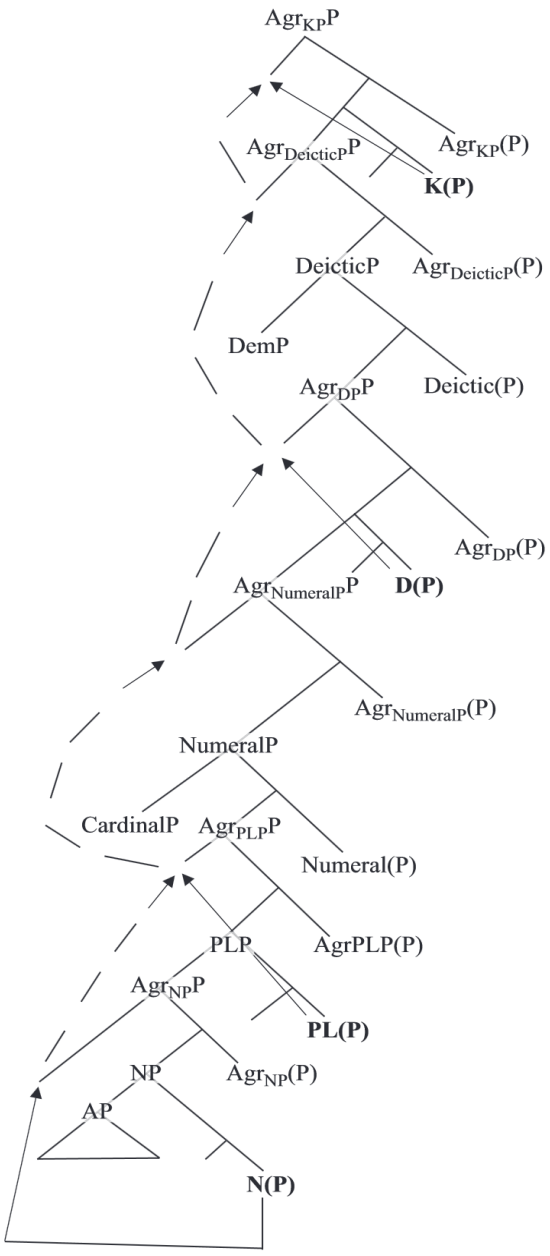
articulated structure, KP is a "substantive functional projection" (Cinque 2023: 41) whose specifier contains DeicticP. Similarly, the unlabeled node immediately dominating NP_{color} and N(P) is the maximal projection NP, which takes the subprojection containing the AP as its specifier, while the intermediate node that in (109) is on the path from NP to DeicticP must be labeled NumeralP, which takes ClassifierP as its specifier.

The heads of DeicticP and NumeralP are never overtly realized, and each of such projections is "associated with a substantive functional projection" (Cinque 2023: 42) which is embedded in the complement of the former. In particular, DeicticP is related to DP whereas NumeralP is associated with PL, i.e. a head that instantiates number marking (Cinque 2023: 11). To conclude the whole depiction of this more articulated structure, it must be observed that the complements of N, PL, D, K are always empty, while the Agr(eement) P(rojection)s already assumed in Cinque (2005) are pervasive in the structure, given that each projection is assigned a functional AgrP layer on its own top (Cinque 2023: 6). This can be obtained by having AgrPs merged to the specifiers of the just mentioned substantive projections, as well as to the complements of DeicticP and NumeralP.

The structure which has just been described can be visualized by means of (110) on the following page, which is taken directly from Cinque (2023: 47), without considering for the moment the arrows (which stand for movements to be discussed shortly). The structure in (110) features a recursive application of the X-bar schema that complies with the Single Complement Hypothesis (Larson 2014: 174-175), as in section 2.2.3 and in Cinque (2005). Cinque (2023: 5) further claims "that complements, like modifiers, are merged in Specifier positions". Although it resonates (due to totally different theoretical reasons) with what is extensively claimed in Adger (2013), this stance weakens a lot the role of complements in the whole proposal, since they are retained only to host, or be hosted by, AgrPs, and be null whenever their sister heads are pronounced. AgrPs, in turn, are necessary to make the three types of movement already posited in Cinque (2005) possible. All the fourteen attested orders can be derived by means of movements that overlap with what Cinque (2005) establishes.[68] This is implies that Dem-Num-A-N is the only permutation of U20 which is accounted in a totally different way in comparison to Cinque's previous work. In fact, it is no longer mapped from the base-generated structure, but rather derived through the consistent application of movements of the *pictures of whom* type.

[68] The issues presented in section 3.2.1 with respect to the derivation of N-Dem-A-Num are simply no longer discussed in Cinque (2023), although they are by no means solved by the revised assumptions.

(110)



Substantive functional projections are added in Cinque (2023: 41-42) to account for those languages where the Pl, D and K heads are realized in discontinuity with the noun. Given that their heads and specifiers are realized overtly, such projections are different from AgrPs. Functional projections in Cinque's account can thus be split in two subgroups that are reminiscent of the distinction between "lexical functional projections" and "functional extensions of lexical projections" made by Haider (2013: 71) and discussed in section 2.4.2. The solid arrows in (110) indicate the possible movement of each substantive functional head to the specifier of "its own AgrP" (Cinque 2023: 46). Whereas substantive functional heads can-

not raise further, N can iterate the raising to the closest Spec,AgrP, as the dashed arrows in (110) indicate. The special status of N thus allows it to raise up to Spec,Agr_{KP}P without pied-piping, so that Cinque (2023: 46-50) manages to account for the N-Dem-Num-A order (as well as Dem-Num-N-A and Dem-N-Num-A).

Although it is an ambitious expansion of Cinque (2005), Cinque (2023) brings new problems to the table. The first is related to the claim that the structures in (109) and (110) are "still in compliance with Antisymmetry" (Cinque 2023: 5) even though no syntactic structure stems "below/to the right of a lexical category" (Cinque 2023: 90). In fact, one of the most impactful elements formulated by Kayne (1994: Sec. 4.1) is that "in any phrase marker [...] specifier and complement are on opposite sides of the head", as it was discussed in section 2.3.1.

Therefore, a syntactic structure based on X-bar schema of the Spec-Comp-H type (as in structure 24c on page 47) is not compliant with an antisymmetric approach in the original sense. One can still define it as "antisymmetric" only if a single factor actually matters, namely that either H or Comp be consistently "null and thus uninfluential for linearization under the LCA" (Cinque 2023: 42). In fact, Chomsky (2014–1995: 310) proposes this line of reasoning to "weaken the LCA", while Koopman (2000a: 330 ff.) draws conclusions that overlap with the claims in Cinque (2023) by giving up the segment/category distinction that plays a pivotal role in Kayne (1994), as discussed on page 49. A highly articulated structure like (110) also proves that a common criticism made by scholars who reject antisymmetry (see section 2.4.1) is ultimately correct, namely that with enough silent functional projections it is possible to respect the LCA and derive linear orders that one would intuitively expect to be ruled out by following Kayne (1994).

Another related issue with the structures proposed by Cinque (2023) is that they feature recursive applications of the X-bar schema where the lexical head is in a final position, as per the alternate approaches to Kayne (1994) discussed in section 2.4.2. This structural overlapping with Barbiers (2000) and Haider (2000, 2013) is explicitly acknowledged in Cinque (2017b: 529 fn. 6) and Cinque (2023: 128 n. 4), but in Cinque's proposal head-final orders require some derivation by movement of the *pictures of whom* type which the former approaches do not assume to occur. As a consequence, movement becomes the *sine qua non* of any derivation, as in the approaches to U20 adopted by Steddy & Samek-Lodovici (2011) and Medeiros (2018).

The consistent application of movements of the *pictures of whom* type in the derivation of Dem-Num-A-N also predicts its lower crosslinguistic frequency with respect to N-A-Num-Dem (Cinque 2023: 51), which is still derived through a total pied-piping of the less marked *whose pictures* type, as was assumed in Cinque (2005). Nevertheless, the exclusion from this line of argument of the raising of the Noun without pied-piping is not motivated anywhere. This represents a problem for the empirical side of the account in Cinque (2023), given that in Cinque (2005: 321) it was claimed to have a degree of markedness which sets it between the two types of pied-piping he proposes.

It is true that the modeling proposed by Merlo & Ouwayda (2018: 12) sets the Noun raising without pied-piping as the most costly syntactic operation among the possible derivations of U20 (*contra* Cinque 2005: 321 himself), but Cinque (2023) does not give hints at a possible reconsideration of the claims he made in this respect (though see Cinque 2017b: 530 fn. 9). This leaves open the question of why there is no linearization derived only by Noun raising without pied-piping whose crosslinguistic frequency lies between N-A-Num-Dem and Dem-Num-A-N.[69]

Cinque (2023) also goes beyond Cinque (2005) in sketching how to derive the crosslinguistic variability in the clausal word order. The new structures he posits for the two domains predict a parallel between Dem-Num-A-N and SOV on the one hand and N-A-Num-Dem and VOS on the other, which depend respectively on whether a language consistently applies pied-piping of the *pictures of whom* or of the *whose pictures* type. Thus, Cinque (2023) manages to account for the fact observed by Szabolcsi (1994: 180) that OV languages tendentially prefer the modifiers of U20 to be consistently prenominal, as opposed to what can be inferred from Cinque (2005).

The consistent *pictures of whom* or *whose pictures* pied-pipings thus give rise to two opposing derivations where the functional heads Pl, D and K consistently follow or precede the noun, respectively (Cinque 2023: 40), and they more generally resonate with the *Consistent Head Serialization* discussed an problematized in 2.2.1. It is not surprising, then, that Cinque

[69] Neither Cinque (2005: 325) nor Cinque (2023: 98) ever give an example of the *wh*-phrase raising alone in English in comparison to other intra-linguistic examples explicitly mentioned to assess the degree of markedness of the two types of pied-piping he proposes. However, Stepanov (2007: 80) and Bernardi (2023: n. vii) give English examples that support the claim that N-raising alone should set between the two types of pied-piping, so that the problem pointed out in the main text seems to be a rather tough one. Still, the latter objection is made by assuming that the whole line of argument given by Cinque (2005: 325) and Cinque (2023: 98) is actually sound. The observations reported in Abeillé (2024: 7 ff.) and references therein show however that the solidity of Cinque's premise needs further careful consideration of its own in the future.

(2023: 68) ends up observing that these two consistent "head-final" and "head-initial" alternatives are only idealizations that are challenged by the data: "the more rigid head-initial VOS languages and the more rigid head-final SOV languages […] display some inconsistencies, and, more strikingly, they appear to be a minority of the languages of the world".[70]

3.2.3 Roberts (2017b)

As anticipated in section 2.3.2, the Num-N-Dem linearization is a case worth discussing in a FOFC perspective, because it is widely attested although not legitimate as soon as the Specifier-Head-Complement Hypothesis (SHCH) and the *Dem* > *Num* > *N* hierarchy proposed by Cinque (2005, 2023) are assumed together with a roll-up account of head-final orders (which was discussed in section 2.3.3 and problematized in section 2.4.1). In fact, no matter if the lexically overt demonstrative is a head or a maximal projection, if Num and N raise together in a *pictures of whom* kind of pied-piping, FOFC will be violated because of the missing internal roll-up movement that would make (a projection containing) N being linearized to the left of Num. Given that the analysis of U20 proposed by Roberts (2017b) is couched on both antisymmetry and FOFC (see section 2.3), it is designed to capture the crosslinguistic variation by demonstrating that widely attested FOFC violations are only apparent, so that FOFC is still completely (or at least widely) obeyed. Therefore, the line of argument followed by Roberts (2017b) represents a good example of the "Pāṇinian approach" defined by Bickel (2015: Sec. 37.1).

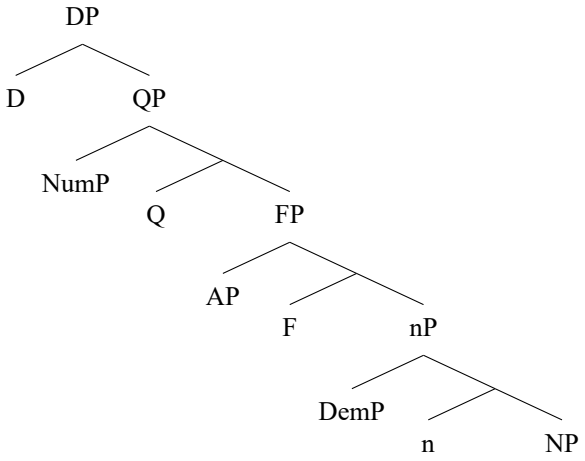
The account of U20 provided by Roberts (2017b) differs from Cinque (2005, 2023) in claiming that demonstratives appearing on the right edge of the nominal domain are in a low position, as argued by Brugè (1996) and Cardinaletti & Giusti (2006) a.o. (see again section 3.1.1). Thus, their appearance on the left edge is deemed to be a result of raising to Spec,DP. To defend such a view, Roberts (2017b) argues that, in its base-generated position, DemP constitutes the external argument of *n*.[71] The structure in (111) illustrates the nominal extended projection that the account in Roberts (2017b) would generate if nothing were moved. The *n* head determines "nonlexical properties of the lexical root (case, agreement, perhaps category, etc.)" (Roberts 2017b: 162) and in exceptional cases it can be lexical-

[70] See Cinque (2017a: Sec. 4) for further discussion related to this issue, together with a sketch of a possible micro-parametric approach to the problem.

[71] Roberts (2017b: Sec. 8.4-8.6) bases this claim on Williams (1980), while Franco, Savoia & Manzini (2015: Sec. 4.1) and Giusti (2018: Sec. 3) make overlapping assumptions by appealing to Higginbotham (1985).

ized by articles (Roberts 2017b: 164). Both Q and F can instead lexicalize a classifier or a plural marker (Roberts 2017b: 165, 181). Although the argument for each of the two heads is grounded on the Greenberg-Sanches-Slobin generalization discussed in section 3.1.2, the structure in (111) actually predicts the contrary of what such a generalization states, i.e. that sortal classifiers and number marker can co-occur by filling distinct heads. See Tatsumi (2020) for an attempt to fix this issue.

(111)



Contrary to Cinque (2005, 2023), Roberts (2017b) does not assume agreement projections in the base-generated structure. On the contrary, a single maximal projection can have multiple specifiers as a consequence of leftward phrasal movement to a c-commanding position.[72] In particular, Roberts (2017b: 153, 181-182) assumes that Comp-to-Spec movement targets an outer specifier (and so, arguably, do A′-moved elements), whereas the inner specifier is occupied either by an externally merged category or by elements that have undergone Spec-to-Spec A-movement. Exactly as it has been observed for Cinque (2023), a claim of this kind represents a weakening of antisymmetry (Kayne 1994) that is plausibly done by tacitly following the suggestion made by Chomsky (2014–1995: 313) in this respect.

As the implicit rejection of the LHH in (111) suggests, Roberts (2017b) argues in favor of a parallel between the nominal and the clausal domains (see section 2.2.4). The aforementioned relation between *n* and DemP mirrors the one between the *v* and its external argument, with the latter being able to appear in Spec,TP after raising. The parallel is further strengthened by assuming that (111) is bi-phasal (see Syed & Simpson 2017, Simpson & Syed 2020 for a similar claim with a different implementation to be discussed in section 4.5.2). This mirrors

[72] See Ura (2001: Sec. 4) and references therein for some relevant comparison of the approaches adopting multiple specifiers as opposed to those relying on AgrPs.

the assertion made in Chomsky (2000, 2008) and much subsequent work, where CPs and (transitive) *v*Ps are phases (but see Keine & Zeijlstra 2023: Sec. 1 for a summary of the alternate views). In particular, Roberts (2017b) argues that their respective equivalents in (111) are D and *n*. However, the parallel between the two extended projections is not exact, as it has been shown for other proposals following the FHH to nominals in sections 2.2.3 and 2.2.4. In fact, Roberts (2017b: 363 n. 12,15) follows Chomsky (2008) in assuming that some properties of the C phase are inherited by the c-selected T, among which the EPP feature that triggers the movement of the external argument of the verb. No similar mechanism is posited between D and Q, even though the latter bears "the basic property of quantifying over the lower phasal category" just like T does (Roberts 2017b: 161).

With this caveat in mind, Roberts (2017b) formulates the following hypothesis: if it is possible to have DemP raised from its *in situ* position to Spec,DP, this should positively correlate with the movement of the external argument of *v* to Spec,TP. This can be captured by a single binary parameter, so that the correlation should also hold for the opposite situation, where neither DemP nor the external argment of *v* are attracted by the higher phase. Roberts (2017b) thus tries to falsify the hypothesized correlation by relying on the relevant data provided by the World Atlas of Language Structures (WALS, Dryer & Haspelmath 2013). In his analysis, Roberts (2017b: 155-158) gathers more than 70 languages and correctly observes that, among them, the VO languages are equally distributed between the VS/SV divide. However, the best he can infer from these data is that "there *seems* to be a link between verb-initial clausal order and LD order. In many cases, LD languages have VSO or VOS order" (Roberts 2017b: 165, italics added), where "LD" stands for "Low-Demonstrative". His hesitation is motivated by the fact that "a rigid VS-LD correlation [...] has counterexamples in both directions." In fact, "there certainly appear to be both SV-LD languages (e.g., Modern Hebrew, Wolof, Indonesian, [...]) and VS-HD languages" (Roberts 2017b: 167).

The part of Roberts's analysis summarized above is perfectly legitimate and sound, since it avoids the methodological circularity criticized by Greenberg (1990: 703), who discusses the subtle difference between correctly using typological data to falsify a hypothesis, as opposed to inappropriately using them for amending one. However, Roberts (2017b: 155-158) strikingly overlooks that, among the more than 70 languages he collects, only two are not VO, namely Tuvaluan and Basque, which Dryer (2013d) classifies as OVS and SOV, respectively

(but see Besnier 2002 for an argument in favor a VSO classification of the former).[73] On a broader scale, this is remarkable, because (as it will be discussed further in section 4.2.3) V-initial languages are attested in less than 15% of the languages of the world, while SOV languages constitute more than 40% of the total.

It is worth observing that Roberts (2017b) neglects the strikingly negative correlation between V-final languages and the two orders of U20 he analyzes (i.e. Num-N-A-Dem and Num-A-N-Dem) in favor of a weaker linking between V-S order and "Low Demonstratives". This fact is an unwelcome consequence of some deductive reasoning based on a not so effective "theory of grammar" (Greenberg 1990). By virtue of this flaw, the sample from Roberts (2017b: 155-158) can be taken as the starting point for two entangled lines of typological inquiry to be further discussed in section 4.1, which will in turn eventually lead to the collection of the data presented in appendix A.

Although FOFC has the merit of capturing the fact that certain structural configurations are remarkably rare crosslinguistically (see section 2.3.2), the account proposed by Roberts (2017b) fails to be as compelling with respect to U20. In fact, the ten orders that are not attested in Cinque (2005) (see page 123) are not ruled out because of some FOFC-violation, but rather because of some independent bans. The first is the prohibition against the raising of AP out of FP, by means of which A-Num-Dem-N is (among others) correctly predicted not to be attested, because it would require the raising of AP to either Spec,QP or Spec,DP in (111). The second ban is the impossibility of DemP to raise to a position different from Spec,DP, so that the non-attestation of Num-Dem-A-N (among others) is accounted by Roberts (2017b: 181) due to the illicit DemP movement to the outer Spec,FP.[74] However, it would be possible to derive this order by A-moving DemP to the inner specifier of DP, followed by a NumP movement to the outer specifier of DP. Therefore, a ban on NumP movement out of QP (similar to what Cinque 2005, 2023 needs to posit) is also required in Roberts (2017b), even though it is not explicitly stated. Moreover, Roberts (2017b: 182) claims that the order that can be mapped from the base-generated structure in (111), i.e. Num-A-Dem-N, is not attested

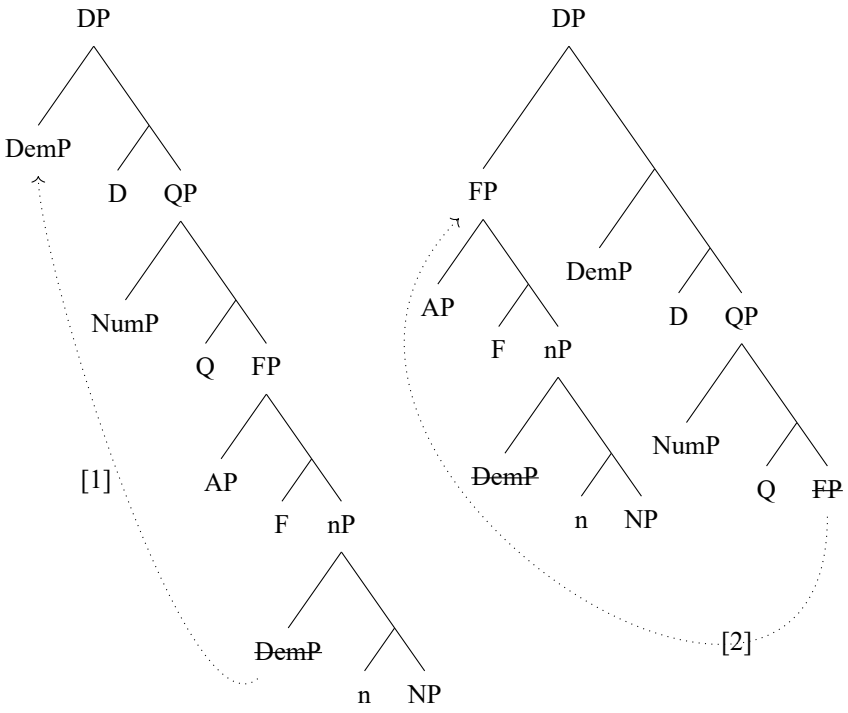
[73] Artiagoitia (2006, 2008) shows that Basque bears postnominal numerals like *bat* 'one' and *bi* 'two', as already discussed in section 3.1.2. However, Dryer (2013d, 2018b) and Cinque (2021) regard such cases as exceptions to the generalization that can be stated with respect to all the other numerals.

[74] The only alternative to DemP raising to Spec,DP thus seems to be its *in situ* position. However, this view represents a further non-parallelism with the clausal domain of V(erb)-initial languages, for which subjects are sometimes assumed to raise to an intermediate position between Spec,*v*P and Spec,TP (Roberts 2017b: 165-166) so that a third alternative does seem to exist in the clausal domain. Roberts (2017b: 163-164) tries to sketch a parallel of this kind with respect to the observations made by Adger (2013) on Scottish Gaelic, but his argument requires an additional functional projection below FP to which DemP must raise.

because of an application of the *in situ* generalization (Alexiadou & Anagnostopoulou 2001, 2007) to the nominal domain, so that at least one between DemP and NP must vacate *n*P.

Due to the aforementioned reasons, the analysis of U20 under a FOFC perspective seems to be both "too strong" and "too weak" (Hawkins 2013). On the one hand, it is too weak because it requires unrelated constraints in order to capture a more severe ban on Num-Dem-N and A-Num-N. Cinque (2005, 2023) and Abels & Neeleman (2012) capture this fact without recurring to FOFC, while Medeiros (2018, 2021) does so by proposing an entirely different mechanism, as sections 3.2.4 and 3.2.5 will further illustrate. On the other hand, the account proposed by Roberts (2017b) is too strong because it rules out linearizations that can be found crosslinguistically, so that FOFC becomes a predictor of their rarity, rather than of their unattestation. A case in point is represented by A-N-Dem-Num, for which a possible derivation is provided in (112), starting from (111).

(112)



According to Roberts (2017b: 181-182), the movement of FP to the outer specifier of DP is a FOFC violation, therefore being as problematic as the Num-N-Dem order derived from the hierarchy in Cinque (2005, 2023). A confrontation with the latter author makes emerge that Roberts (2017b) generally identifies FOFC violations in a way that entails a ban on the *pictures of whom* pied-piping. However, the two concepts do not overlap completely in his analysis, because a single instance of piep-piping can give rise to more than one FOFC violation. In fact, the account of A-N-Dem-Num in (112) differs from Roberts (2017b: 181)

because the latter posits a further string-vacuous movement of NP to the outer Spec,*n*P, given that its absence "would create a further FOFC violation". These observations also point out that the way Roberts (2017b) accounts for FOFC cannot be reconciled with Cinque (2023: Sec. 5.2), because the latter claims that the rarity of FOFC-violating structures is due to a switch from "the less marked *whose-pictures* pied piping in the most deeply embedded domain [...] followed by the more marked *pictures-of-whom* pied piping type at the next higher level" (Cinque 2023: 98-99, italics in the original).

To conclude this section, it is worth highlighting that the roll-up movements which Roberts (2017b) assumes to derive head-final structures are not akin to the Comp-to-Spec movements illustrated in (27) on page 53. Although this fact is never explicitly pointed out, it represents a remarkable difference from the narrow-syntactic explanation of FOFC presented in section 2.3.3 (and pursued by Biberauer 2017b: Sec. 9.4), and it is due to the multiple specifiers that Roberts (2017b) allows as opposed to Kayne (1994) to circumvent the problems discussed in section 2.4.1. If the comparison with Cinque (2005, 2023) is pushed further, roll-up movements can be restated as pied-pipings of the *whose pictures* type where a certain maximal projection targets the outer specifier of the projection by which it is immediately dominated. Although Cinque (2005, 2023) allows NP to raise across multiple projections and then pied-pipe them with this kind of movement, this is not permitted in Roberts (2017b), because in such cases roll-up fails to apply monotonically from the bottom up in the relevant extended projection (Etxepare & Haddican 2017: 140). However, the total application of roll-up/*whose pictures* movements in Roberts (2017b) fails to predict a higher frequency of the derived orders, as opposed to Cinque (2005, 2023). As Tommaso Mattiuzzi (p.c.) remarks, N-Dem-A-Num should be as frequent as N-A-Num-Dem (the only difference being in the absence vs. presence A-movement of DemP), contrary to fact.

3.2.4 Abels & Neeleman (2012) and related works

A formal analysis that rejects both Kayne's antisymmetric approach and the explanation of U20 given by Cinque (2005) is advanced by Abels & Neeleman (2009, 2012). As discussed in section 2.4, they dispute that LCA has effective knock-on restrictions on X-bar theory and thus claim that Kayne's SHCH should be abandoned. According to them, the main flaw on Kayne's argument is that LCA fails to explicitly provide a theory of labeling. For this reason, Abels & Neeleman (2012: 41-42) follow Guimarães (2008) in proving themselves able to

bring back to the picture LCA-compliant structures that Kayne (1994) claims to rule out.

Abels & Neeleman (2009, 2012) thus come to the conclusion presented in section 2.4.1, namely that the hierarchical structure does not need to go forcedly downwards from left to right as in Kayne (1994) and Cinque (2005). Moreover, an effective explanation of U20 based on the typological data reported on page 123 by following Cinque (2005) can be posited by "pruning" Cinque's structures, i.e. after eliminating the silent AgrPs that are supposed to host raised elements (Abels & Neeleman 2009: Sec. 4.4.1).[75]

Of course, one can safely shrink Cinque's structures only if the orders attested on page 123 are still accounted for. Abels & Neeleman (2012: 33) prove their claims by providing an alternative analysis of Cinque's data concerning U20 which is deemed to be better by being more parsimonious on a theoretical level. They begin from the same hierarchical relations that Cinque posits for the elements of the noun phrase, i.e. *Dem* > *Num* > *A* > *N*, where > indicates c-command.[76] However, their refutation of antisymmetry makes them come up with eight base-generated orders that are mutually symmetric. Because of that, Abels & Neeleman (2009: 70), Abels & Neeleman (2012: 68) call their proposal a "symmetric analysis" (see also Ackema & Neeleman 2002: Sec. 2) which is opposed to Kayne (1994). Such eight orders can be illustrated through the following pairs of bracketed structures, where those on the left and those on the right map prenominal and postnominal demonstratives, respectively:

(113)	[Dem [Num [A N]]]	[[[N A] Num] Dem]
	[Dem [Num [N A]]]	[[[A N] Num] Dem]
	[Dem [[A N] Num]]	[[Num [N A]] Dem]
	[Dem [[N A] Num]]	[[Num [A N]] Dem]

As (113) shows, six of these base-generated orders immediately contravene one antisymmetric tenet assumed by Cinque (2005), namely that asymmetrically c-commanding constituents must precede what is c-commanded. This key property of antisymmetric structures is violated whenever demonstratives or numerals are linearized to the right of N, while these two modifiers are not problematic in [Dem [Num [N A]]] and [Dem [Num [A N]]]. However,

[75] Abels & Neeleman (2009: Sec. 4.4.1) actually define the formal operation as *shrinking* (and its reverse as *stretching*), while *pruning* is used within the definition and in the plain text as a synonym. The latter term is preferred here because Heck (2016, 2023) also uses *shrinking* to indicate a totally different syntactic operation to be discussed in section 5.3.2.

[76] In line with Hall (2015: 37 fn. 2) but in contrast with Abels & Neeleman (2012: 33), the 'successor' symbol ≻ will be used here to indicate linear order (i.e. X≻Y should stands for 'X precedes Y') whereas the 'greater than' symbol > indicates structural dominance.

both structures are still not eligible under Cinque's antisymmetric approach because A and N can be both realized as sister nodes that contain pronounced material, so that the structural asymmetry required by the LCA is not necessarily met (see also Holmberg 2017b: Sec. 4.5 and Cinque 2023: Sec. 3.3).

In order to predict all the fourteen orders which are attested according to Cinque (2005), Abels & Neeleman (2012: 33) posit that instances of syntactic movement are possible, as long as they "target a c-commanding position" and "move a subtree containing N". These assumptions, which also prevent the raising of any modifier independently from N, are thus conceptually similar to the ones posited by Cinque (2005), but they still generate more than the six attested orders that do not appear in (113). One problematic example in this respect would be [[[Num [A t_N]] Dem] N], which is not attested but could be generated in Abels and Neeleman's approach as per the assumptions given so far. Starting from the last base-generated structure illustrated in (113), i.e. [[Num [A N]] Dem], one could in fact move N to a c-commanding position where it would dominate all the other elements of U20.

In order to rule out cases like this, Abels & Neeleman (2012: 68) postulate that movement can only occur to the left. Although such a claim traces back to the limits of the human parser discussed in section 2.4.1, it deserves further discussion that will be tackled in section 3.2.5. Still, this restriction makes it is possible to derive the six orders attested in Cinque (2005) that are not listed in (113), by means of movements that take place in one fell swoop.

As anticipated in section 3.2.1, this leads to another crucial distinction from Cinque (2005), namely that Abels & Neeleman (2009, 2012) do not need to assume an exceptional intermediate stranding of A. This follows from the fact that N-Dem-A-Num can be derived from [N [Dem[[A t_N] Num]]] or [N [Dem[[t_N A] Num]]]. The leftward raising of N from one of the structures in (113) further accounts for three out of the five remaining orders, i.e. Dem-N-Num-A, N-Num-A-Dem and N-Dem-Num-A. The remaining two, i.e. A-N-Dem-Num and N-A-Dem-Num, can instead be accounte for by means of leftward movement of {A N} (where the curly brackets indicate that both orders between the two elements is legit).

In sum, the foundational axioms in Abels & Neeleman (2012: 30) set up the basis for an analysis that strongly diverges from the antisymmetric framework, although they need to retain the ban on rightward movement also assumed by Kayne (1994) and Cinque (2005) in order to avoid an overgeneration of orders that are seemingly not attested. Cinque (2023: 52) notes that once movement "is allowed to derive the 'non-symmetrical' orders [...] it is also

available to derive the 'symmetrical' ones" shown in (113) on page 143, "thus exposing the redundancy inherent in any account" that follows the tenets assumed by Abels & Neeleman (2012). However, Abels & Neeleman (2012: 35 ff.) demonstrate to be well aware of this fact, the actual point being that it is not possible to derive any of the orders that are unattested in Cinque (2005) with the premises they set.

Another point worth stressing is that Abels & Neeleman (2009, 2012) and Neeleman (2017) provide no labels for the structures in (113) and the six derived by movement. This follows from the fact that the analysis is deemed to hold "even if no restrictions are imposed on the X-bar theoretical status of modifiers, the range of landing sites for movement or the types of movement involved" (Neeleman 2017: 11 fn. 5). However, Roberts (2017b) and Georgi & Müller (2010) judge the absence of labels in those structures problematic because this makes impossible to state with certainty what is a head and what is a non-head in (113). Roberts (2017b: 178-180) not only observes that it is subsequently hard to evaluate such structures with respect to FOFC, but he also claims that the leftward raisings of N cannot be the result of head movement because they are not sufficiently local.

Locality is however at stake only if one subsumes one universal nominal projection that complies with the FHH (see section 2.2.4), like Roberts does. Georgi & Müller (2010: Sec. 5) argue instead that, once a new way to solve the issues related to the head movement of N is proposed, an explicit LHH-formulation of the structures by Abels & Neeleman (2009, 2012) has the further advantage of offering an independent motivation for the fact that "movement in the nominal projection must involve a subtree containing N", which is instead merely stipulated in Abels & Neeleman (2009, 2012).

Georgi & Müller (2010: Sec. 2) claim that works like Ritter (1991) and Longobardi (1994) (presented in sections 2.2.3 and 2.2.4) provide good arguments for the fact that N can raise as a head, and one can safely extend this to Abels and Neeleman's account of U20. Moreover, the feature(s) triggering the movement of N must be able to percolate, so that the fact that N pied-pipes A can be accounted for (Georgi & Müller 2010: 24).[77] A local chain is however possible only if there is another head position to target, as Roberts (2017b) highlights, but head movement bears some unwelcome theoretical issues discussed by Georgi & Müller (2010: Sec. 3.1) and other sources previously mentioned in footnote 12 on page 37.

[77] See Horvath (2006: 593 ff.) and Georgi & Müller (2010: 24 fn. 28) for an acknowledgement of the problems linked to feature percolation, together with references where the issue is addressed. See also Zelenskii (2020) for a recent redefinition of feature percolation that may overcome some of these shortcomings.

In the system that Georgi & Müller (2010: Sec. 3.2) lay out, these problems can be overcome by proposing that the landing position is created by the moving head itself, which reprojects anew. Any head is endowed with two stacks, each containing a specific set of features ("given the Strict Cycle Condition [in (115)], nonheads cannot have operation-inducing features", Georgi & Müller 2010: 21). By following Heck & Müller (2007: Sec. 2), one stack contains probe features, notated with [∗F∗], which operate under Agree (Chomsky 2000, 2001) in looking for matching features in the c-command domain. The other stack is made of structure-building features, notated with [•F•], which can be divided further into subcategorization features (which trigger External Merge) and movement-type specific specifier features (which trigger Internal Merge). Georgi & Müller (2010: 13) assume that each derivation is regulated by the following principles:

(114) Last Resort:
  a. A syntactic operation must discharge (and delete) [•F•] or [∗F∗].
  b. Only [•F•] or [∗F∗] features that are on top of a feature stack are accessible.

(115) Strict Cycle Condition (SCC):
  Only the head of the present root can have features that trigger operations ([•F•] or [∗F∗]).

(116) Agree over Merge:
  If both [•F•] and [∗F∗] can be discharged, [∗F∗] is given preference.

In Georgi & Müller (2010: 21), the c-command relations consistently assumed for U20 by Cinque (2005, 2023) and Abels & Neeleman (2009, 2012) are nothing but the result of a universal [•*A*•] > [•*Num*•] > [•*D*•] hierarchy on the stack of structure-building features that N bears. Georgi & Müller (2010) thus posit a structure of the nominal domain that follows the LHH and that predates what Cinque (2023: 25) proposes under the FHH. In fact, each nominal modifier of U20 can be (part of) a phrasal subprojection whose head is endowed with its own structure-building features which, however, can never be [•N•], [•A•], [•Num•], or [•D•] (Georgi & Müller 2010: 21 fn. 26). This precondition has three further consequences that put this account on a par with Cinque (2023). The first is the potentially problematic aspect of overlooking the fact that other modifiers of U20 can disrupt the linear continuity between demonstratives and determiners; the second is the implicit endorsement of a unifying classifier-for-the-numeral analysis (see section 3.1.2); while the third is the special status of N, which is the only head in the nominal domain that can move to positions

where it can c-command subprojections that are part of the whole nominal it heads (more on this in section 3.2.5).

N-movement by reprojection is the consequence of a specific type of probe features [∗F∗] that Georgi & Müller (2010) call "Münchhausen features".[78] A Münchhausen feature must have the same value of a corresponding structure-building feature, and its availability "is intrinsically tied to the existence of [such] a [•F•] feature on the same head" (Georgi & Müller 2010: 21 fn. 26).

The mechanism can be illustrated with the derivation of Dem-N-Num-A (see also Georgi & Müller 2010: 22-23). While the stack of structure-building features on N must contain [•A•], [•Num•] and [•D•] (here listed from the topmost to the lowermost) in any derivation where all the elements of U20 are realized, in this specific case the stack of probe features on N contains the Münchhausen feature [∗Num∗], whereas other U20-related probe features are absent.[79] Last Resort in (114) imposes the discharge of structure-building features up to the generation of the unordered structure {Num {A N}}, because [∗Num∗] cannot be checked prior to the External Merge of Num. Once [•Num•] is discharged, Agree over Merge in (116) imposes that [∗Num∗] is preferred over [•D•], which at this point of the derivation is at the top of the stack it is contained in. Given that the Münchhausen feature on N must be checked under c-command, N itself must move to the left (for the reasons assumed by Abels & Neeleman 2009, 2010) and reproject to have [∗Num∗] checked.

As Georgi & Müller (2010: 16) point out, movement by reprojection ensures the LHH of the nominal projection throughout, and the temporary violation of (114) in favor of (115) is justified by the minimal *look-ahead* it ensures for the feature-checking process. This symmetric approach also largely overcomes the already mentioned criticism from Cinque (2023: 52) by minimizing "the occurrence of derivational ambiguities" (Georgi & Müller 2010: 25 fn. 30). Put differently, the attested orders of U20 can be mapped either from base-generated structures or from structures where N or {N A} have raised. Only three cases can be

[78] Münchhausen features are named after a tale where the protagonist, baron Münchhausen, "escapes from a swamp (where he is trapped on the back of his horse) by pulling himself up by his own hair" (Georgi & Müller 2010: 12 fn. 12). See the same footnote for discussion of other syntactic works that predate the proposal discussed in the main text.

[79] The absence of [∗A∗] features is made for the sake of simplicity. However, if one assumes a sisterhood relation between N and AP as for the eight base-generated structures in (113), a [∗A∗] feature could be assumed to be present on top of [∗Num∗], but its effect cannot be discerned by virtue of Agree over Merge in (116), as Georgi & Müller (2010: 14-15, 29-30) observe. Moreover, the presence of multiple [•A•] and [∗A∗] features in the two stacks would lead to the account of ANA or NAA patterns that are not discussed in Georgi & Müller (2010) and that are relevant only in derivations where N does not c-command Num, contrary to what is argued to happen for Dem-N-Num-A.

the result of both options, namely Dem-N-A-Num, A-N-Num-Dem and N-A-Num-Dem.[80] As Georgi & Müller (2010: 25 fn. 30) observe, these three options stand out for being the three cases where "(i) N precedes [Num] and (ii) it is not the case that both D[em] and A precede N."

By following the analyses of U20 proposed by Abels & Neeleman (2009, 2012) and Georgi & Müller (2010), one might expect that the orders that can be mapped from multiple alternate structures are more frequent than the others. However, this expectation is not met in all cases in either analysis. One could alternatively infer that there be a fundamental dividing line between the eight base-generated orders and the six ones derived via leftward movement, the former being consistently more likely than the latter. Merlo & Ouwayda (2018: 27) point out that this is actually not the case (*contra* Neeleman 2017: 14 fn. 6), so that their proposals are not as good as Cinque (2005) at accounting for the data of U20. Cinque (2023: 51) further observes that the frequencies in the pairs of base-generated, mirror-image orders in (113) do not tend to match either. According to Cinque, this fact supports his divide between *pictures of whom* and *whose pictures* derivations, rather than the symmetric analyses presented in this subsection.

3.2.5 Bruening (2018a) and Medeiros (2018)

Georgi & Müller (2010) show that, when applied to U20, head reprojection not only circumvents many of the common problems linked to head movement, but also justifies the pied-piping of nominal modifiers like A, so that the latter type of movement no longer needs to be stipulated as per Abels & Neeleman (2009, 2012). However, there is a further stipulation which the latter three works endorse just like Cinque (2005, 2023) does, i.e. the ban on the independent raising of the modifiers of U20.[81] In such analyses of U20, the latter assumption is necessary so as to account for the fact that the order of prenominal modifiers of U20 is rigidly Dem > Num > A (or a subset thereof, if one of them is postnominal).

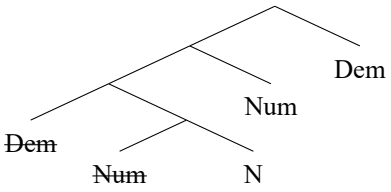
Neeleman (2017: 9-10) provides an amendment to such an issue by illustrating that the eight symmetric orders shown in (113) on page 143 can also be obtained by moving the

[80] In the case of N-A-Num-Dem, there are actually three ways to obtain it from a structure where a subtree containing N raises, so that it is the outcome of four options in Georgi & Müller (2010: 25 fn. 30).

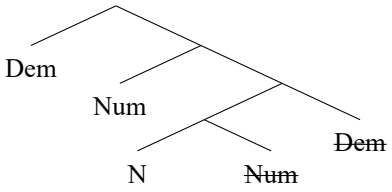
[81] Such a ban is not an axiom in Georgi & Müller (2010), but is still a consequence of a stipulation that makes it apply, namely the claim that no modifier of U20 can have [•N•], [•A•], [•Num•], or [•D•] in its stack of structure-building features (Georgi & Müller 2010: 21 fn. 26). As discussed in section 3.2.3, Roberts (2017b) bans instead the independent raising of adjectives and numerals, but not of demonstratives. Still, the independent movement of the latter is severely constrained.

prenominal modifiers to the right and the postnominal ones to the left. All one has to do is to find a way to legitimize options like those exemplified in (117) while banning their counterparts in (118).

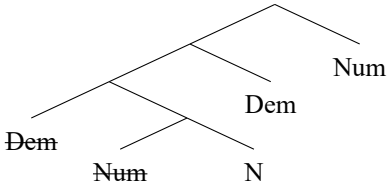
(117) a.



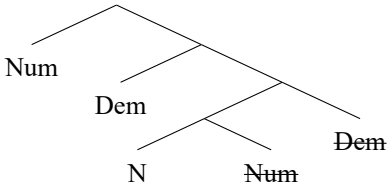
b.



(118) a.



b.



In order to block the "anti-mirror order" effects in (118), Neeleman (2017: 10, 33) adopts the *round-robin constraint* that Starke (2001: 79, 149) formulates as a chain-based extension of Relativized Minimality (Rizzi 1990, 2001):

(119) A chain C_{α} headed by α blocks formation of a chain C_{β} headed by β if and only if

a. C_{α} and C_{β} are of the same type, and
b. all of the links in C_{α} are c-commanded by β and c-command a trace of β.

The application of (119) to U20 suffers however from two problems. The first is acknowledged by Neeleman (2017: 33) as follows:

> [I]t would not be a good idea to try and analyze the mirror effects relevant to Universal 20 using movements, [... because] there is no reason why a trivial chain consisting of a modifier in its base position should be of the same type as a chain created by movement of a lower modifier. Yet this must be the case if we want to account for the fact that prenominally the universal hierarchy of modifiers translates into a universal linear order (Dem-Num-A)[.]

The second problem with the application of (119) to U20 is that it cannot be justified by relying on the extra-grammatical limitations of the human parser discussed in section 2.4.1.

As soon as one assumes (as is often done, see section 3.1) that the modifiers of U20 are phrasal, the latter should undergo movement in the asymmetric way stated by Ackema & Neeleman (2002: Sec. 9). On the one hand, leftward movement is unconstrained except for anti-locality (Grohmann 2012, a.o.), which simply prevents movements that are too local. On the other hand, a highly local movement is exactly the only thing that seems to be allowed for rightward movement, as the latter appears to be phrase-bound. Based on this asymmetry, Georgi & Müller (2010) and other approaches that follow the Lexical Head Hypothesis would neatly predict (117b) but not the phrase-bound counterpart in (117a). Moreover, (118b) would also be allowed as soon as (119) is dispensed with, while (117a) would not.

The analysis of U20 proposed by Bruening (2018a) *de facto* follows the observations just made. The claim is that all one needs to assume to account for the facts is a *Dem* > *Num* > *A* > *N* hierarchy together with the free rightward movement of the nominal modifiers.[82] In Bruening's view, an adequate analysis may accidentally overgenerate, but it cannot undergenerate. Thus, the fact that Cinque (2005: 322 fn. 26) reports the attestation of Num-N-Dem-A in the Eastern Malayo-Polynesian language Kilivila is held as a sufficient argument against the adequacy of the proposals discussed so far, because they do not manage to generate such a fifteenth permutation of U20.

By agreeing with Abels & Neeleman (2009, 2012) on the plausibility of the eight base-generated structures in (113) on page 143, Bruening (2018a) claims that the remaining orders are not the consequence of (a subtree containing) N moving to the left. There are two reasons that make this option theoretically problematic. On the one hand, Bruening (2018a: Sec. 1) agrees with Georgi & Müller (2010: 21) that a phrasal movement of this sort would be an unprecedented stipulation. On the other hand, sections 2.2.4 and 3.1.4 have shown that Bruening takes the Lexical Head Hypothesis as an axiom for the nominal domain. This implies in turn that a moving head N would not manage to find a landing site available amidst it modifiers, and that reprojection *à la* Georgi & Müller (2010) is unnecessary, given that the internal nodes of the structure are all labeled N anyway.

The solution that Bruening (2018a) thus proposes is that N never moves, whereas the nominal modifiers of U20 can be reordered freely to the right as shown in (117b) and (118b), but not to the left of N as in (117a) and (118a). As Bruening (2018a: Sec. 3.3) observes,

[82] Although it is just a draft, the account for U20 proposed by Bruening (2018a) is consistent with the LHH formulated by Bruening, Dinh & Kim (2018) and Bruening (2022), discussed in section 2.2.4 and section 3.1.4. Therefore, his arguments are by no means marginal for a comparison with other analyses of U20.

with these premises the orders Dem-N-Num-A and A-N-Dem-Num can only be generated by not ruling out (118b), so that what Neeleman (2017: 10, 33) formulates in (119) cannot be adopted. The allowance of movements like those in (118b) is what makes Bruening's account generate sixteen permutations, rather than fourteen. One of the two additional permutations is (as anticipated) Num-N-Dem-A, while the other is N-Num-Dem-A. Although Bruening (2018a: Sec. 3.4) is not aware of any attestation of the latter option, he considers this fact not as a problem, but rather as an accidental gap. Potential support in favor of this prediction can be found, however, in Wilhelmsen (2019: 559) and Dryer (2018a: 822), where evidence in favor of the attestation N-Num-Dem-A is provided by the Bantoid languages Mbugwe and Haya, respectively (moreover, the attestation appears as a marked alternative in Lebanese Arabic, Ouwayda & Shlonsky 2017: 183).

Another analysis of U20 that tries to minimize the necessary assumptions in comparison to Cinque (2005) is Medeiros (2018). The latter work is however completely different from the account sketched by Bruening (2018a). In fact, Medeiros (2018: 10) sees grammar as a parser, more specifically as a transducer. The latter is a finite-state machine that relates a string in input with another one in its output by means of an algorithm. According to Medeiros (2018: 4), this machinery becomes a plausible formalization of natural language syntax as soon as one sees both surface orders and base structures as sequences, so that they can be respectively the input and the output strings of the transducer (see Krivochen 2024: 722-723 for related observations). As Medeiros (2018: 6) clarifies (italics in the original):

> Word order and base hierarchy are disconnected sequences, related dynamically. Non-adjacent input elements can perfectly well end up adjacent in the output. Displacement, rather than being the exception, is the rule; *every* element in the surface order is "transformed", passing through memory before retrieval for interpretation [...].

Two further ingredients are fundamental for the proper functioning of what Medeiros (2018) proposes. First of all, the base structure that constitutes the output corresponds to *fseq*, i.e. a universal ordering of functional projections. As already discussed in section 2.3.1, the *fseq* is typical of syntactic models that strongly support the Functional Head Hypothesis (e.g. the cartographic and exoskeletal approaches, see a.o. Putnam 2020 for references and discussion). For the sake of consistency with the usually assumed bottom-up composition of structures, Medeiros (2018: 3) sequences the *fseq* from the bottom up. Given that the *fseq*

for U20 corresponds to the *Dem* > *Num* > *A* > *N* hierarchy assumed by the majority of the approaches discussed so far, its bottom-up sequencing is <N, A, Num, Dem> (Medeiros 2018: 4).

The second ingredient needed for the proper functioning of what Medeiros (2018) proposes is a stack, namely a temporary storage location where the elements in the input can be piled before being mapped to the output. What dictates if an item is directly mapped to the output or first needs to be stacked is the assumed algorithm that rules the transducer. The pseudo-code that Medeiros (2018: 4) provides to illustrate such an algorithm can be adapted as follows:

(120) As long as the input is non-empty, follow the instructions in (120a) and (120b); then move to (120c).

a. If the stack is empty, move onto the stack *I*, i.e. the next item in the input.
If the stack is not empty and *I* does not c-command *S* (where *S* is the item on top of the stack), move *I* onto the stack.
b. If (120a) cannot be satisfied, move *S* from the stack to the output. Repeat this step until the condition in (120a) can be applied.
c. Move *S* from the stack to the output. Repeat this step until the stack is empty.

Crucially, the fourteen orders that are attested in Cinque (2005) converge on one single output, i.e. the <N, A, Num, Dem> sequence mentioned above following Medeiros (2018: 4), while the ten unattested ones (see again page 123) map different output sequences. This can be acknowledged by taking two orders, one that is attested and one which is not, and seeing how (120) parses them (see Medeiros 2018: 4-6 for discussion of all cases).

In the case of Num-N-A-Dem, Num is first placed onto the empty stack (120a). Given that N does not c-command Num, (120a) is satisfied once more and N is piled onto the stack too. Since A c-commands N and the latter is now on top of the stack, (120a) fails and (120b) intervenes by mapping N to the output.[83] At this point, the stack contains an element that A does not c-command (i.e. Num), so that A is piled on top. Since Dem (which is the last element in the input) c-commands all the items in the stack, (120b) maps A and then Num to the output by appending them to N. The eventual parsing of Dem leads to the desired output,

[83] As Medeiros (2018: 4) points out, (120) establishes a last-in, first-out type of sorting. This can be witnessed by the fact that (during the parsing of Num-N-A-Dem) N is the first item to leave the stack, even though it entered it after Num.

i.e. <N, A, Num, Dem>.

In the case of the unattested A-Dem-N-Num, A gets stacked first because of (120a), and then Dem causes the application of (120b). Thus, A is mapped to the output, which is something that already prevents the formation of the desired output sequence. In fact, the output that (120) maps for this permutation is <A, N, Num, Dem>. Medeiros (2018) further observes that the ten orders that according to Cinque (2005) are unattested bear a common pattern. Based on the usually assumed *fseq*, one of the first two items in the input sequence c-commands one of the last two, and there is a third item in between that c-commands both. In the case of A-Dem-N-Num, the three elements that match with this pattern are A, Dem and N.

By adopting a numbering system where the higher the number is, the lower it is in the *fseq*, Medeiros (2018: 3) addresses the just mentioned property of the unattested orders as a "213-avoidance" in "information-neutral word order possibilities". Medeiros (2018) further argues that the latter is not limited to U20, but also applies to other phenomena like the West Germanic verb clusters (which will be briefly discussed in section 5.4), as well as to FOFC (see again sections 2.3.2 and 2.3.3). However, Medeiros (2018: 12 fn. 7) observes that his approach can capture the latter syntactic phenomenon only if one assumes that the output of verbal extended projections (or spines, as per Biberauer, Holmberg & Roberts 2014: 211) is sequenced from the top down, rather than in the bottom-up fashion assumed for U20. Based in the same numbering system just mentioned, FOFC represents in fact a case of "231-avoidance", rather than of "213-avoidance" (see also Svenonius 2024: Sec. 4.2).

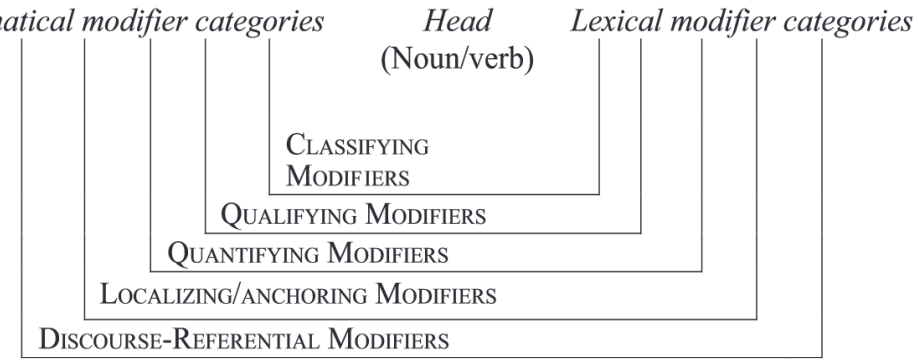
3.2.6 Other relevant works on U20

Rijkhoff (2002: Ch. 7-9) represents a typologically relevant analysis of Greenberg's U20 that predates all the works discussed in the previous subsections and that bears interesting parallels with some of them. Rijkhoff (2015: 654) refines and summarizes his previous publication by providing the layering in (121) on the following page (see Primus 2001: 862 and references therein for predecessors of this idea). The latter is an iconicity principle (Dryer 2018a: 827) that applies to both the nominal and the clausal domains, so that they end up paralleling each other in several respects. In particular, in both domains "[m]odifiers tend to occur next to the part of the expression that they have in their scope", so that it becomes possible to formulate "a formal, meaning-based representation" in which the "modifiers are distributed

over 'nested' layers around the head constituent, reflecting differences in semantic scope" (Rijkhoff 2015: 653-654).

Rijkhoff (2002: 216) further warns that the representation in (121) "must not be confused with some underlying [...] syntactic structure", and does not force grammatical and lexical modifiers to precede and follow the head they modify, respectively. What matters is that the nestedness assumed in (121) be respected in the surface word order in an essentially exceptionless manner (as Dryer 2018a: 827 points out).

(121)



As soon as one follows Rijkhoff (2002: Sec. 7.4) in understanding Dem as a localizing/anchoring (or discourse-referential) modifier, Num as a quantifying one, and A as some qualifying (or classifying) modifier, (121) provides a conceptual structure (Martin et al. 2020: 3) that not only agrees with the *Dem* > *Num* > *A* > *N* hierarchy that (excluding Roberts 2017b) is assumed by all the works discussed so far, but also predicts the eight permutations of U20 that are base-generated in Abels & Neeleman (2009, 2012). Such orders, previously listed on page 143, are repeated in (122) for convenience.

(122)

[Dem [Num [A N]]]	[[[N A] Num] Dem]
[Dem [Num [N A]]]	[[[A N] Num] Dem]
[Dem [[A N] Num]]	[[Num [N A]] Dem]
[Dem [[N A] Num]]	[[Num [A N]] Dem]

Given that Rijkhoff (2002, 2015) argues that (121) is a universal layering that is not strictly syntactic, the eight orders that are derived from it in (122) are regarded as both necessary and sufficient so as to capture the crosslinguistic attestations of U20. Aside from Greenberg (1966–1963) and Hawkins (1983), Rijkhoff's analysis is thus the one which most restricts the number of attested permutations. In this respect, Rijkhoff (2015: Sec. 3.3.2, 2016: Sec. 5)

further contends that the other accounts (discussed so far and to be discussed shortly) report a greater number of attested syntactic arrangements by collecting false attestations, which are to be excluded as soon as functional, semantic, and syntactic criteria are applied together in a manner that is both strict and consistent.

However, Rijkhoff's claims suffer from at least two distinct problems. The first is shared with Abels & Neeleman (2009, 2012), as it consists in observing that (121) predicts an equal crosslinguistic distribution of the eight orders in (122), contrary to fact. In this respect, another observation is worth reporting. Within the sample he scrutinizes, Rijkhoff (2002: 348 fn. 15) does not manage to find two of the orders listed in (122), i.e. A-N-Num-Dem and Dem-A-N-Num, so that the remaining six turn out to be enough to cover the data he collects. Still, Rijkhoff (2015: 654) argues that these two orders are attested, even though they were not spotted by his previous inquiry.

The second problem for Rijkhoff's analysis is raised by Velde (2022: 900-902, 912). Among the languages belonging to the Bantoid genus (as defined in Glottolog, see section 4.2.1), anti-iconic phenomena that violate (121) are the norm. Possessive pronouns represent one first example in this sense. Albeit the latter are localizing modifiers, they occur in "the immediately postnominal position in very many languages", so that they are closer to the noun than the other postnominal modifiers that they should have in their scope, i.e. Num and A (Velde 2022: 912).

The observation just made is closely related to another fact, namely that the original formulation of Greenberg's U20 (presented on page 79) establishes the anti-iconic pattern N-Dem-Num-A as one of the possible options precisely because of a Bantoid language featuring this unmarked order, namely Kikuyu (Greenberg 1966–1963: 87). As a consequence of this, Rijkhoff (2002, 2015) needs to "make sense of these violations of his principle of iconicity" (Velde 2022: 901) and therefore suggests that "it may be the case that certain modifiers are not part of the integral NP but rather appositions" (Rijkhoff 2002: 274-275). However, Dryer (2018a: 827) observes that Rijkhoff offers so "little evidence that the apparent exceptions he is excluding involve apposition" that it soon becomes unclear on what basis much of the data that typologists collect should be discarded. Velde (2022: 901) further argues that "Dryer's skepticism proves to be valid when we consider the Kikuyu language" once again. In fact, every departure from the both default and non-iconic N-Dem-Num-A order "asks for special comma intonation". Thus, "an intonational pattern that is indicative of apposition" (Velde

2022: 901-902) is paradoxically required for any iconic alternative that can be obtained.

Still, the intuition behind (121) is something that is largely agreed upon in one way or another, as already seen. A line of research worthy of mention in this respect is Artificial Grammar Learning (AGL). The latter consists in running experiments to test the participants' expectations and learning strategies when exposed to bits of a novel (artifically constructed) language (Culbertson & Adger 2014, Culbertson 2017, Culbertson et al. 2020, Martin et al. 2019, 2020). In experimental contexts where the elements of U20 are learnt, participants are generally observed to prefer the eight orders in (122) or in a subset thereof, depending on the number of modifiers to which they are exposed. These facts thus support once again the often assumed hierarchical relations in (121).

In an experiment where English native speakers had to learn the postnominal nature of a pair of nominal modifiers in the constructed language, the results lead Martin et al. (2020: Sec. 2.3) to observe that "participants preferred noun phrases with the order N-Adj-Dem [...] over the order N-Dem-Adj", as expected. However, they "showed no preference between orders N-Adj-Num [...] and N-Num-Adj [...], nor any preference between orders N-Num-Dem [...] and N-Dem-Num". Martin et al. (2020: Sec. 5) follow Culbertson & Adger (2014) and Martin et al. (2019) in arguing that this effect hints at the presence of "more structure intervening between Dem and Adj than between Dem and Num, or Num and Adj", even when no other modifier of N is overtly present aside from Dem and Adj. In addition to this possible explanation of the facts, one can start from the examples in (123) and (124) to make a different observation:

(123)	a.	<Dem>	**N**	<Dem>	**A**	<Dem>	**Num**	<Dem>
	b.	<A>	**N**	<A>	**Num**	<A>	**Dem**	<*A>
	c.	<Num>	**N**	<Num>	**A**	<Num>	**Dem**	<Num>
(124)	a.	<Dem>	**N**	<Dem>	**Num**	<*Dem>	**A**	<Dem>
	b.	<A>	**N**	<A>	**Dem**	<A>	**Num**	<A>
	c.	<*Num>	**N**	<*Num>	**Dem**	<Num>	**A**	<Num>

The three cases in (123) correspond to the homomorphic arrangements of two overt (and boldfaced) postnominal modifiers of U20, with the third modifier placed wherever it can occur, no matter if the outcome is iconic or not. In fact, the null hypothesis is that learners have no bias with respect to the position of silent elements. In (124), the two postnominal modifiers are instead reversed so as to be surely non-homomorphic (for discussion of the

latter term, see footnote 58 on page 122). As long as the attested orders are those reported in Cinque (2005), the former group shows one single permutation in (123b) that is ruled out, whereas (124) shows three, two of which occuring in (124c). Crucially, the latter is the only non-homomorphic option that turns out to be significantly disfavored by the participants in the experiment discussed by Martin et al. (2020: Sec. 2.3, 5). Thus, whatever the property that makes participants disprefer (124c) is, it is likely to be the same property that rules out Num-N-Dem-A and N-Num-Dem-A (*contra* Bruening 2018a).

Compared to Rijkhoff (2002, 2015), the analysis of U20 provided by Steedman (2020) represents the opposite extreme. Even though the two analyses agree on the unmarkedness of the eight orders listed in (122), Steedman (2020: Sec. 4) assumes a formalism whose fundamental principles allow twenty-two out of the twenty-four mathematically possible permutations of U20. In the admitted cases, *at least one* of the elements of U20 is immediately preceded or followed by another element that, according to the usual $Dem > Num > A > N$ hierarchy, either immediately dominates it or is immediately dominated by it. Thus, Num-N-Dem-A and A-Dem-N-Num are the only two orders that must be excluded, as they are the only cases where *none* of the elements of U20 satisfies the combinatory principle just mentioned (Steedman 2020: Sec. 5).

Steedman (2020: Sec. 5-6) argues that his claim is empirically verified as soon as one no longer distinguishes data based on their markedness, so that linearizations obtained by means of focalizations and topicalizations are as good as those that require none. What really matters is whether or not an order can be generated under any circumstances, because the fundamental principle that Steedman (2020: Sec. 4.4) proposes is a formal, exceptionless universal (Steedman 2020: Sec. 2). To support this view, Steedman (2020) merges the analysis given by Cinque (2005) with the one provided by Nchare (2012: Ch. 3), whose main focus is Shupamem, a Bantoid language spoken in Cameroon.[84] The latter language allows a lot of syntactic variation in the nominal domain, up to the point of showing seven alternatives that Cinque (2005) would rule out (due to their markedness) while Steedman's model predicts them (together with an eighth one, i.e. N-Num-Dem-A).

Given the range of the admitted data and the argued solidity of the formal universal that is assumed, even the slightest evidence against the utter ban of the two aforementioned permu-

[84] This language is recorded as Bamun by both Glottolog (Hammarström et al. 2024) and WALS (Dryer & Haspelmath 2013). As chapter 4 will show, this is one of the reasons that makes the identifiers known as "glottocodes" (Forkel & Hammarström 2022) extremely useful. In the case of Shupamem/Bamun, the glottocode is bamu1253.

tations would represent a serious issue for Steedman (2020).[85] A potential problem of this kind is indirectly provided by Joo (2015: Sec. 3-4), where Korean native speakers are asked to participate in an experiment aimed at identifying the most unmarked linearization of U20 for that language. In such a context, participants judge A-Dem-N-Num as marked, though still acceptable and less marked than A-Num-Dem-N. Although the markedness of both orders can be taken as supporting any analysis discussed so far, it represents a twofold challenge for Steedman (2020). On the one hand, A-Dem-N-Num should be rejected completely, so that A-Num-Dem-N should result as a better alternative on the other hand.

This chapter will conclude by presenting the analysis of U20 provided by Dryer (2018a). To account for the data he collects, Dryer (2018a: 801) proposes five principles that are descriptive and breakable, as they "are in themselves explanatory only in a superficial sense. To fully explain the relative frequencies, one needs to provide explanations" for such principles, which are only posited to "test how well they account for" the data (Dryer 2018a). The first three descriptive principles *de facto* establish aspects that have already been pointed out for many of the analyses of U20 discussed so far. The first two principles in Dryer (2018a: 801) are "iconicity principles", whose outcome corresponds to (121), discussed on page 154 by following Rijkhoff (2002, 2015). The third principle acknowledges instead the asymmetry whereby "exceptions to the iconicity principles will occur only with postnominal modifiers".

Both the fourth and the fifth principles presented in Dryer (2018a: 801) deserve instead further discussion. The former captures the fact that "[n]oun-adjective order is preferred over adjective-noun order". Given the high crosslinguistic frequency of Dem-Num-A-N, this asymmetry can appear surprising at first. However, it can be seen as a typological tendency that follows from Universal 18, given by Greenberg (1966–1963: 87, 111):

(125) When the descriptive adjective precedes the noun, the demonstrative and the numeral, with overwhelmingly more than chance frequency, do likewise.

As Dryer (2018a: 828) observes, the formulation in (125) manages to capture the crosslinguistic rarity of the attested permutations where A is prenominal whereas Num and/or Dem

[85] One can easily extend a similar argument with respect to Steddy & Samek-Lodovici (2011). As already said in section 3.2.5, Wilhelmsen (2019: 559) and Dryer (2018a: 822) provide evidence for the attestation of N-Num-Dem-A in the Bantoid languages Mbugwe and Haya, respectively (and it appears as a marked alternative in Lebanese Arabic, Ouwayda & Shlonsky 2017: 183). Regardless of whether they are marked constructions or not, they seriously challenge the proposal by Steddy & Samek-Lodovici (2011), as in their model N-Num-Dem-A and A-Num-Dem-N cannot be generated at all, not even as sub-optimal permutations of U20.

are postnominal, i.e. Dem-A-N-Num, Num-A-N-Dem, A-N-Num-Dem and A-N-Dem-Num. Moreover, these four orders of U20 "are all less frequent than the corresponding orders with the adjective following the noun", i.e. Dem-N-A-Num, Num-N-A-Dem, N-A-Num-Dem, N-A-Dem-Num (Dryer 2018a: 828).

Two observations are worth mentioning in this respect. The first is that the rarity of Dem-A-N-Num, Num-A-N-Dem and A-N-Num-Dem can also be captured in an approach *à la* Abels & Neeleman (2009, 2012). All one needs to argue is that there is a significant set of languages that realizes both the prenominal adjective and (at least) one of the postnominal modifiers of U20 by means of syntactic heads that are consequently constrained by FOFC (see Bernardi 2022 for related discussion). This would also provide an explanation for the aforementioned difficulty faced by Rijkhoff (2002: 348 fn. 15) in finding attestations of A-N-Num-Dem and Dem-A-N-Num.

The second observation is that Dryer (2018a: 828) sees Greenberg's U18 in (125) as a consequence of both the fourth and the fifth principles given in Dryer (2018a: 801). This thus allows to turn to the latter of the two, which establishes that "the demonstrative, numeral, and adjective tend to all occur on the same side of the noun". Given that it takes inspiration from the concept formulated by Greenberg (1966–1963) and discussed in section 2.2.1, Dryer (2018a: 801) defines this fifth principle as *intracategorial harmony*, which summarizes the fact that N-A-Num-Dem and Dem-Num-A-N are by and large the most common orders of U20 crosslinguistically.

Dryer (2018a) further stands out for providing a typological inquiry of U20 that samples 576 distinct languages (fully reported in Dryer 2018b). In opposition to Cinque (2005), Dryer (2018a: Sec. 6) argues that eighteen orders of U20 are attested. With the exception of Num-N-Dem-A (whose frequency is superior to the rarest order also attested by Cinque, i.e. A-N-Dem-Num), the additional orders that Dryer (2018a: Sec. 6) finds are less frequent than those on which he and Cinque (2005, 2021, 2023) agree upon, and correspond to the following: Dem-A-Num-N, Num-Dem-A-N and N-Num-Dem-A.

For his part, Dryer (2018a) analyzes the crosslinguistic data based on semantic categories, and highlights two methodological problems for Cinque (2005) in this respect. On the one hand, "Cinque does not provide information on the total number of languages in his sample" (Dryer 2018a: 822 fn. 25), while on the other hand he does not explicitly state the criteria adopted for his data collection (Dryer 2018a: 804). In absence of such methodological clari-

fications, Dryer (2018a: 804, 824) supposes that Cinque (2005) selects the empirical data by resorting to syntactic categories, and that this methodological distance is what underlies the different number of attested permutations in the two samples (see Dryer 2018a: Sec. 4, 6 for further discussion).

Dryer's supposition is not unmotivated, given that Cinque (1999) is a work that has fostered a revival of the "[u]niversality of linearization **and** categorization" in the generative tradition (Wiltschko 2014: 11, boldface in the original).[86] However, the methodological contrast between Cinque and Dryer is not as sharp as argued by the latter. In fact, Cinque (2023: 35-36) supports his own theoretical claims by resorting to Cinque (2021), where the collected languages are fully listed and classified based on their genealogical relatedness, as it is done in Dryer (2018b) (see Cinque Forthcoming: 1 fn 2 for further discussion). Moreover, in Cinque (2021) the basic order of U20 is determined for each language by resorting to a methodology that largely overlaps with the one adopted by Dryer (2018a,b). This fact is explicitly admitted in Cinque (Forthcoming: 1-2 fn. 3), as the following quotation further clarifies:

> I will keep to the traditional terms demonstrative, (cardinal) numeral and adjective despite the apparently different morphosyntactic category with which each may be expressed [...] This should not be interpreted as a major problem for the present analysis, as the different morphosyntactic forms in which they appear in different languages can be taken to correspond to the same abstract modifier type on the basis of their semantics and especially on the basis of their order with respect to each other, possibly a consequence itself of their relative semantic scope [...].

Given that Cinque (2023) argues that the attested permutations of U20 are the same fourteen that are reported in Cinque (2005), the disagreement with Dryer (2018b) is not determined by some different sampling methodology between the two, in opposition to what Dryer assumes. On the contrary, Cinque (Forthcoming: 1) contests the four additional attestations reported in Dryer (2018a: Sec. 6) by arguing that the data provided in this respect are "far from constituting clear counterevidence". According to Cinque (Forthcoming: 18), such permutations do not represent the "**unique** canonical order (in an unequivocal way)" in any of

[86] Because of what has been pointed out in the main text, Wiltschko (2014: 11-19, 39-41) argues that Cinque (1999) and related works endorse what is often addressed as the *Universal Base Hypothesis*, which consists in positing a single functional structure that holds for all languages.

the languages where they are found (boldface in the original).

In conclusion, what has just been reported paves the way for a comparison between the data collected by Dryer (2018a,b) and Cinque (2021), and the methodology to be discussed in chapter 4 will show how this is performed so as to come up with the sample presented in appendix A.

Chapter 4

Novel typological observations

This chapter builds upon the content of Bernardi (2025). Hence, it presents an inquiry aimed at collecting languages that have been classified in the literature with respect to both the basic order concerning Greenberg's U20 and the unmarked order of arguments in mono-transitive clauses. The goal is to create a convenience sample (Körtvélyessy 2017: 127) whose amplitude be comparable to other samples related to U20 (e.g., Dryer 2018b, Cinque 2021), as further discussed in section 4.2. Appendix A lists the languages collected according to the methodology to be laid out in section 4.1, while section 4.3 assesses whether it is possible to observe some parallels across the two domains based on this sample of languages. After a summary of the observed tendencies in section 4.4, the focus of section 4.5 will shift towards a parameterization that represents a first attempt to capture the facts in a formal and concise way.

4.1 Constructing the database

This section will be concerned with the methodology adopted to construct the sample reported in appendix A. The discussion in section 4.1.1 focus on WALS and its limits, thus motivating the integration of additional samples that will be presented and discussed in section 4.1.2. The method that will be adopted to classify the basic word order in mono-transitive clauses will be supported by the generalization in section 4.1.3 and the observations to be made in section 4.1.4.

4.1.1 Sources and methods: WALS as a starting point

Given that the objective of the present chapter is inspired by typological facts that are neglected by Roberts (2017b: 155-158), a first exploratory observation has been made by using the same resource used by him, namely the World Atlas of Language Structures (WALS, Dryer & Haspelmath 2013). With respect to the nominal domain, WALS presents samples related to the order of Demonstrative and Noun (Dryer 2013c), Numeral and Noun (Dryer 2013d) and Adjective and Noun (Dryer 2013b). Given that WALS offers the opportunity to cross-reference and visualize data coming from up to four samples present in the database, the three that have just been mentioned have been cross-referenced together with Dryer (2013f), which is concerned with the order of Subject, Object and Verb in mono-transitive clauses whose arguments are overtly realized and are not pronouns.[87]

It should be noted that Roberts (2017b: 155-158) does cross-reference the samples that WALS provides for the nominal domain, but it seems that Dryer (2013f) is kept separate from this operation. In fact, Roberts (2017b: 155-158) presents the order of the arguments in the clause without mentioning the source generally used for the purpose, although one can infer from Roberts (2017b: 157 fn. a) that the data provided are taken from Dryer (2013f) and then fine-tuned by means of complementary sources.

However, using WALS to extend Roberts's search for parallels between the nominal and clausal domains is an effort that presents several issues. First, the classification made by Dryer (2013f) with respect to the word order in the clause suffers from the problems already discussed in section 2.2.2. A second issue is that nothing more can be said with respect to Num-N-A-Dem and Num-A-N-Dem than Roberts (2017b: 155-158), because he relies on the same database. Although the more than 70 languages he gathers represent a noteworthy sample, Cinque (2023: 35-36) claims to have gathered more than 270 languages whose basic order of U20 is either Num-N-A-Dem or Num-A-N-Dem, which means that there might be around 200 unanalyzed counterexamples to what WALS allows to observe.

A third problem is that "WALS contains information on over 2600 languages overall, but most chapters cover only a few hundred languages" (Ivani & Bickel 2023: 20). This fact does not necessarily represent an issue on its own, given that each sample on WALS contains more than 100 languages selected so as to enhance internal consistency and both areal and

[87] In the present note, the word *WALS* can be clicked so as to open the result obtained by cross-referencing Dryer (2013c,d,b,f).

genealogical diversity (see Ivani & Bickel 2023: 20-21), which are factors that are also taken into account in crosslinguistic surveys performed by syntacticians (see Pleshak & Polinsky Forthcoming: Sec. 2.2 and references therein for discussion). However, cross-referencing several chapters from WALS can worsen the situation, because a certain language is subsequently represented only if it is valued in each of the original groupings used for the scrutiny. Differently put, each initial sample can be viewed as a set of languages, while the cross-comparison represents the intersection of all the sets of languages at hand, so that the overall number of elements thus included in the intersection can be dramatically reduced.

Cross-referencing chapters from WALS also raises another problem, namely that languages are divided into the partitions determined by those used in the initial sets, so that the number of languages linked to one of such composite values can be so small that any inferrable generalization is rather tentative. By taking the present survey as an example, languages where adjectives are the only prenominal modifiers of U20 can be valued as N-Dem/N-Num/A-N/VSO, N-Dem/N-Num/A-N/SVO, N-Dem/N-Num/A-N/SOV, and so on.[88] The data from such a specific subset of U20 show no occurrence of SOV languages, but the attested alternate cases are only found among 13 languages in total (in particular, 11 languages are SVO, while the remaning two are VSO and alternating between SVO and VSO, respectively).

Another relevant problem linked to those pointed out for the aforementioned cross-referencing is best explained by the following quotation taken from Cinque (2013: 52) (italics in the original):

> [C]orrelation pairs, though important, do not suffice to reconstruct the 'ideal' harmonic orders. They fall short of giving the *total* order of functional heads, arguments, circumstantials, and modifiers of the clause, and of the other major phrases in 'head-initial' and 'head-final' languages.[...] Exclusive focus on correlation pairs can even mislead one into attributing to the same type word-order types that should be kept distinct. To take one example, if one considers only the orders of *pairs* of elements like NA/AN, NNum/NumN, NDem/DemN,

[88] Another related issue is that the initial chapters may contain certain properties that are so rare that they give rise to several uninformative zeros in subsequent cross-referencing operations. The most notable example in this sense comes from Dryer (2013d: Sec. 1), where only 2 of the 1154 languages evaluated there belong to one of the four established partitions, namely a subset whose members are shown to never modify the noun directly, but rather use cardinal numerals exclusively in adverbial constructions that affect the event denoted by the matrix verb. See Gil (2001: Sec. 3) for related discussion.

> without considering their *total* order, one is led to put three languages like Lalo [...], which has N A Dem Num, Luo [...], which has N Num A Dem, and Gungbe [...], which has N A Num Dem, in one and the same class, as all of them are: NA, NNum, NDem. Yet, while the order found in Gungbe is the overwhelmingly prevalent postnominal order of these elements, the orders found in Lalo and Luo are quite rare in the languages of the world [...]. Thus, one runs the risk of not singling out the correct subtypes and of misrepresenting the number of languages belonging to each. Cases like this, where attention is limited to lists of word-order pairs of elements, rather than to the complete sequence of these elements in each phrase, are unfortunately the norm.

The combination of samples on WALS in relation to the nominal modifiers perfectly reflects the quoted problem.If all modifiers are prenominal, one could argue that their total order is necessarily Dem-Num-A-N based on Cinque (2021) (though see Dryer 2018a,b for the claim of rare Dem-A-Num-N attestations). In a similar fashion, there are cases where one postnominal modifier and two prenominal ones (or vice versa) arguably lead to one single order out of the two possible alternatives. This is the case of N-Dem/Num-N/A-N, which arguably hints at the total order Num-A-N-Dem since A-Num-N-Dem is never found (Cinque 2021, Dryer 2018b). However, there are cases where both options are attested (as for N-Dem/N-Num/A-N, which could refer to both A-N-Num-Dem or A-N-Dem-Num), and the situation is even worse when all modifiers are postnominal, as it has already been highlighted by quoting Cinque (2013: 52). In addition, nothing guarantees that the partial order between one modifier and the noun remains unchanged when they co-occur with other modifiers. This is a doubt that cannot be resolved by a cross-checking on WALS, but rather invites the use of sources that explicitly take into account the total order under consideration, as the next sections will illustrate.

4.1.2 Sources and methods: Beyond WALS

With respect to the nominal domain, the sampling provided in the present work is taken from the database of languages that Cinque (2023: 35-36) collected for his most recent analysis of Greenberg's Universal 20, namely Cinque (2021). The latter sample is still in expansion, and its latest developments can be retrieved at ARCA, the free institutional archive of the

university Ca' Foscari of Venice (Italy).[89] Cinque (2021) contains over two thousand languages and is an (almost) proper superset of Dryer (2018a,b), which is in turn a sample of 576 languages on the same topic.[90] As already observed on page 160, the methodologies that guide the two samples largely overlap, so that it is legitimate to compare the data therein (as Cinque Forthcoming does), and at the same time it is not surprising that the two sources provide classifications that are by and large convergent.

With respect to the word order in mono-transitive clauses, Dryer (2013f) presents several issues already discussed in section 2.2.2, two of which are crucial. First, it is a sample in the WALS database (Dryer & Haspelmath 2013), which is in turn monocategorizing. This means that each linguistic phenomenon is mapped "onto a set of concrete types (categories)" (Ivani & Bickel 2023: 3), so that each item is assigned to exactly one type. Thus, Dryer (2013f) classifies each language as bearing a specific word order (or to none). However, this operation can be 'procrustean' in the sense of Krivochen (2018a: 240), i.e. it can force a specific classification for a certain language, even though the latter would receive a different classification in a multicategorizing database.

As already pointed out in section 2.2.2, the second problem in Dryer (2013f) is related to the way in which Subject and Object are defined. However, all issues related to the nature of the direct arguments can be set aside (momentarily) as soon as one defines the crosslinguistic typology of mono-transitive clauses by exclusively focusing on the position of the verb itself. Hengeveld, Rijkhoff & Siewierska (2004: 542) represent one possible starting point for defining such an approach:

> [W]e use a three-way typology of predicate-initial, predicate-medial and predicate-final clausal orders. This three-way typology based on the position of the predicate has two major advantages over the six-way typology. First of all, since languages exhibit more variation in the location of the transitive arguments relative to each other than in the location of the predicate phrase relative to both of the arguments [...], it is easier to assign a basic order to languages in terms of the three-way typology than in terms of the six-way typology. Sec-

[89] Moreover, a systematization of such data is under development at *TerraLing* (see Koopman & Guardiano 2022).

[90] Note however that this does not imply that Cinque (2021) is using the data from Dryer (2018b) as the main source for his classification. On the contrary, the statements from Dryer (2018a) are always confronted with other works, and relevant discussions (and examples) are often provided in a footnote. It is therefore not surprising that a handful of languages collected by Dryer (2018b) have not yet been reported by Cinque, or that the two provide diverging classifications.

> ondly, of the three logically possible locations of the predicate phrase in transitive clauses, the predicate-initial and predicate-final positions are directly applicable to the classification of clause ordering in intransitive clauses. Consequently, the predicate-based typology allows one to capture any consistencies in the ordering of transitive and intransitive predicates that may obtain in a transparent manner.

The first advantage highlighted in the quotation above represents a good point which will find a further empirical motivation in sections 4.1.3 and 4.1.4, while the second advantage reported there is problematic. In fact, the discussion in section 2.2.2 showed that even Dryer (2013f: Sec. 4) acknowledges that in a single language intransitive and mono-transitive verbs can display different word orders with respect to the arguments they select, so that eventual overlappings need to be treated as carefully as diverging cases. Anyway, it is important to note that this three-way typology keeps the advantages of the four-way one proposed by Dryer (2013a) and eliminates the distinction between SVO and OVS languages, which have been pointed out as problematic orders for the various reasons discussed in section 2.2.2. One might argue that such a move down to three options is bad, because the arguments reported in section 2.2.2 pave the way for more fine-grained distinctions between different types of languages that have been mistakenly labeled either 'SVO' or 'OVS'. However, another perspective on the same phenomena expects them to behave non-uniformly as opposed to the V-initial and the V-final languages, which are on the contrary predicted to be more consistent in some typologically relevant sense (see Dryer (1991: 445) and references therein). And in fact, this is what the remainder of the chapter will contend.

Although the three-way typology proposed by Hengeveld, Rijkhoff & Siewierska (2004: 542) can be applied to the data from WALS (Dryer 2013f), the issue linked to its monocategorizing nature (Ivani & Bickel 2023) are not solved. In particular, a remark needs to be dedicated to those languages that show some internal variation that involves the surface displacement of the verb with respect to its arguments. Would a VSO language that shows some relevant alternation with SVO be counted as verb-initial (henceforth, V-initial) or V-medial? According to Hengeveld, Rijkhoff & Siewierska (2004: 542), it has to be either one of the two, provided that such an order is twice as frequent as opposed to the other. If this criterion is not met, then the language is labeled as having 'no basic order'. Thus, the problems derived from the monocategorizing nature of Dryer (2013f) are seemingly enhanced by the three-way typology proposed by Hengeveld, Rijkhoff & Siewierska (2004: 542).

However, Grambank (Skirgård, Haynie, Blasi, et al. 2023, Skirgård, Haynie, Hammarström, et al. 2023) is a different database that largely solves this issue by being multicategorizing. As Ivani & Bickel (2023: 4) explain, "[w]here monocategorization reduces a domain to a single type, multicategorization reduces the domain to an intersection of multiple (overlapping) categories". Thus, Grambank investigates the word order of pragmatically unmarked mono-transitive clauses that contain lexically expressed arguments by asking whether a language can be V-initial (*GB131*), V-medial (*GB132*), or V-final (*GB133*).[91] Given that these three options are tackled separately, a language can be marked positive for more than one of these alternatives at the same time, so that the three-way typology from Hengeveld, Rijkhoff & Siewierska (2004: 542) can be taken as a starting point and further refined in ways to be discussed in sections 4.1.3 and 4.1.4.

With these premises in mind, the data related to the nominal domain (Cinque 2021, Dryer 2018b) have been cross-referenced with the ones concerning the clausal domain provided by WALS (Dryer 2013f) and Grambank (*GB131*, *GB132*, *GB133*). Several criteria have been adopted so as to enhance the reliability of the sample (Gast & Jędrzejowski 2022: 598). First, only languages reported by either Cinque (2021) or Dryer (2018b) were first considered. Sign Languages mentioned by Cinque (2021) were excluded because their syntax is characterized by some simultaneity (see e.g. Slonimska, Özyürek & Capirci 2022) which is in contrast with the linearity of spoken languages. Of the languages thus filtered, if neither Grambank (Skirgård, Haynie, Blasi, et al. 2023, Skirgård, Haynie, Hammarström, et al. 2023) nor WALS (Dryer 2013f) could provide the relevant data for their clausal word order, they were discarded. If WALS and Grambank provided opposing generalizations that could not be judged by means of a third source (or by checking the mentioned ones), the language was excluded from the sample. If a similar conflict could not be solved between the data provided by Cinque and Dryer, the language was not listed either.

Languages that satisfy all these criteria were then grouped together based on the way they linearize the nominal modifiers captured by Greenberg's U20. Given that the syntax of the nominal domain is not necessarily more stable crosslinguistically than the clausal domain is, languages can show a flexible word order as long as nominal modifiers are involved. Although some of the languages collected may feature some degree of internal variation

[91] Both WALS (Dryer 2013f) and Grambank (Skirgård, Haynie, Blasi, et al. 2023, Skirgård, Haynie, Hammarström, et al. 2023) explicitly aim to exclude pronouns, null anaphora, topicalization and focalization from the relevant samples.

in the latter respect, this could not be captured in the present sampling. In fact, assigning multiple orders of U20 to a single language would offset the calculations, because each copy of the same language would count as one distinct language. For this reason, the sample in appendix A is necessarily monocategorizing (Ivani & Bickel 2023) with respect to the elements of Greenberg's U20. Still, this problem has been curbed by avoiding languages whose way of linearizing nominal modifiers was highly uncertain.

The need for a convergence of views between Cinque's and Dryer's cataloging inevitably implies that the languages in appendix A are distributed among 14 different linearizations of Greenberg's U20. As already pointed out in section 3.2.6, there are eight languages that Dryer (2018a,b) assigns to four further possibilities, while Cinque (2023: 110 fn. 1, Forthcoming) claims they are spurious cases that need to be inserted among the 14 options on which he and Dryer agree. Because of this conflicting views between the two authors, such languages have been excluded from the present sample without trying to take a stance on the debate related to their classification.

As the remainder of the chapter will show, the gathering of a convenience sample (Körtvélyessy 2017: 127) based on very large databases like Cinque (2021) and Grambank (Skirgård, Haynie, Blasi, et al. 2023, Skirgård, Haynie, Hammarström, et al. 2023) allows to check if the correlations that can be observed in the sample from Roberts (2017b: 155-158) hold even after collecting more languages. Moreover, this empirical verification can become part of a broader inquiry aimed at verifying if similar correlations with the clausal domain can be formulated for other linearizations of U20, and what speculations can be made as a consequence.

4.1.3 One first generalization

Languages sampled according to the aforementioned criteria have been subsequently valued for the way they linearize the arguments in a mono-transitive clause. One first general trend that can be observed is that, although theoretically possible, no linearization of Greenberg's U20 forms a subset where only V-initial and V-final langauges are found (let alone equally distributed). In fact, SVO languages can be found across all the orders of U20, and they are rare only if only one of the alternatives predominates.

The tendency that can be observed in the sample thus mirrors the fact that on Grambank each language is positive for both V-initial and V-final patterns only if this also holds for

the V-medial one.[92] This is an observation that is also alluded to on WALS (Dryer 2013f: Sec. 3). When dealing with "languages with two primary alternating orders", Dryer lists five possibilities that comply with what has just been said: (i) SOV or SVO; (ii) VSO or VOS; (iii) SVO or VSO; (iv) SVO or VOS; (v) SOV or OVS.[93] This crosslinguistic tendency with respect to U20 finds empirical support not only language-internally, but also in a diachronic perspective. In fact, the evolution of word order proposed by Gell-Mann & Ruhlen (2011: 17291) highlights that a language cannot become V-initial if it was V-final in an immediate precedent evolutionary state. Thus, they claim that a VOS language can only originate from an ancestor that was either SVO (namely, the most common V-medial option) or VSO (namely, already V-initial).[94]

4.1.4 Mono-transitive clauses: back to six types?

Data provided by WALS and Grambank were then classified along the three-way typology used by the latter. Such three options were thus scored by means of numbers: 1 for V-initial, 2 for V-medial and 3 for V-final. Languages characterized by the V2-phenomenon were assigned to one of these types by following Holmberg (2015) and S. Müller (2023a: 22). In opposition to Hengeveld, Rijkhoff & Siewierska (2004: 542), languages that varied among two possible types have not been classified as having "no basic order". By virtue of the multicategorizing nature of Grambank discussed by Ivani & Bickel (2023: Sec. 1), the intermediate possibilities for a single language could be V-initial and V-medial but not V-final, which was scored 1.5, or V-medial and V-final but not V-initial, which was labeled with a 2.5, or showing all the three options at once, thus being classified as a 0.

---

92 Although no counterexample is found in the data used for the present inquiry, a specific search made in this respect throughout all the Grambank data shows three instances of this kind that would need further check (that is, languages where V-initial and V-final patterns are found with no pragmatically unmarked V-medial counterparts). Such languages are Antankarana Malagasy (anta1255), Mocoví (moco1246), and Tule-Kaweah Yokuts (tule1245). These results can be obtained by first combining Grambank features GB-131 (which parameterizes languages with respect to the V-initial position in unmarked transitive clauses) and GB-133 (does the same with respect to the V-final position), and then comparing the list of 61 languages positive for both parameters with the values listed in GB132 (Grambank's feature focusing on the unmarked V-medial position).

93 See also Steele (1978), where SVO is already pointed out as the only recurring alternate ordering for the SOV, VSO and VOS basic orders.

94 Walkden (2019: 7-8) highlights that Gell-Mann & Ruhlen (2011: 17291) hold a position against uniformitarianism by claiming that diachronic changes from VO to OV are exceptional if not utterly absent, a fact which finds however relevant counterexamples discussed by Walkden (2019: 8 fn. 15). The observations reported in the main text do not force a commitment to a view *à la* Gell-Mann & Ruhlen (2011: 17291), since a uniformitarian perspective can be maintained with the proviso that V-medial languages represent an intermediate stage in both directions, be it from V-initial or V-final languages.

Hence, the three-way typology proposed by Hengeveld, Rijkhoff & Siewierska (2004: 542) becomes more flexible, up to the point that could be termed as a six-way typology in disguise. Languages that are purely V-initial (type 1) can be grouped together or be kept apart from languages that also allow V-medial alternatives (thus being of type 1.5), and the same can be done with respect to verb-finality (i.e. with types 3 and 2.5). This also implies that only languages that are listed with a 0 could be termed as having "no basic order" in the present work, thus restricting the number of languages that comply with this classification as opposed to what Hengeveld, Rijkhoff & Siewierska (2004: 542) suggest.

4.2 A comparison between the present sample and others

The merger of data from different sources does not come without problems, in particular for matters of comparability. This is what the present section will discuss, by starting from the seemingly harmless problem of language classification. In fact, while Cinque (2023), Dryer (2018a,b) and WALS (Dryer 2013f) follow the linguistic classification provided by Ethnologue (Eberhard, Simons & Fennig 2024), Grambank (Skirgård, Haynie, Blasi, et al. 2023, Skirgård, Haynie, Hammarström, et al. 2023) follows the one provided by Glottolog (Hammarström et al. 2024), and so does the sample presented here. Section 4.2.2 will show where and how the sorting of the languages in appendix A differs from Cinque (2021) and Dryer (2018b) with respect to U20, while section 4.2.3 will present a similar comparison with samples that focus on the unmarked order of mono-transitive clauses.

4.2.1 Glottolog over Ethnologue

There are multiple reasons for preferring Glottolog's (Hammarström et al. 2024) classification over Ethnologue's (Eberhard, Simons & Fennig 2024). First, Glottolog provides each linguistic variety with a unique identifier, a *glottocode* made of four letters followed by four numbers (Forkel & Hammarström 2022). Given that a certain language can be addressed in the literature by means of different spellings and different names (Körtvélyessy 2017: 128),[95]

[95] A good example in this sense is the Mayan language Tojolabal (tojo1241), which Cinque (2021) lists twice (and under two different ways of linearizing U20) because the sources he consulted spell the language name in different ways, namely Tojolabal and Tojol-ab'al. This language is not part of the present sample because it does not satisfy the selectional criteria established in section 4.1.2.

or (vice versa) that a specific name can be used to address several unrelated languages,[96] a coded identifier is extremely useful for typological purposes. It is true that WALS (Dryer & Haspelmath 2013) and Ethnologue (Eberhard, Simons & Fennig 2024) propose solutions going in this direction, ISO-369-3 being the most useful and widely adopted (see Forkel & Hammarström 2022: 918 for discussion). However, there is a one-to-many mapping between an ISO-369-3 and different languages, and some languages are still not coded in this way (an example is the Afro-Asiatic language Malgwa, malg1251). These facts make the ISO codification prone to the problems just discussed for bare language names, as opposed to unambiguous glottocodes, which have the added advantage of identifying as much a dialect as a language family, thus being an "ID type oblivious to level of linguistic abstraction" (Forkel & Hammarström 2022: 918).

A second reason for preferring Glottolog (Hammarström et al. 2024) is that it provides the user with the relevant literature that has been used for identifying language families (and isolates). A direct consequence of this fact is that Glottolog's choices tend to be less speculative, so that the Niger-Congo hypothesis is rejected (in favor of a split among Atlantic-Congo languages and other minor families) as well as the Nilo-Saharan one. More generally, many languages who are supposed to be cognates because of some borrowings or weak lexical similarities are kept apart under Glottolog's classification.

What is more, it is easier to check and operate with Glottolog's family trees than with Ethnologue's. By means of Glottolog is possible to recognize when to conflate, split or rename genera provided by Cinque (2021). For example, Cinque identifies a genus of Semitic languages and a genus of Ethio-semitic ones. However, Glottolog clearly illustrates that the handful of Ethio-semitic languages is nothing but an immediate branching daughter of the former, which therefore should not be treated simultaneously as separate genera.[97]

It may look like the classification provided by Glottolog (Hammarström et al. 2024) produces more distinctions than Ethnologue (Eberhard, Simons & Fennig 2024) does. Although this may be somewhat true for the family level, there is room for adaptations at the genus level, so as to make genera comparable both within a single linguistic family and across samples. As a general rule, the genera appearing in appendix A are decided by selecting groupings

[96] Aside from different varieties addressed with the same name, such as Mixtec and Zapotec (which both identify subfamilies respectively coded as mixt1427 and zapo1437 from the Otomanguean family, otom1299), a good instance provided by the present sample is Kom, which is a name referring to both a Sino-Tibetan language (komi1270) and an Atlantic-Congo one (komc1235).

[97] Moreover, one could better name the Semitic languages sampled here as West Semitic, in order to keep Akkadian, Eblaite and other Semitic isolates separate.

that were at the same depth in the tree of a certain linguistic family. This implies that, in some cases, languages that Cinque (2021), Dryer (2018b) and WALS (Dryer & Haspelmath 2013) treat as separate genera have been clustered together.[98]

However, there are some exceptions to this criterion, which have been applied due to a twofold reason. First, a further subdivision of a certain group might be required by the vast area or the high number of languages that it contains. An example in this sense is represented by the Malayo-Polynesian subfamily. It is an immediate daughter of the Austronesian family, and it is in a sisterhood relation with the other Austronesian subfamilies that contain a handful of languages each. One may thus infer that this depth is the best that one could choose according to the criterion that is being discussed. However, the Malayo-Polynesian group is exceptional with respect to its sisters because it contains twelve hundred languages spreading over three continents. Therefore, splitting the Malayo-Polynesian cluster among its subfamilies is a move that accounts for its richness and variety. The second reason for overriding the sisterhood rule is that it subsequently becomes possible to define genera that largely coincide with the ones used by Cinque (2021), Dryer (2018b) and WALS (Dryer & Haspelmath 2013) as long as it is made reasonable by the breadth/richness criterion just illustrated.[99] In fact, in most cases the classifications provided by Glottolog and Ethnologue overlap, although family and genera might be addressed by different names.[100] This also means that in general, the different language classification adopted here might be concretely affected only by fringe cases, which are more likely to appear as linguistic isolates in Glottolog's classification than in Ethnologue's.

98 An example in this sense is the Balto-Slavic group, which Cinque (2021), Dryer (2018b) and WALS (Dryer & Haspelmath 2013) further split between the Slavic and the Baltic languages. This division is however troublesome, not only because it would put the Balto-Slavic subfamily on a par with the Indo-Iranian one (which requires further subdivisions due to the richness and the breadth of languages it contains), but also because the compactness that these sources assume for the Germanic group would be left unexplained.

99 One particular case in this sense is represented by creoles and pidgins, which are treated as a separate genus by all sources, while their individual families still resonate with the ones of their lexicalizing languages. My sample follows this trend by clustering them at the genus level while keeping them separate at the family level.

100 In fact, the Atlantic-Congo family defined by Glottolog (Hammarström et al. 2024) contains most languages that belong to the Niger-Congo family assumed by Ethnologue (Eberhard, Simons & Fennig 2024), so that the distinction affects the sample only in few cases.

4.2.2 U20 in Cinque's sample, Dryer's, and in Appendix A

The present sample gathers 1,141 languages from 325 genera, distributed among 171 families in 6 macro areas.[101] 701 languages come exclusively from the sample in Cinque (2021, 2023), while the remaining 440 languages are part of the one from Dryer (2018a,b). This means that, on average, more than half of the languages collected by Cinque (which are more than 2,100) are gathered in appendix A, as well as four-fifths compared to Dryer's sample. However, the fewer number of languages picked up does not imply an equally lower typological variability. In fact, the comparison with each possible linearization of U20 in Cinque's sample shows that the proportion of genera that are collected in appendix A is always higher than the respective proportion of languages. This can be seen in table 4.1, which compares the number of languages and genera listed in appendix A.1 with the ones reported by Cinque (2023: 35-36).

A confrontation for each linearization of U20 of the ratio (in percentages) of languages with the respective ratio of genera shows that the former is always higher than the latter. As was argued in section 4.2.1, the higher proportion of genera in the sample presented here cannot be exclusively due to the different classification that has been adopted. Instead, it indicates that the fewer number of languages investigated here is often caused by a lower availability or records on both Grambank and WALS for languages belonging to the same genus, either because of excessive typological similarity/areal proximity between such varieties or because of a general lack of bibliography for the linguistic genus at hand.[102] By adopting the convention in Dryer (2018a,b), U20 linearizations in table 4.1 are shortened by indicating nominal modifiers with capital letters (so that, contrary to what is done in Georgi & Müller 2010, uppercase *N* indicates cardinal numerals, while lowercase *n* indicates nouns).

Linearizations of U20 reported in table 4.1 are listed in decreasing order according to the number of languages reported by Cinque (2023: 35-56). As one can see, absolute values in other columns may end up consequently unordered at some point. In this sense, a comparison among different samples (i.e. the present one together with Cinque 2021, 2023, Dryer 2018a,b) can be made also with respect to the absolute frequency of each order of

[101] The six macro areas are Africa, Australia, Eurasia, North America, South America, Papunesia (a portmanteau of *Papua New Guinea* and *Austronesia*). To my knowledge, there is no disagreement between Glottolog (Hammarström et al. 2024) and Ethnologue (Eberhard, Simons & Fennig 2024) in the definition of these areas.

[102] The bibliographic bias is a serious issue that is discussed, among other works, in Rijkhoff et al. (1993) and Guzmán Naranjo & Becker (2022: Sec. 3).

	Languages per sample			Genera per sample		
U20 orders	**Cinque**	**App. A**	**Ratio**	**Cinque**	**App. A**	**Ratio**
nAND	606	283	47%	137	111	81%
DNAn	425	224	53%	110	87	79%
NnAD	226	155	69%	47	35	74%
DnAN	207	121	58%	79	61	77%
DNnA	182	77	42%	67	47	70%
nADN	101	64	63%	34	27	79%
nDNA	81	49	60%	22	15	68%
nNAD	67	34	51%	35	20	57%
NAnD	51	30	59%	22	14	64%
DnNA	47	20	43%	29	15	52%
DAnN	46	39	85%	24	23	96%
AnND	31	19	61%	13	11	85%
nDAN	30	14	47%	18	8	44%
AnDN	20	12	60%	8	8	100%

Table 4.1: Comparison of Cinque's sample and appendix A

U20. First of all, the three samples reach a consensus for the two most frequent orders, i.e. N-A-Num-Dem followed by Dem-Num-A-N, and the fourteenth most frequent one, namely A-N-Dem-Num, both at the language and the genus levels. Similarly, the three samples consistently agree on the third, the fourth and the fifth most frequent orders by number of languages, namely Num-N-A-Dem, Dem-N-A-Num, and Dem-Num-N-A, respectively. However, across the three samples the ranking of Num-N-A-Dem is consistently lowered to the fifth position as soon as genera are counted instead of languages. The latter order then concretely represents a possible case of oversampling, according to which the varying number of languages contained in a single genus may induce an over- or an under-representation of a certain typological factor. However, the way of dealing with this issue is still a matter of debate (see Rijkhoff et al. 1993 and Guzmán Naranjo & Becker 2022).

N-A-Dem-Num is the last order of U20 about which all the samples considered here (i.e. appendix A together with Cinque 2021, 2023, Dryer 2018a,b) largely agree on. It is the sixth most common order per number of languages, and the situation does not vary in both the present sample and Dryer's by adding up genera. From this point on (and up to the thirteenth most common order), the counting done by the three samples largely diverges on any basis. As it is acknowledged by Dryer (2018b: 99 ff.), this phenomenon is yet another example of the fact that, as soon as the numbers become smaller, the criteria adopted and the biases discussed by Rijkhoff et al. (1993) and Guzmán Naranjo & Becker (2022: Sec. 3) affect each sampling, so that a definitive consensus on crosslinguistic frequency of certain typological factors becomes extremely difficult.

4.2.3 Different samples w.r.t. the linear order in the clause

Before moving to the observations that can be made with respect to the sample presented here, it is worth confronting the overall data that have been collected in appendix A with other samples that are also concerned with the basic word order in mono-transitive clauses. First of all, table 4.2 on the next page shows the frequency of each type defined in section 4.1.4 among the languages collected in appendix A. Both absolute values and percentages are reported for languages, genera, and families. Values related to the six macro areas are not shown simply because each alternative is attested among all of them. The row of *sums* adds up the values from the rows above it, while the row of *real totals* reports the actual counting of languages, genera and families within the sample prior to the typological subdivision discussed here.

As anticipated in section 4.1.2, it is not by chance that only the number of languages matches (both in absolute values and percentages) across the two rows at the bottom. In fact, while there are no repetitions of a single language across the sample, it may well be the case that a single genus or family occurs across multiple types of clausal word order or among various linearizations of U20. The latter fact explains why the yield of rough sums is always higher than the overall real values in all columns of table 4.2, but those whose header is either *#Lgs* or *Lgs%*.

One may wonder whether type 1 and type 1.5 are found across the same subsets of genera and families, or whether a similar point holds for types 2.5 and 3. If this were the case, the six-way typology proposed so far could be reduced to a three-way counterpart that is more consistent with what Hengeveld, Rijkhoff & Siewierska (2004: 542) propose. However,

	#Lgs	#Gen	#Fam	Lgs%	Gen%	Fam%
Type 0	11	11	11	1%	3%	6%
Type 1	131	48	25	11.5%	15%	15%
Type 1.5	27	20	11	2.4%	6%	6%
Type 2	464	122	46	40.7%	38%	27%
Type 2.5	23	20	18	2%	6%	11%
Type 3	485	166	113	42.5%	51%	66%
Sums	1141	386	223	100%	119%	130%
Real totals	1141	325	171	100%	100%	100%

Table 4.2: Frequencies in the languages reported in appendix A, sorted by type

neither of the two possibilities holds. On the one hand, a merger of types 1 and 1.5 gathers 158 languages (14% of the total) from 58 genera and 29 families, which represent the 18% and the 17% of the whole sample, respectively. On the other hand, types 2.5 and 3 are attested in 508 languages (45% of the total) from 179 genera and 121 families, representing 55% and 71% of the total, respectively. As discussed in sections 4.1.3 and 4.1.4, type 1 and type 3 do not co-occur in a language without type 2 (and if all types are attested in one, then the latter is classified as an instance of type 0). Hence, one can safely sum the values just presented for types 1/1.5 and 2.5/3 with the number of genera and families of types 0 and 2. By doing this, the rough sum of genera falls down from 119% to 114%, and families do so from 130% to 121%. The reduced values indicate that in some cases the two pairs of types do co-occur in the same genera and families. However, the obtained values still exceed the 100% threshold, which means that there exist genera and families where type 1 is attested while type 1.5 is not, or vice versa; ditto for types 2.5 and 3.[103]

The sample in appendix A can also be compared with others that are concerned with the basic order in mono-transitive clauses. This is done in table 4.3, where the related distributions in the present sample are contrasted with those reported in other four (see Cinque 2023: 138 fn. 7, Arikawa 2013: 278 for a look at even more samples). Given that the works reported in table 4.3 divide the languages they collect based on a Greenbergian typology of

[103] This result holds even if one assumes a different overlapping for types 1.5 and 2.5, namely by grouping one of them or both together with type 2 languages. Since these counterproofs go along the lines that have just been discussed, they will not be illustrated here.

word order, their results need to be adapted by grouping together VSO and VOS orders under type 1, SVO and OVS under type 2, and SOV and OSV under type 3. Languages that are not classified for various reasons are thus grouped together under type 0 (except where otherwise indicated, as in the case of Dryer 2013f: Sec. 3).

	0	**1**	**1.5**	**2**	**2.5**	**3**
Ruhlen (1975: 151)	0.0%	10.8%	0.0%	37.4%	0.0%	51.7%
Mallinson & Blake (1981: 148)	11.0%	11.0%	0.0%	36.0%	0.0%	42.0%
Dryer (2013f) (WALS)	8.9%	9.7%	1.5%	36.3%	2.1%	41.5%
Hammarström (2015: 10)	2.4%	12.9%	0.0%	41.0%	0.0%	43.7%
This sample (appendix A)	1.0%	11.5%	2.4%	40.7%	2.0%	42.5%

Table 4.3: The frequency of different types across different typological samples

As far as V-initial and V-final languages are concerned, table 4.3 shows that their distribution in appendix A.1 does not diverge from what can be observed in other similar works. In agreement with Hammarström (2015: 10), the present sample classifies (proportionally) more V-medial languages than other authors do, but considerably fewer languages with no basic word order. Languages that in other works would be classified as having "no dominant order" (here labeled as "type 0") are interspersed among types 0, 1.5 and 2.5 within the sample presented here, as a consequence of the multicategorization (Ivani & Bickel 2023) provided by Grambank (Skirgård, Haynie, Blasi, et al. 2023, Skirgård, Haynie, Hammarström, et al. 2023).

4.3 An analysis of the data

Starting from the inquiry pursued by Roberts (2017b), where two linearizations of Greenberg's U20 correlate positively with V-initial and negatively with V-final word orders in the clause, the present work expands the investigation to see if similar trends can be observed for other options of U20. For a start, section 4.3.1 will verify whether the observations discussed for Robert's sample (section 3.2.3) are confirmed by the data collected in appendix A. Sections 4.3.2 and 4.3.3 will instead be concerned with the analysis of the other attested permutations of U20, while section 4.3.4 will further discuss the observed results, which will be

summarized in section 4.4.

4.3.1 The cluster in Roberts (2017b), extended

Among Num-N-A-Dem and Num-A-N-Dem, which are the only two attested orders of U20 with both prenominal cardinal numerals and postnominal demonstratives, only two languages with V-final orders in unmarked transitive clauses are found. The first is the Ethiopian language Northern Gumuz (gumu1244), which is of type 2.5, whereas the second is standard Basque (basq1248), which is of type 3 (see again section 3.1.2 for further discussion).[104] Given that V-final languages are the most common crosslinguistically, this is perhaps the most striking fact related to these two linearizations of U20. Table 4.4 summarizes the results related to these two permutations. Percentages are referred to the whole sample, and only real totals are shown (because rough sums would encounter the same issues discussed for table 4.2 on page 177). Furthermore, types 1 and 1.5 are grouped together; ditto for types 2.5 and 3.

	#Lgs	#Gen	#Fam	Lgs%	Gen%	Fam%
Type 0	1	1	1	0.1%	0.3%	0.6%
Types 1 AND 1.5	88	24	10	7.7%	7.4%	5.8%
Type 2	94	27	10	8.2%	8.3%	5.8%
Types 2.5 AND 3	2	2	2	0.2%	0.6%	1.2%
Real totals	185	45	17	16.2%	13.8%	9.9%

Table 4.4: The frequencies in appendices A.4 and A.11 for the two orders in Roberts (2017b).

By relying on their overall average frequency worldwide, one can expect to find around 10-15% of V-initial languages for each linearization of U20 (see table 4.3 on the facing page). However, Num-N-A-Dem and Num-A-N-Dem disregard this expectation. Indeed, they seem to be already sufficient to capture more than half of the overall languages (88 out of 158; 7.8% out of 13.8%) of type 1 or 1.5. However, this estimate may still be affected by an oversampling of languages belonging to the Otomanguean and Austronesian families. In fact,

[104] There were three other possible exceptions (i.e. potentially both SOV and Num-N-Dem) which were excluded because the sources considered were not agreeing on the classification of at least one of their domains: Isirawa (isir1237), Wajarri (waja1257), and Akkadian (akka1240).

the rest of the sample contains 33 more genera for types 1 and 1.5 (namely, three-fifths of the total), and 18 other families (almost two-thirds of the total). Put differently, the remaining 69 languages of types 1 and 1.5 are more varied from a phylogenetic point of view.[105]

Before moving on to the rest of the sample, it should be noted that the languages in table 4.4 are not evenly divided between Num-N-A-Dem and Num-A-N-Dem. In fact, the latter is represented by only 30 languages (namely less than one-sixth of the total in table 4.4) distributed among 14 genera and 10 families (see also appendix A.11). What is more, the two permutations of U20 show opposite trends. In Num-A-N-Dem (i.e. the smaller cluster) V-initial languages, genera, and families slightly prevail over the V-medial ones, while in Num-N-A-Dem the opposite holds.

4.3.2 Four orders with peculiar properties

Dryer (2013a: 275) claims that the linear order of numeral and noun is statistically significant with respect to what are here called type 1 and type 3 languages. In particular, "the numeral **follows** the noun more often in verb-final languages than it does in verb-initial languages" (boldface in the original). Although this observation fits well with the patterns that have just been discussed, it is possible to find a relevant number of exceptions to the correlations between Num-N and type 1 on the one hand and N-Num and type 3 on the other. The sample discussed here can shed some light in this respect by virtue of the fact that, contrary to what is done on WALS (Dryer & Haspelmath 2013), it does not split U20 into pairs.

The two best candidates that could confirm the tendency just sketched are the two most common linear orders of U20, namely N-A-Num-Dem and Dem-Num-A-N. *Contra* the generalizations discussed in section 2.2 and in section 3.2.2, Dryer (2013a: 275) expects the former to occur among many V-final languages, while the latter among many V-initial ones. However, these predictions turn out to be rather weak. First, they are clearly disproved by the 224 languages featuring Dem-Num-A-N. Of those, only 20 languages (from 14 genera) are V-initial (i.e., of type 1 or 1.5), while 162 (from 55 genera) are V-final (i.e., of type 2.5 or 3). This also means that V-final languages, which are predicted to be unlikely, constitute nonetheless more than 70% of the languages in this cluster (and over 50% of genera and families therein). At the same time, it is true that V-final languages are likely to co-occur

[105] This may be linked to another property that the two linear orders of U20 in section 4.3.1 share, i.e. both cannot be found in Australia.

with N-A-Num-Dem, but in opposition to what could be observed for Dem-Num-A-N, they are outnumbered by V-medial languages: 110 languages of types 2.5 and 3 from 51 genera against 162 languages of type 2 from 59 genera.

The properties illustrated for the two most frequent orders of U20 outline two general patterns that hold throughout the sample in appendix A. The first is that, in general, the attested orders of U20 that feature a prenominal demonstrative co-occur more frequently with V-final orders.[106] This specific trend is not new in the literature, since it is also observed by Dryer (2019: 87), but the readers will shortly see that this sample allows to find further relevant insights in this respect.

The second pattern that holds throughout appendix A can be formulated only as soon as one looks at mono-transitive clauses and nominal modifiers of U20 in their entirety, and it consists of the fact that the V-medial word order (type 2) most commonly occurs among the attested linearizations where both the demonstrative and the numeral are postnominal (see section 4.3.3).[107]

Although they *prima facie* comply with the generalizations just proposed, two linearizations of U20 require in-depth discussion, i.e. N-Dem-Num-A and Dem-Num-N-A. The former seems to obey the generalization concerning it: V-medial languages are way more frequent than V-initial and V-final ones. However, this first observation gets blurred at the levels of genera and families, as table 4.5 shows. At a closer look, the higher number of type 2 languages in table 4.5 is in fact due to an overrepresentation of the Bantoid genus (from the Atlantic-Congo family). As soon as one is concerned with genera and families, V-initial languages appear to be more frequent than V-final ones. Given the crosslinguistic low frequency of the former (see again table 4.3 on page 178), this fact is as striking as it is for the orders presented in table 4.4 as an expansion of Roberts (2017b).

Two further interesting features of the data in table 4.5 can be noted. First, no languages with multiple word orders in the clause occur, i.e. types 0, 1.5, and 2.5 are not attested. Second, rough sums of genera and families match the real values.[108] These facts combined may indicate that languages which linearize the nominal domain in this way are not shifting

[106] Such orders (by shortening them *à la* Dryer 2018a,b, where capital letters are used for modifiers) are: DNAn, DnAN, DnNA, DAnN, and DNnA.

[107] Orders of this kind (by shortening them as in footnote 106) are: nAND, nADN, nNAD, nDAN, AnND, AnDN, and nDNA.

[108] As the mismatch between sums and real values indicates, the situation for the macro areas is different. All types can be found in Africa, while one language of type 1 and one of type 3 were found in Australia. The sample provides no Dem-Num-N-A language for either South America or North America.

	#Lgs	#Genera	#Families	#Areas
Type 0	0	0	0	0
Type 1	7	6	5	2
Type 1.5	0	0	0	0
Type 2	33	5	5	2
Type 2.5	0	0	0	0
Type 3	9	4	4	3
Sums	49	15	14	7
Real totals	49	15	14	4

Table 4.5: The frequencies in appendix A.8 for the N-Dem-Num-A order

from one word order in the clause to another, and that they may have reached the observable situation through distinct historical developments.

Let us now turn to Dem-Num-N-A, which is predicted to be more likely to occur in V-final languages. As table 4.6 shows, this prediction holds at all levels of analysis, except the macro areas. However, this higher frequency is only slimly so. In comparison to the data discussed in table 4.5 for N-Dem-Num-A, data in table 4.6 are similarly distributed in an even way among types 1, 2, and 3, with a surprisingly high frequency of the former.

	#Lgs	#Genera	#Families	#Areas
Type 0	2	2	2	2
Type 1	18	15	14	6
Type 1.5	4	3	3	3
Type 2	25	18	12	5
Type 2.5	1	1	1	1
Type 3	27	19	18	3
Sums	77	58	50	20
Real totals	77	47	35	6

Table 4.6: The frequencies in appendix A.6 for the Dem-Num-N-A order

However, there are some remarkable differences between table 4.5 and table 4.6. In the latter, some genera and families coincide across types, a fact that can be inferred from the mismatch between sums and real totals. Most overlappings occur between type 1 and type

2, which are actually attested in 27 genera from 18 families (instead of 33 genera from 26 families). What is more, Dem-Num-N-A is also found among languages with multiple word orders (i.e., types 0, 1.5, and 2.5), contrary to what has been seen for table 4.5.

Table 4.7 clusters Dem-Num-N-A and N-Dem-Num-A together, by replicating the methodology adopted in table 4.4 on page 179 with respect to Num-N-A-Dem and Num-A-N-Dem. The percentages in table 4.7 further summarize what has been discussed so far and present the proportions with respect to the whole sample of languages.

	#Lgs	**#Gen**	**#Fam**	**Lgs%**	**Gen%**	**Fam%**
Type 0	2	2	2	0.2%	0.6%	1.2%
Types 1 AND 1.5	29	23	19	2.5%	7.1%	11.1%
Type 2	58	22	16	5.1%	6.8%	9.4%
Types 2.5 AND 3	37	24	22	3.2%	7.4%	12.9%
Real totals	126	61	46	11.0%	18.8%	26.9%

Table 4.7: Aggregated frequencies from table 4.5 and table 4.6

4.3.3 Two more groupings

The discussion can now move to the remaining orders that obey the observations proposed above in a more consistent way. For this reason, two further clusters can be made, so as to better analyze the positive correlations between Dem-N and V-final languages on the one hand and the distribution of clausal word orders in the remaining attested permutations of U20 on the other hand. The group outlined by the former consists of attested orders where the demonstrative appears prenominally while the numeral is postnominal, i.e. Dem-N-A-Num, Dem-A-N-Num, and Dem-N-Num-A, which have just been listed from the most common to the least common in the sample presented here. As appendix A.5 shows, Dem-N-A-Num co-occurs with only 2 V-initial languages and 14 V-medial ones (the latter belonging to 12 genera), which are outnumbered by 105 V-final languages from 53 genera. Dem-A-N-Num is then found in 39 languages that are listed in appendix A.9, 34 of which are of type 3 (the remaining 5 are interspersed among types 0, 1, 1.5, and 2). Dem-N-Num-A closes the list with 20 languages presented in appendix A.12, 11 of which are V-final. This slim superiority with respect to V-medial languages becomes more evident by looking at genera and families:

11 genera from 10 families for V-final languages as opposed to 3 genera from 3 families for the V-medial ones.

	#Lgs	#Gen	#Fam	Lgs%	Gen%	Fam%
Type 0	1	1	1	0.1%	0.3%	0.6%
Types 1 AND 1.5	5	5	4	0.4%	1.5%	2.3%
Type 2	24	15	9	2.1%	4.6%	5.3%
Types 2.5 AND 3	150	73	51	13.1%	22.5%	29.8%
Real totals	180	87	59	15.7%	26.8%	34.5%

Table 4.8: Aggregated frequencies for the three attested Dem ≻ N ≻ Num orders

Table 4.8 aggregates the values for the three orders just presented by following the criteria already applied to table 4.4 on page 179. As the readers can see, the higher proportion of V-final languages occurring in this cluster becomes even more visible once the four linear orders of U20 are grouped together. The values shown in table 4.8 sharply contrast with what can be observed by grouping together N-A-Num-Dem, N-A-Dem-Num, N-Num-A-Dem, N-Dem-A-Num, A-N-Num-Dem, and A-N-Dem-Num, whose aggregated values are illustrated in table 4.9. With the exception of N-Dem-Num-A (whose distribution has been reported in table 4.5), the latter six orders represent all the attested cases of U20 where both Dem and Num are postnominal. In a similar way to what happened with the values in table 4.8, the aggregated values in table 4.9 effectively summarize the general trends that can be described for any of the orders that constitute it. Although the number of languages is slightly higher in the latter group than in the former, the number of genera in table 4.9 is lower, and this is even more evident with families. Moreover, the six cases of U20 grouped there are more likely to be found in V-medial languages than in V-final ones, but the latter come from a higher number of families as opposed to the former (which in most cases is still more frequent across genera).[109]

If one goes past the frequencies that have already been discussed in section 4.3.2 for N-A-Num-Dem, N-A-Dem-Num is the linear order of U20 that is crosslinguistically more frequent among the six at hand. It occurs in 65 languages from 28 genera, and 33 of those languages

[109] A further observation to highlight, as per the recommendations in Dryer (2018a: 802), is that these orders are absent in one macro area, i.e. South America.

	#Lgs	#Gen	#Fam	Lgs%	Gen%	Fam%
Type 0	1	1	1	0.1%	0.3%	0.6%
Types 1 AND 1.5	16	10	4	1.4%	3.1%	2.3%
Type 2	252	70	23	22.1%	21.5%	13.5%
Types 2.5 AND 3	157	66	44	13.8%	20.3%	25.7%
Real totals	426	131	57	37.4%	40.3%	33.3%

Table 4.9: Aggregated frequencies for N > Dem, N > Num orders (except N-Dem-Num-A and *N-Num-Dem-A)

are V-medial (and occur in 13 genera), while 28 languages are V-final (and occur in 15 genera). The 4 remaining N-A-Dem-Num languages are of type 1 or 1.5, and they add up to 11 N-A-Num-Dem languages of the same type. This surprisingly low number of verb-initial languages co-occurring with N-A-Num-Dem also disconfirms the claim that such a permutation of U20 is typical of VOS/VSO languages (Cinque 2023: 42-43).[110]

Gude (gude1246) is the only language of type 1 that occurs among the other four rather rare orders of U20 included in table 4.9, i.e. N-Num-A-Dem, N-Dem-A-Num, A-N-Num-Dem, and A-N-Dem-Num.[111] Table 4.10 on the following page gathers the frequencies for the 79 languages belonging to these four orders that negatively correlate with V-initial orders, which is a peculiarity that they share with Dem-N-Num languages gathered in table 4.8. What is more, A-N-Dem-Num is the least frequent order of U20 in the sample presented in this chapter, and it is also the only one that occurs almost exclusively in V-medial languages (out of 12 languages from 8 genera, only one language of this kind is V-final, i.e. Puxi Qiang, sout2728). The almost complete absence of verb-initial languages among these four permutations of U20 is an aspect that will be discussed further in section 4.5.

4.3.4 Some further remarks

The data presented and discussed in the previous sections are not put to test with respect to their statistical significance. Although this might be a limit of the present work, it can

[110] In this respect, it is worth observing that one of the two examples that Cinque (2023: 41) mentions in support of the rigid head-initial structures that he posits is the Polynesian language East Uvean, which already represents a partially divergent case by featuring the N-A-Dem-Num order (see Cinque 2023: 126 en. 49 and appendix A.7).

[111] For further discussion on the classification of Gude, see footnote 244 in appendix A.14 on page 482.

	#Lgs	#Gen	#Fam	Lgs%	Gen%	Fam%
Type 0	0	0	0	0.0%	0.0%	0.0%
Types 1 AND 1.5	1	1	1	0.1%	0.3%	0.6%
Type 2	58	20	8	5.1%	6.1%	4.7%
Types 2.5 AND 3	20	15	13	1.8%	4.6%	7.6%
Real Totals	79	35	18	7%	11.0%	11.0%

Table 4.10: Aggregated frequencies for
N-Num-A-Dem, N-Dem-A-Num, A-N-Num-Dem, A-N-Dem-Num

be fixed in the future. Moreover, the validity of this kind of calculations is still a matter of debate. In fact, Piantadosi & Gibson (2014) argue that the number of unrelated languages needed to obtain reliable statistical results is far beyond the scope of any sample available from currently documented languages.

On the other hand, Guzmán Naranjo & Becker (2022) take a different view by underlying the importance of a carefully constructed sample for the purposes of a good statistical assessment. The discussion reported therein, however, shows that there is no general consensus on whether rich samples are better than selective ones, nor on whether crosslinguistic variability should be enhanced at the expense of crosslinguistic frequency (see Guzmán Naranjo & Becker 2022: Sec. 3). What is more, some of the works that use a statistical approach to linguistic typology are unlikely to give a blueprint for the results presented here, because they focus on the correlations between OV and VO that have been put aside here (e.g. Bickel 2007, 2015).

In any case, the literature offers some methodologies that one could apply in the future to the data presented in this chapter. As a first example, Maurits & Griffiths (2014) use a Bayesian approach to hypothesize the word order of mono-transitive clauses in the ancestors of the language families that are most spoken nowadays. This approach could be applied to investigate which orders of U20 were more likely to be used in the past, and the results coming from the two domains could then be confronted.

It is also true that certain linguistic families show so little internal variance that one could argue for a specific syntax of the ancestors even without some statistics, as Gell-Mann & Ruhlen (2011) do. These different degrees of variability are however lineage-specific, and

can also be observed in this sample with respect to U20: the members of certain families highly differentiate their way of linearizing nominal modifiers (e.g. Austroasiatic, Indo-European, and Sino-Tibetan), while other families show a rather consistent internal behavior (e.g. the Dravidian consistently show the Dem-Num-A-N order).[112] These lineage-specific trends can thus be investigated along the lines illustrated in Shcherbakova et al. (2023), who focus on the richness of verbal and nominal morphology and come to the conclusion that the evolution of such phenomena follows different patterns across families, although being consistent among related languages.

Statistics aside, the data discussed in this chapter are sufficient to put to question many claims made before by the relevant literature. Cinque (2023: 70-71) argues that modifiers are less compliant than syntactic complements when it comes to pattern with their selecting head in a harmonic way. Given that the absence of such patterning is treated as an exception that needs some other explanation, this approach may at least partially be off track, because in this chapter it was nonetheless possible to argue for some clear trends, albeit not the expected ones.

On a theoretical point of view, the order of modifiers within the nominal domain turns out to be somewhat related to the order of arguments in the clausal domain, which in turn suggests that Manning (2003: Sec. 3.1), McInnerney (2022) and other works mentioned on page 122 (in section 3.1.4) might be right in arguing for a relaxation of the syntactic distinction between arguments and modifiers (understood as adjuncts), as long as word order is concerned (see instead Kosta & Krivochen 2012: Sec. 3.1 for theoretical reasons in favor of the split).

Moreover, the data commented in the previous sections strongly invoke an overhaul with respect to the head-initial or and head-final patterns that have been taken for granted up to now in the literature. The most recent example of this seemingly off track trend is Cinque (2023: 40, 63-64), who posits consistently head-initial or head-final structures that are expected to be the least marked for any language and thus highly likely crosslinguistically. However, this is not the case: Cinque (2023: 68) admits that even "the more rigid head-initial VOS languages and the more rigid head-final SOV languages […] display some inconsistencies, and, more

[112] Of course, language numerosity within families can affect these observations, as Dryer (2018a: 802-803) points out. Note however that similar differences can also be seen across comparable genera, with Nakh-Daghestanian and Germanic languages being consistently Dem-Num-A-N as opposed to the highly variable Core Chibchan and Kuki-Chin-Naga languages. Cinque (Forthcoming: 1 fn. 2) makes a similar observation, and uses it to argue in favor of convenience samples as opposed to other sampling methods: "It is in any event important to try and document as many languages as possible for each genus, as it turns out that languages belonging to the same genus often show different orders".

strikingly, they appear to be a minority of the languages of the world". Under his view based on the expectation of harmonic patterns determined by binary-branching structures, these are "two puzzles that remain to be understood" Cinque (2023: 68).

As far as the accounts of U20 provided by Cinque (2005, 2023) are concerned, it is worth observing that the aforementioned correlations between the nominal and the clausal domains cannot be captured by means of the three types of NP-movements proposed there either. This fact can be easily acknowledged by means of three attested permutations of U20, i.e. Dem-N-A-Num, Dem-N-Num-A and Dem-A-N-Num. Although table 4.8 on page 184 shows that there are good reasons to cluster them together, Cinque (2005, 2023) derives them in the three distinct manners that he posits. Dem-N-Num-A follows from NP-raising, while Dem-N-A-Num and Dem-A-N-Num are derived from pied-pipings of the *whose picture* and *picture of whom* type, respectively.

The alternate structural approach to Greenberg's U20 that Abels & Neeleman (2009, 2012) give (section 3.2.4) does not fare better in capturing the observed tendencies either. The major distinction that they propose is between eight base-generated structures and six counterparts that are derived via leftward movement. Both subsets group together linearizations that have been shown to correlate in quite different ways, or vice versa split permutations of U20 that should be clustered together. This can again be exemplified by means of Dem-N-Num-A, which is derived by leftward N-raising, as opposed to Dem-N-A-Num and Dem-A-N-Num, which are base-generated. These three orders also show that the branching directionality of the modifiers would not help solve the matter. On the one hand, Num must be a postnominal modifier in the base-generated structures for Dem-N-A-Num and Dem-A-N-Num, whereas it is has to be prenominal in the one where N raises to derive Dem-N-Num-A.

With respect to the analysis conducted in the previous sections, another general observation that can be made is that moving beyond the widespread pairwise analysis of typological phenomena has truly been beneficial. In fact, most of the observed correlations between the clausal and the nominal domains could only be formulated by considering demonstratives, cardinal numerals, and nouns at the same time. However, two objections may be formulated in this respect. The first is that one of the observations made relies on a pairwise relation that has already been illustrated by Dryer (2019: 87), namely the positive correlation between prenominal demonstratives and V-final orders. The second is that adjectives are not particularly helpful in providing insightful correlations, which is a matter already observed by Dryer

(2013b).

Both objections can be dismissed by focusing on two other aspects that have been remarked, namely the correlation between V-initial languages and Num-N-A-Dem and Num-A-N-Dem on the one hand, and the exceptional patterns shown by Dem-Num-A-N, N-Dem-Num-A, and Dem-Num-N-A on the other hand. In fact, the correlation between prenominal demonstratives and V-final languages is nothing but a superset which also includes the three orders where the demonstrative and the numeral are respectively prenominal and postnominal, i.e. Dem-N-Num-A, Dem-A-N-Num and Dem-N-A-Num. Hence, one could state that V-final languages better correlate with the latter three orders of U20, thus gaining two advantages. The first is that it becomes thus possible to claim that the latter correlation is nothing but the mirror-image of the relation that holds for the cluster first gathered by Roberts (2017b). The second advantage is that the exceptional behavior of Dem-Num-A-N and Dem-Num-N-A follows from the prenominal nature of Num. N-Dem-Num-A can instead be related with the latter two orders by intuitively observing that there is a precedence relation $Dem \succ Num \succ A$ that unites them. However, this intuition is shaky and must be reformulated on different grounds, not only because it excludes Dem-N-Num-A for no reason, but also due to additional factors that will be discussed in section 4.5.1.

4.4 Interim conclusions

The data commented in this chapter stem from the cross-referencing of publicly available sources, i.e. Dryer (2013f) and Dryer (2013f), Skirgård, Haynie, Blasi, et al. (2023), and Skirgård, Haynie, Hammarström, et al. (2023) for the word order in the clause and Cinque (2021, 2023) and Dryer (2018a,b) for the linear order of nominal modifiers. These sources also demonstrate that numerous scholars have tackled Greenberg's Universal 20 and/or the order of the arguments in mono-transitive clauses, and the explanation of the patterns found is an ongoing matter of debate. The research of correlations or parallels across the two domains is thus not a brand-new idea. The innovation of this chapter rather lies in the scope of the presented data, which (except for Bernardi 2025) is unprecedented in the literature.

Such a sample (fully reported in appendix A) aimed at questioning a widespread methodological blueprint discussed since chapter 2, namely the assumption of certain linguistic elements being heads selecting their dependents in a pairwise fashion, whose relative order

is expected to be a predictor of the order of other pairings. A different treatment is thus pursued by trying to adopt an approach to both the clausal and the nominal domains that is somewhat holistic and pre-theoretic, in order to see whether a look at wider patterns could be more promising or insightful than the more common decomposition into smaller, binary subparts. This was done by dividing the languages of the sample among the fourteen largely acknowledged ways of linearizing the elements of U20 (see again section 3.2).

The groups obtained through this method were then confronted with a revised typology of the clausal domain which focuses on the pragmatically unmarked position of the mono-transitive verb without worrying about the structural or semantic distinction between the two lexically expressed arguments it selects (see again section 4.1). The analysis of the data came to four main observations, which are *prima facie* surprisingly unrelated from one another.

The first observation is that V-medial patterns are attested in pragmatically unmarked mono-transitive clauses as soon as both V-initial and V-final patterns are equally valid alternatives. Section 4.1.4 showed that this holds true not only at the intra-linguistic level, but also crosslinguistically as long as languages are sorted according to the way they linearize the elements of Greenberg's U20. Second, table 4.9 on page 185 illustrated that any linearization where both demonstratives and cardinal numerals are postnominal is more likely to pattern with V-medial languages and is unlikely to occur in V-initial ones (however, the latter option is not completely ruled out). Third, attested orders where Num precedes N and Dem follows it are likely to occur in V-initial languages and are almost never found in V-final ones (section 4.3.1).

The last observation stemming from the data discussed in this chapter is that prenominal demonstratives positively correlate with V-final languages, as already observed by Dryer (2019: 87). However, this generalization can be formulated in a better way by splitting such cases into two subgroups. On the one hand, the correlation between the attested $Dem \succ N \succ Num$ orderings and V-final languages could be contrasted in a more pronounced way with the correlation between the attested $Num \succ N \succ Dem$ options and V-initial patterns. Although they are generally consistent with the trends outlined above, $Dem \succ Num \succ A$ orderings require on the other hand some further analysis of their own. This point represents a first hint at the fact that the observations just summarized cannot be adequately captured without appealing to structural considerations (as argued by Wiltschko 2014: 39-41 and references therein). This issue will be tackled in section 4.5.1.

To conclude this section, one can remark that the notions of *prefield* and *postfield* proposed by Rijkhoff (2002: 251-252) might be useful so as to capture the last three observations just reported (as will be argued further on page 372, in footnote 191). These two terms indicate "a position in the area before or after the head in a domain, respectively", so that Num and Dem correlate with V-medial orders if they are both in the postfield. If one of the two is instead in the postfiled and the other in the prefied, they correlate with V-initial or V-final languages depending on whether it is the numeral or the demonstrative that goes in the prefield, respectively. Alternatively, these two trends can be both addressed by means of a third term, namely *counterfield*, which Rijkhoff (2002: 252) uses to refer to "the Postfield in a Prefield or head-final language and [to] the Prefield in a Postfield or head-initial language". Thus, it can be said that both V-initial and V-final languages positively correlate with the numeral being placed in the counterfield, provided that the demonstrative is not occurring there as well.

4.5 On the need of a different theoretical approach

The analysis of U20 provided by Roberts (2017b), which has already been presented and discussed in section 3.2.3, represents the major inspiration for the typological data presented so far in this chapter and collected in appendix A. Moreover, section 4.3.4 has shown that the analyses of U20 given by Cinque (2005, 2023) and Abels & Neeleman (2009, 2012) would struggle at capturing the observed correlations between the nominal and the clausal domains.

In section 4.5.1 it will be argued that the data presented in this chapter need an analysis formulated on structural terms. Section 4.5.2 will then try to reformulate the premises laid out by Roberts (2017b) so as to effectively capture the observed connections between the verbal and the nominal extended projections. A formal parameterization will be proposed and the subsequent pros and cons of such a formulation will be assessed. Even though the observed correlations suggest that an attempt along these lines is well motivated, section 4.5.3 will show that a brief look at a specific subset of languages analyzed by Tvica (2017) is sufficient to demonstrate that connecting a certain binary parameter with a specific position in the structure might be too deterministic. Section 4.5.4 will show that Frank & Hunter (2021) provide a theoretical reasoning that allows to have it both ways, once the determinism connected to a derivational and successive-cyclic approach to structure building is replaced

by the formalism that will be discussed in chapter 5, where countercyclic operations are allowed in constrained and well-defined manners.

4.5.1 Beyond a linear statement

The sampling collected by Roberts (2017b) is gathered to test the assumption that the Num-N-Dem attested orders are not the result of pied-piping of [Num [N]] to a specifier c-commanding a demonstrative base-generated in Spec,D, because Roberts (2017b) further argues that this would lead to a FOFC violation (see again section 3.2.3). The focus on languages whose nominal domains feature prenominal numerals and postnominal demonstratives makes Roberts (2017b) observe that articles and demonstratives frequently co-occurr discontinuously and that the V-initial order in the clause appears in half of the sampled languages. Roberts then tries to capture these facts by assuming that both the verbal and the nominal extended projections contain two phasal heads, i.e. C and *v* in the former and D and *n* in the latter, so that the higher phase could attract the external argument of the lower phase, and that this syntactic relation is in turn subject to parameterization.

The Specifier-Head-Complement-Hypothesis adopted following Kayne (1994) and the assumptions just summarized make Roberts (2017b) neglect the more striking absence of OV languages (with very few exceptions) from the aforementioned sample, while the data discussed section 4.3 hint at a ban on VS holding within another subset of orders of U20. Given that the latter includes those languages whose nominal domains bear prenominal demonstratives and postnominal numerals, one first legitimate hypothesis would be to say that OV and VS orders in mono-transitive declarative clauses respectively show a negative correlation with two sets of nominal domains that seemingly mirror each other on linear grounds, i.e. Num-N-Dem as opposed to Dem-N-Num.

However, the latter formulation appears to be insufficient and needs to be restated on structural grounds, for three reasons. First, adding adjectives to the picture shows that the linear mirroring captured by the two correlations is only apparent, given that the Dem-N-Num-A order has no attested A-Num-N-Dem counterpart. Second, if the two correlations must be formulated in linear terms, this should also apply in the clausal domain. Thus, one would expect either a contrast between SOV and VOS that cannot be extended to head-finality and head-initiality more generally (a hypothetical scenario that would be *contra* Cinque 2023), or a split along the SV/VS or OV/VO dichotomies already discussed in section 2.2.2. However,

none of these options adequately captures the facts, mainly because SVO languages can be found in each attested permutation of U20 (see again section 4.1.3).

The third factor that hints at a structural analysis of the aforementioned correlations is that four among the rarest orders of U20, i.e. N-Num-A-Dem, N-Dem-A-Num, A-N-Num-Dem, and A-N-Dem-Num, almost never appear in VS languages, as it was illustrated in table 4.10 on page 186. As soon as the number of languages belonging to such a group (i.e. 79) is regarded as large enough to argue that the restriction on VS orders also applies there, then what determines this phenomenon can no longer be reduced to the presence of a prenominal demonstrative and a postnominal numeral within the nominal domain.

4.5.2 A parametric restatement

Given that Roberts (2017b) himself tries to account for the typological data he gathers by establishing parameters whose setting depends on structural properties of the analyzed languages, one first way to restate the observed correlations in non-linear terms boils down to pursuing his set of assumptions and reformulate the whole parameterization accordingly. The effectiveness of this process can then be assessed by evaluating the theoretical and empirical challenges that emerge.

As already discussed in section 3.2.3, Roberts (2017b) posits that Dem is first-merged as a specifier of a low phasal head in the nominal domain, i.e. *n*, and that the language-specific relation between Dem and the high phasal D head is determined based on a binary parameter, which can be addressed as A whenever Dem remains structurally low. By following these premises, Roberts (2017b: 180-183) shows that both Num-A-N-Dem and Num-N-A-Dem are derived if A holds while NP raises to the outer specifier of either *n*P or FP, as follows:[113]

(126) a. $[_{DP}$ **D** $[_{QP}$ NumP Q $[_{FP}$ AP F $[_{nP}$ NP DemP **n** [~~NP~~]]]]]

b. $[_{DP}$ **D** $[_{QP}$ NumP Q $[_{FP}$ NP AP F $[_{nP}$ ~~NP~~ DemP **n** [~~NP~~]]]]]

Whenever Dem raises to Spec,DP, the alternate realization of the aforementioned parameter holds, i.e. ¬A is true. According to Roberts (2017b: 180-183), the latter is realized for

[113] In (126) and in the subsequent derivations, phasal heads are boldfaced as in Roberts (2017b: 180-183). In this respect, it is worth recalling that the assumed phasal nature of the *n* head forces NP to stop at its edge before targeting any specifier c-commanding *n*P. It is for this reason that NP cannot be assumed to raise to Spec,FP in one fell swoop in (126b); ditto for other derivations to be discussed later on in the present section.

instance by A-N-Dem-Num and N-Dem-A-Num, as well as by the orders where the demonstrative is prenominal and the numeral is postnominal, i.e. Dem-A-N-Num, Dem-N-A-Num, Dem-N-Num-A. These five derivations are shown in (127) and can be opposed to those in (126) not only because of the different parameterization with respect to the structural position of Dem, but also because in (127) the NP raises (either alone or by pied-piping a projection containing it) to the outer specifier of Q, whereas in (126) the Q head always c-commands NP.

(127) a. $[_{DP}\ [_{FP}$ AP F $[_{nP}$ NP ~~DemP~~ **n** [~~NP~~]]] DemP **D** $[_{QP}$ NumP Q ~~FP~~]]
b. $[_{DP}$ NP DemP **D** $[_{QP}\ [_{FP}$ AP F $[_{nP}$ ~~NP~~ ~~DemP~~ **n** [~~NP~~]]] NumP Q ~~FP~~]]
c. $[_{DP}$ DemP **D** $[_{QP}\ [_{FP}$ AP F $[_{nP}$ NP ~~DemP~~ **n** [~~NP~~]]] NumP Q ~~FP~~]]
d. $[_{DP}$ DemP **D** $[_{QP}\ [_{FP}$ NP AP F $[_{nP}$ ~~NP~~ ~~DemP~~ **n** [~~NP~~]]] NumP Q ~~FP~~]]
e. $[_{DP}$ DemP **D** $[_{QP}$ NP NumP Q $[_{FP}$ AP F $[_{nP}$ ~~NP~~ ~~DemP~~ **n** [~~NP~~]]]]]

The contrast between (126) and (127) with respect to the relation between NP and the Q head can be captured by positing a binary parameter along the lines of A and ¬A, so that in (126) the NP overtly realized below the Q head instantiates B while the derivations in (127) are examples where ¬B is true, i.e. NP raises past the Q head. If one further assumes C and D, which respectively stand for momentarily unspecified bans on OV and on VS, it becomes possible to state the following:

(128) $(A \wedge B \rightarrow C) \wedge (\neg A \wedge \neg B \rightarrow D) \quad \equiv$
$(A \wedge B \rightarrow C) \wedge (A \downarrow B \rightarrow D)$

Each material implication in (128) bears an antecedent made of the conjunction of the two premises discussed for (126) and (127). Thus, eight possible outcomes per material implication are possible, but only one of them is not verified. In the implication where C is the consequent, the specific scenario that is ruled out is that where both A and B are true but C (i.e. the ban on OV) is false. The implication whose consequent is D is instead falsified if both A and B are false (i.e. both ¬A and ¬B hold) and D is also false (so that the band on VS does not hold).

The exclusion of the latter two scenarios is what forces C to hold in (126) and D to hold in (127), as in such groups of orders A and B are both true and both false, respectively. At this point, it is worth remarking that the truth value of C or D is not a concern as soon as A and B

do not bear the same truth value. For this reason, (128) seems to be a first good approximation of the facts also when extended to Dem-Num-A-N and Dem-Num-N-A, as in the latter two orders Dem is arguably high while NP does not raise past Q (i.e. A is false while B is true), and the two orders can be found among verb-final languages as well as among verb-initial ones (see again section 4.3.2). The fact that A-N-Num-Dem further aligns with the other five orders in (127) suggests that it should be derived as shown in (129). The latter should represent a welcome innovation with respect to the manner in which Roberts (2017b: 183) derives such an order of U20. As soon as the problematic way in which Roberts (2017b) treats FOFC exceptions is neglected (see again section 3.2.3), the derivation in (129) is not only plausible, but also circumvents the problems related to the constituency of FP that arise by arguing that "DemP appears not to move or at least to remain in a low position" (Roberts 2017b: 183).

(129) $[_{DP}\ [_{QP}\ [_{FP}$ AP F $[_{nP}$ NP ~~DemP~~ **n** [~~NP~~]]] NumP Q ~~FP~~] DemP **D** ~~QP~~]

However, the derivations for the four attested orders of U20 that have not been discussed yet (i.e. N-Dem-Num-A, N-A-Dem-Num, N-Num-A-Dem and N-A-Num-Dem) raise further theoretical problems to the present reworking of Roberts (2017b). The derivations that Roberts (2017b: 180-183) assumes for N-Dem-Num-A and N-Num-A-Dem are shown in (130a) and (131a), respectively. Although the application of (128) to (130a) leads to the expectation of a negative correlation between N-Dem-Num-A and verb-initial orders, this does not seem to be the case as per the data in appendix A.8 that have been previously summarized in table 4.5 on page 182. Conversely, the derivation in (131a) dictates that A is true while B is false. The parameterization in (128) thus expects the findability of N-Num-A-Dem with any possible linear order in mono-transitive clauses, but the absence of verb-initial languages from the data collected in appendix A.10 does not confirm this prediction. The evidence just recalled for both orders of U20 would be more in tune with (130b) and (131b), which derive the two orders of U20 in alternate manners that still comply with the premises in Roberts (2017b).[114]

(130) a. $[_{DP}$ NP DemP **D** $[_{QP}$ NumP Q $[_{FP}$ AP F $[_{nP}$ ~~NP~~ ~~DemP~~ **n** [~~NP~~]]]]]

b. $[_{DP}$ **D** $[_{QP}\ [_{nP}$ NP DemP **n** [~~NP~~]] NumP Q $[_{FP}$ AP F ~~nP~~]]]

[114] In (130b), NP or *n*P can also raise to Spec,DP (see also footnote 115 on page 196). Such alternatives would lead to the desired parameterization as per (128) nonetheless.

(131) a. [$_{DP}$ NP **D** [$_{QP}$ NumP Q [$_{FP}$ AP F [$_{nP}$ ~~NP~~ DemP **n** [~~NP~~]]]]]

b. [$_{DP}$ [$_{QP}$ NP NumP Q [$_{FP}$ AP F [$_{nP}$ ~~NP~~ ~~DemP~~ **n** [~~NP~~]]]] DemP **D** ~~QP~~]

The application of (128) to the derivations that Roberts (2017b: 180-183) proposes for N-A-Num-Dem and N-A-Dem-Num leads to the expectation of a negative correlation between these orders and verb-initiality. On the one hand, N-A-Num-Dem is derived by means of a sequence of consistent roll-up movements with DemP raising to Spec,DP, while N-A-Num-Dem is accounted for as in (132a). As already pointed out in section 4.3.3, these two orders appear in 11 and 4 verb-initial languages, respectively (see also appendices A.2 and A.7). Although such counterexamples can be regarded as exceptions to be investigated in the future, it is worth observing that no alternate derivation of N-A-Num-Dem seems possible to predict them by means of (128), whereas an alternative to (132a) that would predict such exceptions can be provided in (132b). However, both alternatives in (132) can be regarded as something that overfits the data (Merlo & Ouwayda 2018: 27), depending on one's analysis of the data collected in appendix A.7.

(132) a. [$_{DP}$ [$_{FP}$ NP AP F [$_{nP}$ ~~NP~~ ~~DemP~~ **n** [~~NP~~]]] DemP **D** [$_{QP}$ NumP Q ~~FP~~]]

b. [$_{DP}$ [$_{FP}$ NP AP F [$_{nP}$ ~~NP~~ DemP **n** [~~NP~~]]] **D** [$_{QP}$ NumP Q ~~FP~~]]

The discussion carried out so far shows that the parameterization in (128) would capture the collected data for at least ten of the fourteen attested orders of U20, provided that they are derived along the lines given by Roberts (2017b: 180-183). As long as the four remaining permutations are concerned, the ways in which two of them are derived (i.e. N-Dem-Num-A and N-Num-A-Dem) can be reworked so as to make them compliant with the present formal account. However, the inadequacy of the original derivations given by Roberts (2017b: 180-183) in this respect, i.e. (130a) and (131a), would remain a mystery to add to the other shaky aspects discussed in section 3.2.3 with respect to his account of U20.[115]

At this point, a bit more must be said with respect to the theoretical reasons that may back up the material implications proposed in (128). A first aspect of (128) that is perhaps surprising is the relevance that the Q head plays in establishing B. However, Syed & Simpson (2017) and Simpson & Syed (2020) provide a precedent in this sense by claiming that in

[115] Another potential problem is represented by the status of *n*P-raising. Roberts (2017b: 180-183) allows it, and in the main text it turned out to be necessary so as to provide the alternate derivation for N-Dem-Num-A given in (130b). However, if such an operation is legitimate, then nothing hinders the *n*P-raising to Spec,QP for Dem-N-Num-A in (127e), as well as the FP-internal raising of *n*P for Dem-N-A-Nm and N-A-Dem-Num in (127d) and (132a), respectively.

Bangla a Q head *à la* Roberts (2017b) is phasal. Their argument is grounded on the contrast between (133) and (134), which are adapted from Simpson & Syed (2020: 63).[116]

(133) a. *ei tin Te notun tupi*
DEM three CLF new hat
b. *ei notun tupi tin Te*
DEM new hat three CLF
'These three new hats'.

(134) a. *ei sat Ta notun tupi*
DEM seven CLF new hat
b. **ei notun tupi sat Ta*
DEM new hat seven CLF
'These seven new hats'.

The examples in (133a) and (134a) show that the most neutral order of U20 in Bangla is Dem-Num-A-N (Simpson & Syed 2020: 62). When the numeral is strictly lower than *five*, it is possible to raise [A N] to a position below Dem as in (133b),whereas the same is not possible if the numeral is equal or higher than five as in (134b).[117] By assuming that lower numerals like *tin* in (133) occupy Q^0 while the higher ones like *sat* in (134) occupy Spec,QP, Syed & Simpson (2017: Sec.5.1) and Simpson & Syed (2020: 4.2) argue that in Bangla Q bears the crucial phasal property of being an escape hatch, as any extraction of A and/or N from nominals must take place successive-cyclically through its specifier.

The claims reported following Syed & Simpson (2017: Sec.5.1) and Simpson & Syed (2020: 4.2) converge with the parameterization in (128) on page 194 in highlighting the relevance of Q in the derivation of a nominal extended projection *à la* Roberts (2017b). However, these observations cannot be easily reconciled with what Roberts (2017b) assumes to account for U20, nor with the reworking of his proposal that has been discussed in the present section. A first difficulty is represented by the fact that Roberts (2017b) crucially allows the presence

[116] Hsu & Syed (2020: Sec. 3.3) further discuss the difference between (133a) and (133b). In a nutshell, the presence of the demonstrative ensures the definite interpretation for both alternatives, but (133b) bears an inclusive reading that is not present in (133a). *Inclusiveness* is understood as in Lyons (1999), i.e. as "reference to the totality of the objects or mass in the context that satisfy the description" (Hsu & Syed 2020: Sec. 3.3).

[117] Simpson & Syed (2020: 63) further show that it is also possible to focalize A from (133a) by fronting it above Dem, whereas this is impossible for cases like (134a). In absence of Dem, (133b) keeps a definite interpretation, whereas that of (133a) is specific indefinite. See Simpson & Syed (2020: 63) and Syed & Simpson (2017: Sec. 1, 5.1) for further discussion.

of multiple specifiers per phrase, whereas Simpson & Syed (2020: 66 fn. 1) assume "that QPs in Bangla only have a single specifier position".

A second problem is represented by the fact that both proposals argue for the bi-phasal nature of DPs, but in Roberts (2017b) the phase heads are D and *n*, while in Syed & Simpson (2017) and Simpson & Syed (2020) they are D and Q. This fact confirms once more that scholars supporting the Functional Head Hypothesis propose divergent implementations of the parallelism between nominal and clausal domains (see again section 2.2.4). Moreover, assuming that Q is a phase in addition to (or instead of) *n* would not help solve the problems observed for the derivations of N-Num-A-Dem and N-A-Dem-Num in (131) and (132), respectively. If anything, the addition of an intermediate raising through Spec,QP would be problematic for the independent raising of Dem to Spec,DP. In fact, all the derivations shown in (127) on page 194 would no longer be incontrovertible cases where Dem is high, i.e. where A holds true as per the parameterization proposed on page 194.

Another interesting aspect of the material implications proposed in (128) on page 194 is that they relate parameters that apply in the nominal and in the clausal domains. This idea is in line with Longobardi, Crisma & Guardiano (2020: Sec. 3.3, 4.2), where implicational structures among distinct parameters are favored so as to reduce the workload needed for parameter setting. Given that the set of empirically grounded parameters can become rather large, it becomes implausible to think that people set all of them at the same time in order to acquire the target language properly, as this scenario raises an unmanageably huge number of competing alternatives. As soon as such parameters are instead related by means of implicational relations, the burden of the task is reduced by making certain choices irrelevant in the setting procedure.

If what is stated in (128) is on the right track, the acquirers infer that the target language cannot have both A and B hold as soon as they have evidence that OV is an unmarked order in the clause, or viceversa. Similarly, A and B cannot both be false in a language that allows verb-initiality, or viceversa. The fact that in (128) the parameters affecting the nominal domain are in the antecedent should not suggest a correspondent directionality in time. In fact, the statements in (128) have an equivalent counterpart in (135) by virtue of contraposition, so that the parameters affecting the clause can also be set first, followed by those assumed for the nominal domain:

(135) $(\neg C \rightarrow \neg(A \wedge B)) \wedge (\neg D \rightarrow \neg(\neg A \wedge \neg B)) \equiv$

$(\neg C \rightarrow \neg A \vee \neg B) \wedge (\neg D \rightarrow A \vee B)$

4.5.3 Different structures despite the same paremeter setting

Besides the problems already discussed for some of the attested permutations of U20, (128) and (135) raise an issue related to the evidence that the acquirers need to set C or D. As already discussed in section 3.2.3, Roberts (2017b: 165-167) explores the hypothesis that the value of A is so closely related to whether the extenal argument raises to Spec,TP that two of the four possible cases should not empirically findable. However, this prediction has counterexamples in both directions, which can be connected once again to the pervasiveness of the SVO order that was pointed out in section 4.1.3.

An alternate approach aimed at establishing how a ban on OV (i.e. parameter C) can be applied consists of verifying if the languages (discussed in section 4.3.1) that have attracted the interest of Roberts (2017b) bear some common structural property that can be pivotal in the definition of the parameter itself. An exploratory attempt in this respect can be pursued by looking at a subset of languages analyzed by Tvica (2017), whose explicit main goal is to compare the derivational paths of different SVO and VSO languages by testing the Rich Agreement Hypothesis. The latter follows the tradition paved by Pollock (1989) and related works already discussed in section 2.2.4, according to which the following statement given by Tvica (2017: 2) holds:

> [I]f the verb in a particular language has a rich paradigm of inflectional forms for subject agreement, the verb must always appear to the left of adverbs, as a consequence of 'v-to-I' movement, v-to-I movement being the operation that accounts for the correlation between verb movement and agreement morphology.

A central problem for the aforementioned hypothesis is represented by the definition of the exact properties that are supposed to determine the richness of the verbal paradigm (for discussion, see Tvica 2017: Sec. 2, Petzell 2021: Sec. 4). Without going into the complications of this theory-internal debate, there are two aspects of the crosslinguistic exploration made by Tvica (2017) that are worth considering here. On the one hand, Tvica (2017) does not investigate SOV languages due to a rejection of antisymmetry (Kayne 1994). In symmetric approaches like those presented in section 2.4.1, head movement is string-vacuous as long as it occurs among consistely head-final projections. As a consequence, Koeneman & Zei-

jlstra (2014: 604-605) and Tvica (2017: 90-91) argue that it is impossible to assess whether in head-final languages the verb raises past the adverbials that are usually taken (following Pollock 1989 and related works) to represent a diagnostics for its displacement within the clause.[118] Because of this theoretical premise, the cluster discussed in section 4.3.1 (where SOV is seen to negatively correlate with Num-N-A-Dem and Num-A-N-Dem) is the only one that can be tested against the languages sampled by Tvica (2017) in an effective manner.

On the other hand, Tvica (2017: Ch. 5) follows the sampling methodology provided by Rijkhoff & Bakker (1998) and collects a restricted set of languages where typological and genetic distance are enhanced. The members of such a set are subsequently divided according to whether they are "poor agreement languages" (Tvica 2017: Ch. 6) or "strong agreement languages" (Tvica 2017: Ch. 7). Six among the twenty-four languages he collects are characterized by prenominal numerals and postnominal demonstratives, namely Vietnamese (viet1252), Haitian (hait1244), Hmong Njua (hmon1264), Quiegolani Zapotec (sant1451), Egyptian Arabic (egyp1253) and Wolof (nucl1347). This is illustrated by the data in (136)-(141).

(136) *ba con trầu trắng của tôi này* [Vietnamese]
three CLF buffalo white of 1SG DEM
'these three white buffaloes of mine' (adapted from Nguyễn 2004: 3)

(137) *M wè de/twa/kat vwati sa yo sou wout la* [Haitian]
1SG see two/three/four cars DEM PL on street DEF
'I see those two/three/four cars on the street.' (adapted from Glaude 2013: 234)

(138) *ob phau ntawv loj nua* [Hmong Njua]
two CLF book big DEM
'those two big books' (from Harriehausen 1990: 144)

(139) *Per n-an y-rup mgyeey gin* [Quiegolani Zapotec]
but STAT-know POT-two man DEM
'But those two men knew...' (from Black 1994: 107)[119]

[118] For different views on the properties of head-final functional verbal projections, see Haider (2010: Ch. 2), Zeijlstra (2023: 196-197) and references therein.

[119] See Black (1994: 124 140) for examples showing that adjectives are postnominal in Quiegolani Zapotec. Black (1994: 329) further observes that the potential marker (glossed as POT) before the numeral bears an inclusiveness interpretation, while its absence fosters a non-inclusive reading.

(140) *it-talat sinīn signi dōl* [Egyptian Arabic]
DEF-three year.PL imprisonment.GEN DEM.PL
'these three years of imprisonment' (adapted from Woidich 2006: 230)

(141) *ñetti téére yu xonq yooyu* [Wolof]
three book REL.PL red DEM.PL
'those three red books' (adapted from Munro 2020: 4, cited in Cinque 2021)[120]

These six languages are already sufficient to show that what one would expect by elaborating on Roberts (2017b) is disproven once again by the facts. Not only these six languages do not belong to the same group as per the distinctions that Tvica (2017: Ch. 6-7) makes with respect to the richness of their verbal paradigms (the first three bear poor agreement whereas Quiegolani Zapotec, Egyptian Arabic and Wolof show rich agreement), but one also needs to make distinct structural assumptions to account for their clausal domains. This can be seen by starting from Haitian, which is an SVO language that lacks any agreement morphology between the Subject and the Verb, as the example from Tvica (2017: 120) in (142a) shows. Moreover, (142b) adapts the examples that Tvica (2017: 120-121) cites to illustrate that adverbs mandatorily precede the VP, while they cannot be linearly intraposed between Verb and Object. On these grounds, Tvica (2017: 121) argues that the most conservative analysis is to assume that the object does not leave the VP and the verb does not raise beyond *v* either.

(142) a. *mwen/ou/li/nou pale ak Gabi* [Haitian]
1SG/2SG/3SG/1PL spoke with Gabi
'I/you/he/she/it/we spoke with Gabi.'

b. *Boukinèt <preèske> kite <*preèske> Bouki*
Boukinèt <almost> leave <*almost> Bouki
'Boukinèt (almost) left Bouki'

Tvica (2017: 162-163) discusses examples that are summarized in (143), which shows that Hmong Njua is similar to Haitian in being an SVO language lacking Subject-Verb agreement, but that the former is different from the latter in allowing its adverbs appear exclusively to the right of the *v*P. As Tvica (2017: 167-168) observes, this fact can be captured in a twofold

[120] See Diouf (2009: 63, 101, 173) for related examples, and McLaughlin (2004) for discussion concerning the relativizer that precedes the adjective.

manner. On the one hand, one could allow right-adjunction of the adverb(s) to take place, but in such a case adverbs no longer represent a diagnostics of verb-raising in Hmong Njua. On the other hand, one could argue that adverbs are stranded by means of leftward movement of VP, which is an alternative that would comply with both symmetric and antisymmetric theories. In either case, it is not possible to state that the derivations in Hmong Njua and Haitian are perfectly convergent.

(143) *kŭ/kâo/nŵ/pé* <*qéng> *hlai* <*qéng> *mblê* <*qéng* [Hmong Njua]
1SG/2SG/3SG/1PL <*slow> harvest <*slow> rice <slow>
'I/you/he/she/it/we slowly harvest rice.'

Vietnamese also lacks Subject-Verb agreement, but adverb placement is different from both Haitian and Hmong Njua. The example in (144) follows Tvica (2017: 153) in showing that the definite object can be separated from the verb by a certain subset of adverbs. In case the object is indefinite, adverb intraposition is instead completely ruled out, and all adverbs can only occur in the other two places where they are also allowed with a definite object, namely either before or after the VP.

(144) *Tôi sẽ đọc cẩn thận/ nhanh chóng/ thướng xuyên/ *luôn luôn/ *chắc hẳn*
1SG will read careful/ quick/ often/ *always/ *probably
cuốn sách này
CLF book DEM
'I will carefully/quickly/often/always/probably read this book.' [Vietnamese]

Tvica (2017: 154-155) accounts for the facts just presented by assuming that head movement (if ever occurs) must be *v*P-internal as in Haitian when an indefinite object is involved. On the other hand, the facts in (144) can be explained by arguing that the definite object can undergo rightward extraposition by virtue of its own definiteness. Given that the featural makeup of the verb does not change depending on the (in)definiteness of the object, Tvica (2017) argues that an alternate analysis resorting to verbal raising would not be justified (in both symmetric and antisymmetric derivations).

Quiegolani Zapotec is a VSO language that bears rich agreement morphology and an obligatory presence of aspectual morphology. The latter factor and the linear distance between verb and object lead both Black (1994) and Tvica (2017: 170-171) to argue that verb-raising is

mandatory in the langauge. These facts are accompanied by the impossibility to have adverbs between the verb and its arguments, as shown in (145) following Tvica (2017: 170):

(145) *<yzhe> ts-a <*yzhe> noo <*yzhe> Mejiko <yzhe>*
<tomorrow> POT-go <*tomorrow> 1.EXCL <*tomorrow> Mexico <tomorrow>
'Tomorrow I will go to Mexico.' [Quiegolani Zapotec]

Wolof is an SVO language that features both *pro*-drop and a rich agreement morphology, "marking distinctions for subjects in (at least) three persons and two numbers" (Tvica 2017: 232). The behavior of manner adverbs further shows that the linear adjacency between Verb and Object is not mandatory in Wolof. All these facts are presented in (146), which summarize the data from Tvica (2017: 231-232).

(146) *(xác bi) <*ndànk> lekk -na <ndànk> jën wa <ndànk>* [Wolof]
(dog DEF.SG) <*slowly> eat -3.SG <slowly> fish DEF <slowly>
'The dog/he slowly ate the fish.'

By following the standard assumption that manner adverbs are *v*P adjuncts, Tvica (2017: 234-237) infers from (146) that the discontinuity between Verb and Object can be due to the raising of the former outside of *v*P, an analysis which is further grounded on the Subject-Verb agreement that Wolof shows. As shown in (147) by adapting the data from Tvica (2017: 209), the same observations also apply to Egyptian Arabic (though see Tvica 2017: Sec. 7.2.6 for detailed discussion).

(147) *(Layla) <*bi-sur'a> katab-it <bi-sur'a> gawab <bi-sur'a>*
(Layla) <with-speed> wrote-3.SG.F <with-speed> letter <with-speed>
'Layla/she quickly wrote a letter.' [Egyptian Arabic]

As anticipated, in the six languages just surveyed both the verb and the object do not show a consistent behavior that suggests a common property that can be taken as explanatory for the patterns observed both in Roberts (2017b) and in section 4.3.1. By virtue of this failed attempt, the next section will explore a conceptual alternative towards the account of the empirical facts.

4.5.4 The conceptual basis for a different solution

The different approach mentioned at the end of the previous section can be illustrated by means of the word order typology proposed by Broekhuis (2010: 17, 26) and already discussed on page 52 in section 2.3.1. The various ways in which he explicitly derives SVO, SOV, VSO and VOS can be reproposed here by sorting them anew, depending on the number of projections that are assumed to be dominated by IP. One can thus start with table 4.11 on the current page, which groups together the derivations where V-ext is projected while *v*-ext is not.

	Spec$_1$	**I**	**Spec$_2$**	***v*-ext**	**Subj**	***v***	**Spec$_3$**	**V-ext**	**V**	**Obj**
SVO	S	#			t_S	V + *v*	O	t_V	t_V	t_O
	S	V + *v* + I			t_S	t_{V+v}	O	t_V	t_V	t_O
SOV	S	#			t_S	*v*	O	V	t_V	t_O
VSO		V + *v* + I			S	t_{V+v}	O	t_V	t_V	t_O

Table 4.11: Orders in table 2.3 where V-ext is projected while *v*-ext is not.

As table 4.11 shows, an immediate way to ban VS (so that D assumed in section 4.5.2 holds) would be to force the raising of S to spec,IP. However, the data discussed in section 4.5.3 following Tvica (2017) question the adequacy of such a possibility for languages where such movements would be unnecessary or unmotivated. In a similar fashion, the derivations in table 4.12 (on on the facing page) effectively ban OV by having neither V-ext nor *v*-ext projected. In fact, in the typology proposed by Broekhuis (2010) it is the (optional) presence of such projections that feeds the raising of O past V. However, this would also prevent the derivation of VOS which requires both V-ext and *v*-ext, as shown in table 4.13 on the next page. The required premises to derive this order are in contrast with two facts. On the one hand, it is often an intra-linguistic alternative to VSO (see section 2.2.2), while on the other hand it can be found among the sampled languages that have either Num-N-A-Dem or Num-A-N-Dem as their unmarked order of U20 (see appendices A.4 and A.11).

Despite the limitations that have just been pointed out, the approach given by Broekhuis (2010) provides the premises by means of which a novel strategy can be pursued to formulate the bans on OV and VS. Such an alternative consists in counting how many times O and V raise within a single derivation. Starting from the verb-initial orders, it can be observed

	Spec$_1$	**I**	**Spec$_2$**	***v*-ext**	**Subj**	***v***	**Spec$_3$**	**V-ext**	**V**	**Obj**
SVO	S	#			t_S	*v*			V	O
	S	#			t_S	V + *v*			t_V	O
	S	V + *v* + I			t_S	t_{V+v}			t_V	O
VSO		V + *v* + I			S	t_{V+v}			t_V	O

Table 4.12: Orders in table 2.3 where neither V-ext nor *v*-ext are projected.

	Spec$_1$	**I**	**Spec$_2$**	***v*-ext**	**Subj**	***v***	**Spec$_3$**	**V-ext**	**V**	**Obj**
SVO	S	V + *v* + I	O	t_{V+v}	t_S	t_{V+v}	t_O	t_V	t_V	t_O
SOV	S	#	O	*v*	t_S	t_V	t_O	V	t_V	t_O
	S	#	O	V + *v*	t_S	t_{V+v}	t_O	t_V	t_V	t_O
VOS		V + *v* + I	O	t_{V+v}	S	t_{V+v}	t_O	t_V	t_V	t_O

Table 4.13: Orders in table 2.3 where both V-ext and *v*-ext are projected.

that O remains in its first-merged position while V raises twice (up to I) to derive VSO in table 4.12; O raises instead once (to Spec$_3$) while V raises three times to derive the same order in table 4.11, while VOS in table 4.13 is obtained by raising O twice (to Spec$_2$) and V four times. If *n* were the number of movements made by V while *m* the number of times O raises, one could say that what unites the three derivations just mentioned is that $n = (m + 2)$. Consequently, a ban on VS can be formulated along these lines:

(148) $n < (m + 2)$

The application of (148) also rules out three ways to derive SVO, namely those where V raises all the way up to I (i.e. the second derivation in table 4.11, the third in table 4.12 and the first in table 4.13). If these were needed in cases where (148) seems to hold otherwise, one could posit the raising of S to Spec$_1$ as an additional factor that prevents these SVO derivations from being excluded together with VSO and VOS.

A similar line of reasoning can be applied to the ways in which SOV is derived. In table 4.11, both V and O raise once (so that $n = m = 1$), whereas in the third derivation in table 4.13 V raises three times and O raises twice (so that $n = 3$ and $m = 2$). The second derivation in table 4.13 is slightly different, as V moves to V-ext while *v* raises independently

to *v*-ext, but still *n* and *m* would have the same value, so that $n = m = 2$. Based on these observations, a counterpart of (148) that bans the SOV derivations proposed by Broekhuis (2010: 10, 26) would be the following:

(149) $n > (m + 1)$

Again, (149) would rule out some SVO derivations, in particular the first in table 4.11 and the first two in table 4.12. However, the latter two can be salvaged (if needed) by further stipulating that (149) applies only if $m \neq 0$, namely only if O does not remain in its first-merged position.

Rather than attempting to argue for or against some of the possibilities just highlighted, the present discussion is aimed at showing that typologically relevant parameters may not be related to a specific head or feature, as they can instead be accounted for by assuming some counter that allows to better capture the facts.

Although the variable values of *n* and *m* in both (148) and (149) may suggest the presence of two independent counters, one can adapt the reasoning that Frank & Hunter (2021: Sec. 2.2) propose to account for multiple *wh*-fronting in Bulgarian, by means of which one single counter turns out to be sufficient. In fact, one can posit that each movement of V in the derivations *à la* Broekhuis (2010) increases the value of a given counter by 1, whereas each movement of O reduces the value of the same counter by 1. Thus, both (148) and (149) are checked against the value of such a counter at the end of the derivation, rather than by counting the absolute number of movements that both V and O require in each derivation.

However, the assumption of a counter that works along the lines just sketched raises a difficulty that can be introduced by following the abstraction presented by Frank & Hunter (2021: Sec. 3.1):

> Suppose we need to construct an arrangement of white and black pebbles in a line, so that there are an equal number of each, with all the white pebbles preceding all the black ones. One can imagine a variety of ways of doing this, of which two seem particularly canonical. In the first of these, we proceed from left to right, always extending the line of pebbles at its (right) end. We maintain a counter to keep track of the number of white pebbles already laid down. Once we start laying down black pebbles, we reduce the counter by one, and we conclude when the counter returns to zero. [... This implies w]orking at one end of the

> pebbles [... with] no bound on the number of white pebbles that can be laid down, [so that] there is no bound on the value of the counter. [...]
>
> The alternative derivation eschews the left-to-right construction of the pebble sequence, but instead alternates the placement of a white pebble to the left of the sequence so far created with the placement of a black pebble to the right of the sequence. [...] The possibility of working at both ends of the sequence [...] means that we no longer need to maintain a counter, or at least not one which counts any higher than 1.

As anticipated, neither of the two strategies that Frank & Hunter (2021: Sec. 3.1) describe hinges on two counters. In the first strategy, the string (of pebbles) is extended only on one side, while the second allows to have the string extended on both sides. The former option corresponds to the string analog of the Extension Condition assumed in Chomsky (2014–1995: 227 ff.), which states that the structure is always built at its root, while the second alternative corresponds to formalisms where countercyclic operations are allowed.

The latter fact is already sufficient to understand that the counter that would rule (148) or (149) on the derivations proposed by Broekhuis (2010: 10, 26) corresponds to the first strategy that Frank & Hunter (2021: Sec. 3.1) describe. The contrast with the alternate strategy shows however that there is a second factor worth considering. As long as the Extension Condition is respected, the exact value of the counter is relevant exclusively at the end of the derivation. In fact, if one wants to have the same amount of white and black pebbles, what matters is that the *final* value of the counter equals zero. Similarly, (148) is respected as soon as the counter is eventually not greater than 1, whereas (149) is respected as soon as it is eventually greater than 1.

However, under these conditions nothing bounds the value of the counter *during* the derivation.[121] Within the derivations proposed by Broekhuis (2010: 10, 26), one can add indefinitely many projections that trigger the raising of V without affecting O. Although the ban on VS in (148) rules out such an option (as $n \geq m + 2$ would hold), it would be perfectly legitimate under the ban on OV in (149), or if none of these bans apply at all.

As will be shortly illustrated, the converse scenario would also raise a similar problem. Such a case can occur, however, only by assuming that negative numbers are legitimate values

[121] The problem posed here thus seems to raise the same difficulties that would emerge if filters were applied to the obtained representations. See Chesi (2024: 8, 28) and references therein for further discussion.

in the counter, or alternatively by positing the following. If an operation that would reduce the (positive) value of the counter takes place when its value already equals zero, the syntactic operation is carried out while leaving the value of the counter unchanged. With either one of these premises, indefinitely many functional projections can be assumed where each of them triggers one raising of O without affecting V. If on the one hand the ban on OV in (149) would rule them out (as the value of the counter would be equal or smaller than zero), on the other hand their presence in the structure would be legitimate under (148), or when no ban holds.

As Frank & Hunter (2021: Sec. 2.2) observe, one would intuitively counterargue that the Phase Impenetrability Condition (see a.o. Chomsky 2001: 13) would prevent the occurrence of these scenarios. This is not the case, however, as under the PIC these indefinitely many projections are still theoretically possible phase-internally. As Frank & Hunter (2021: Sec. 2.2) put it (italics in the original), the PIC "constrains *where* the highlighted things are allowed to be, but not *how many*" of them can be relevant to the counter.

The alternate strategy that Frank & Hunter (2021: Sec. 3.1) provide (so as to have the pebbles arranged as desired) fixes the latter problem by imposing a "fixed finite bound on the amount of derivational state" (Frank & Hunter 2021: Sec. 1). Thus, the crucial aspect of this alternative consists in defining local domains where the number of movement-triggering projections is constrained. As will be discussed in sections 5.1.5 and 5.3.2, this comes in turn at the cost of stating either that the Extension Condition must be given up, or that it is not an ubiquitous principle in syntactic derivations.

The next chapter will present a formalism that endorses this latter strategy, i.e. Tree Adjoining Grammars (TAGs). Given that such an approach suffers from many long-standing theory-internal problems, an exploration of its merits must go hand in hand with an analysis of such issues. A number of solutions in this respect will be provided in section 5.4, in order to strengthen the grounds under which an alternate analysis of U20 will be proposed (in section 5.5). Although a TAG adaptation of the parameterization proposed in section 4.5.2 will turn out to be beyond the scope of the present work, the hope is that the upcoming discussion will provide a useful contribution to the continuation of this research. The conceptual reasoning just provided and the thorough revision of the formalism that will be proposed are thus aimed at showing (*contra* Longenbaugh 2014 and Georgi 2019) that the question is not *whether* TAGs can capture linguistically relevant facts, but rather *how* they can do that.

Chapter 5

Tree Adjoining Grammars

This chapter will present the formalism it is named after and will eventually provide a novel account for Greenberg's Universal 20 that is consistent with the insights that emerged in chapter 4. Before getting to that point (i.e. section 5.5), a thorough introduction of the formalism's basic properties is needed, which will be covered in sections 5.1 and 5.2.1.

The aspects to be presented in those starting sections actually allow to reframe some previous analyses of U20 in a TAG perspective, as sections 5.2.2 and 5.2.3 will show. Making TAG adaptations of previous proposals means that remergers will be reduced to the minimum and that successive-cyclic movements will be dispensed with entirely. However, the theoretical premises to be presented in sections 5.1 and 5.2 turn out to be theoretically or empirically inadequate in some respects, so that section 5.3 will be concerned with the proposals that have been put forth in the literature to overcome these difficulties.

Many problems have still been left open by the TAG literature, and section 5.4 will show that the content in section 5.3 sets the theoretical basis for a novel promising version of Tree Adjoining Grammars. Section 5.4 will thus illustrate how the to-be-proposed formalism allows to recast many insights that the generative tradition provides with respect to the clausal domain. Section 5.5 will then illustrate that the same can be done with respect to Greenberg's Universal 20 and with some of the problematic points that have been raised in section 3.1.

5.1 Setting the stage

As Findlay (2023: Sec. 2) observes, some properties of natural language syntax have been captured in a formal way by means of Context-Free Grammars (CFGs) at least since the late

1940s. However, CFGs fail to provide an adequate structural description of certain crossing dependencies that are attested in some Germanic languages (more on this in sections 5.1.2 and 5.4.3) and cannot derive structures that are "recursive in the representational sense" (Frank 2013: 227). The latter issue was already sufficient to make the syntactic model in Chomsky (2002–1957) start the derivations "with the restricted use of context-free rewriting [...] to produce non-recursive *kernel sentences*" which "were then combined using operations called *generalized transformations*" (Frank 2013: 228, italics in the original). While speculating on how automata theory might shed light on linguistic concerns, Chomsky (2004: 42-43) lists the following pros and cons with respect to the usage of CFGs and transformations for modeling natural language syntax (see also Frank 2004 for related discussion):

> Certainly context-free grammars represent some of the properties of languages. This seems to me what one would expect from applied mathematics, to see if you can find systems that capture some of the properties of the complex system that you are working with, and to ask whether those systems have any intrinsic mathematical interest, and whether they are worth studying in abstraction. And that has happened at exactly one level, the level of context-free grammar. [...] The systems that capture other properties of language, for example transformational grammar, hold no interest for mathematics. But I do not think that that is a necessary truth. It could turn out that there would be richer or more appropriate mathematical ideas that would capture other, maybe deeper properties of language than context-free grammars do. In that case you have another branch of applied mathematics which might have linguistic consequences. That could be exciting.

A branch like the one envisioned in the quotation above exists as a computational model since Joshi, Levy & Takahashi (1975) and has been used for linguistic purposes at least since Joshi (1985): Tree Adjoining Grammars (henceforth TAGs). What makes any TAG different from the other approaches presented thus far is that it is a tree rewriting system, which means that already constructed trees are taken as the simplest syntactic objects of the formalism (Weir 1988: 15, Joshi 2003: 245, Kuhlmann & Satta 2012: 618). Trees given at the beginning of a TAG derivation are called *elementary trees* and can be combined to one another by following precise formal constraints (discussed in Joshi & Schabes 1991, Kallmeyer 2010, Lichte 2015, a.o.). Any syntactic tree obtained as a result of the combination of various

elementary trees is called *derived tree*.

TAGs feature two further differences from the others approaches discussed so far. The first is that they are not forced to comply with structural uniformity (see Krivochen 2021 and Krivochen & Padovan 2021: Sec. 3). In fact, the elementary trees of a TAG need not be derived by means of previous combinatorics nor have to comply with strict binary branching or one single assumed syntactic template (such as the X-bar schema). These choices are up to the researcher, who can independently assume these kinds of structure and build the formal account accordingly. This is by no means a marginal aspect of the formalism, given that the second point of departure from the approaches previously discused is that "the set of elementary trees in the grammar of a language is the only locus of crosslinguistic differences, imposing substantial restrictions on the kind of variation that we will find crosslinguistically" (Frank 2013: 238). Because of this, TAGs can be included among the formalisms that free the researcher from typological uniformity if they do not need it, as no single syntactic template is imposed to all the languages of the world, and syntactic categories (and their properties) can vary across languages.

However, this is not to say that TAG approaches deny the existence of language universals that make any two natural languages comparable on a solid basis. On the contrary, Findlay (2023: 2089, smallcaps in the original) observes that the aforementioned view is "a position very much in keeping what M. C. Baker (2008a: 353) calls the Borer-Chomsky Conjecture, after Borer's (1984) proposal and Chomsky's (1995) later adoption of it, whereby parametric variation is restricted to the lexicon". Even more importantly, Abeillé & Rambow (2000: 42) state the following:

> The relevance of a principles-and-parameters methodology is, as observed, independent of whether one adopts a transformational approach in the style of Generative Grammar. The use of TAG as a language for formulating a theory of syntax is in some sense already a very strong theoretical position in favor of language universals, since the claim is that the constrained grammatical formalism represents (or captures) universal truths about syntax: we can use TAG to model natural languages syntax because natural languages syntax universally obeys certain principles. The difference between TAG on the one hand and [...] Generative Grammar on the other hand is that a portion of the universal is captured by the mathematical formalism and is not subject to linguistic theorizing.

More specifically, the unquestionable part of TAG is represented by the ways in which elementary trees combine to one another. This is done by means of two fundamental operations, which section 5.1.1 and section 5.1.2 will introduce. However, in the present chapter it will be argued that linguistic theorizing can, and in fact must, sneak in the mathematical formalism used by TAGs, even though the latter's core must be taken for granted as in the quotation above. This bears a consequence that is neatly remarked by Bleam (2000: 193):

> Any TAG-based theory of grammar must have two components: the formal system and a set of linguistic constraints. In explaining a given phenomenon there arises a tension as to which of these components carries the burden in accounting for the various properties associated with that phenomenon.

In section 5.1.3 and throughout section 5.3, it will be shown how this tension has been carried out in much of the TAG literature.

5.1.1 Context-freeness: Tree Substitution Grammars

The elementary trees and the subsequently derived tree generated by any TAG should look somewhat familiar to many syntacticians, as they satisfy "the requirement that every node in a tree have a single mother node (i.e. be dominated by only one node, distinct from itself)" (Krivochen & Padovan 2021: 19, Krivochen 2023b: Sec. 1.6 and references therein). The totality of elementary trees (i.e. the simplest units) allowed by a TAG is made of the union of a set *I*, whose members are *initial trees*, together with a set *A*, whose members are *auxiliary trees* (Joshi & Schabes 1991: Sec. 2, Kallmeyer 2009: 59). The nodes in the frontier, i.e. the leaves, of initial trees can be both terminal or non-terminal symbols, and in this respect auxiliary trees can be seen as initial trees enriched with a specific peculiarity to be discussed in section 5.1.2. In fact, the latter are pivotal in the second TAG operation which is not the topic of this subsection.

In TAGs, the fact that both initial and auxiliary trees contain terminal symbols is a crucial property due to a peculiar formal requirement of TAGs used for linguistic purposes, namely that any elementary tree must be *lexicalized*. This means that they must contain "*at least* one terminal symbol in their frontier" (Joshi & Schabes 1991: Sec. 2, italics added), which is called the *anchor* of that tree. As Kuhlmann & Satta (2012: 617) observe, this is what allows to obtain "combinations of elementary structures from a finite lexicon". Moreover, a

lexicalized grammar not only "cannot generate the empty string", but is also "always finitely ambiguous, because the length of the generated strings provides an upper bound on the size of the associated derived trees" (Kuhlmann & Satta 2012: 620). Finite ambiguity has not only the advantage of making the grammar decidable (Kallmeyer 2010: 21, De Santo, Aksënova & Graf 2016: 93), but also allows to devise efficient parsing algorithms Osterholzer (2017: 4).[122]

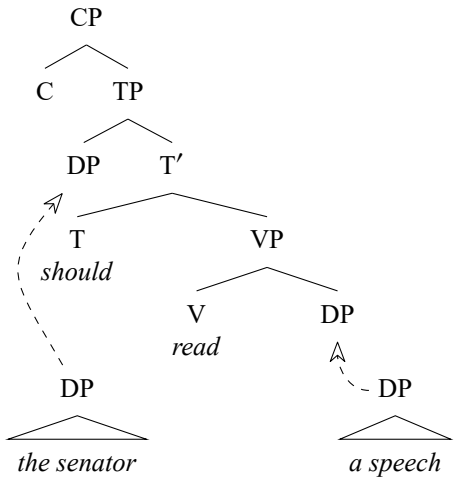
Take (150), which is an adaptation from Frank (2006: 153), as a concrete example of what has just been said. For each of the three elementary trees shown there, one of the nodes in their frontier that is associated with a phonetically pronounced element is the anchor of such trees. The leaves with terminal symbols cannot be the target of a TAG operation (Abeillé & Rambow 2000: 10), whereas the other nodes in the frontier, which are non-terminals, can.

In (150), an example of the latter is potentially represented by the C and the two DP nodes in the tree rooted in CP, and in fact, such two DPs are involved in the first fundamental TAG operation, namely *substitution* (see Kallmeyer 2009: 59 for a formal definition). An initial tree (say, the DP *the senator*) can combine with another elementary tree (like the one forming the main clause) if the label of the root in the former concides with a non-terminal node in the frontier of the latter. Thus, the former tree targets such relevant node (DP in Spec,TP) and, technically speaking, expands it by rewriting it (Abeillé & Rambow 2000: 4 ff.). As Krivochen (2023d: 339) observes, substitution "is consistent with a characterisation of structure composition as graph union", as it "does not create new nodes" which have the same label (so that there remains a single DP in Spec,TP even after a tree has substituted there). The two substitutions in (150) are represented with dashed arrows for a reason which will be explained in section 5.1.3, whereas the thus obtained derived tree is illustrated in (151).[123]

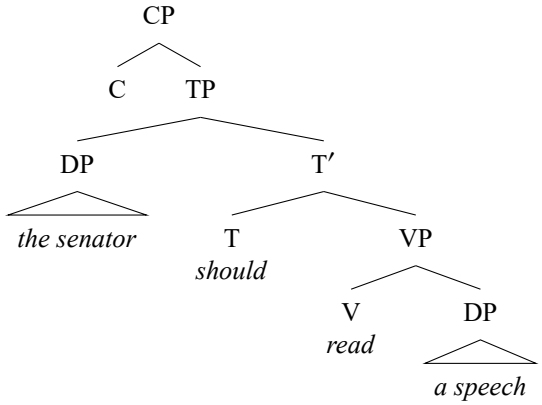
[122] See Joshi & Schabes (1991) and references in Lichte (2015: 131) for further formal motivations behind the need of lexicalization in TAG. In this respect, see also Kuhlmann & Satta (2012) and Engelfriet & Maletti (2017), whose results will be discussed in section 5.4.3.

[123] The fact that the C head in (150) is never targeted for substitution (so that the derived tree meets the desideratum of leaving C unpronounced) is not problematic. As Abeillé & Rambow (2000: 10) observe, "not all terminal symbols must or even may be rewritten".

(150)



(151)



A generative syntactician may wonder where the standardly assumed movement from Spec,*v*P to Spec,TP has gone in (150) and (151). This is a matter that will be discussed in section 5.1.5. For now, other related aspects of the structures just proposed will be clarified, namely what determines the number and the nature of the units used in the former structure to derive the latter. An answer in this respect can be given by relying on Frank (2002, 2004, 2006, 2010, 2013), where the following definitions are adopted.[124]

(152) Fundamental TAG Hypothesis (FTH), as per Frank (2006: 150, 2013: 233): Every syntactic dependency is expressed locally within a single elementary tree.[125]

(153) Condition of Elementary Tree Minimality (CETM), as per Frank (2002: 54, 2010: 375, 2013: 239): The syntactic heads in an elementary tree and their projections must form an extended projection of a single lexical head.

[124] See Lichte (2015: 139 ff.) for further discussion and for alternate definitions aimed at the same formal goals.

[125] Frank (2002: 22) omits the word *single*, but the uniqueness that (152) makes explicit is nonetheless implied. See also the Syntactic Lexicalization Hypothesis in Frank (2010: 373), which extends the FTH in (152) by stating that in any lexicalized grammar "syntactic relations entered into by a lexical item L must be expressed within the elementary structure that is associated with L".

(154) Θ-criterion (TAG version), as per Frank (2002: 55, 2006: 152, 2010: 375, 2013: 240):

a. If H is the lexical head of an elementary tree T, H assigns all of its θ-roles in T.
b. If A is a frontier non-terminal node of elementary tree T, A must be assigned a θ-role in T.

As one might guess, the fact that *the senator* substitutes at Spec,TP while *a speech* does so at the sister of V is partly due to (154), and it can be dictated by means of further specifications with respect to the features involved (as section 5.4.4 will show) or the semantics (see Krivochen & Padovan 2021 on the latter). Furthermore, the CETM in (153) is a way of determining the anchor of any elementary tree as opposed to the other leaves with terminal symbols, while the synergy of the definitions in (152), (153) and (154) are stated to establish the number and form of the elementary trees involved in a well-formed derivation, whose ultimately derived tree is said to be *saturated* (Kallmeyer 2010: 57).

For the present work, the relevance of (153) is twofold. First, it relies on the notion of extended projection given in Grimshaw (1991, 2005) and discussed in sections 2.2.3, 2.2.4 and 2.3.2, which means that a TAG approach *à la* Frank explicitly gives the tools for an interface with the proposal made by Roberts (2017b) for Greenberg's U(niversal) 20. Second, CETM goes beyond establishing the anchor of an elementary tree by also determining its lower and upper bounds, as well as the types of non-terminal symbols within.

The underlying principles regulating the derivation in (150) and (151) could well be part of a Tree Substitution Grammar (TSG), namely a formal system whose generative power is more restricted than TAG (as section 5.1.2 will discuss). Still, some formal properties of TSGs are relevant for the present chapter, as they foster the introduction of some key terms that will also appear in other parts of it (especially section 5.3.1). In fact, TSGs are weakly equivalent to CFGs, which means that the two "generate the same string languages" (Abeillé & Rambow 2000: 6).

Given that TSGs and CFGs characterize equivalent set of strings, they are said to bear the same *weak* generative power (see Joshi 2003: 243). Weak equivalence among different alternatives is important because it is an often used criterion to locate them "within one or another hierarchy of formal grammars" (Joshi 2003: 243. See also Maletti & Engelfriet 2012: Sec. 1 for further references and discussion). What makes TSGs preferable over CFGs is their increased *strong* generative power, which "relates to the set of structural descriptions

(derivation trees, directed acyclic graphs, proof trees, etc.) which a formal system is capable of assigning to the strings it specifes" (Joshi 2003: 243).[126]

The increased strong generative power of TSGs as opposed to the weakly equivalent CFGs is due to the former being *tree* rewriting formalisms, whereas the latter are *string* rewriting ones (Abeillé & Rambow 2000: 9 fn. 8, 10 fn. 9 and Frank 2004: Sec. 3; see also footnote 129 on page 221). With respect to linguistic purposes, TSGs thus bear a twofold advantage that Abeillé & Rambow (2000: 6) neatly illustrate (see also Joshi 2003: Sec. 3.1, Bangalore & Joshi 2010: 4 ff., Findlay 2023: 2075). Firstly, they have an "increased domain of locality" which "allows the linguist to state linguistic relationships (such as subcategorization, semantic roles of arguments, case assignment and agreement)" in a single tree rather than across rewrite rules. Secondly, the flexibility provided by a TSG is such that whenever the linguist using a CFG is forced by the formalism to make certain choices, "the TSG linguist may choose to do so or not on purely linguistic grounds" (Abeillé & Rambow 2000: 6).

5.1.2 Mild context-sensitivity: Tree Adjoining Grammars

Frank (2013: 259 fn. 4) observes that "substitution is capable of creating recursive structures" and cites "the substitution of the clausal complement of a verb into an elementary tree" as an example in this sense. In fact, by a slight modification of (150) it would be easy to account for the sentence *[The senator should read [$_{CP}$ that they are protesting]]*. It is the subcategorization in the initial tree anchored by *read* that must be modified. Hence, such a lexical verb selects for a CP node as its complement, so that the initial tree formed by the embedded clause can substitute there (given that the label of its root matches with the label of the sister of the lexical verb in the main clause).

However, a rewriting system can be recursive, but still not "in the representational sense" of the term, if it fails at providing "an infinite class of representations" from "a finite specification" (Frank 2013: 227). This happens to be the case for the weakly equivalent CFGs and TSGs, as Frank (2013: 227, 259 fn. 4) and Findlay (2023: 2075) show, because they both lack "a single application of the rewrite [...] operation that produces recursive structure" by containing "the same symbol twice" (Frank 2013: 227). In fact, initial trees are not supposed

[126] The fact that TSGs have a greater strong generative power as opposed to CFGs does not mean that the former strongly lexicalize the latter (see Joshi 2003: Sec. 3.2 for a demonstration), but rather that "there are some tree languages which can be described by a TSG for which an equivalent CFG does not exist"(Findlay 2023: 2074). See also Boullier & Sagot (2011) and Maletti & Engelfriet (2012) for related reflections where other formalisms are also considered.

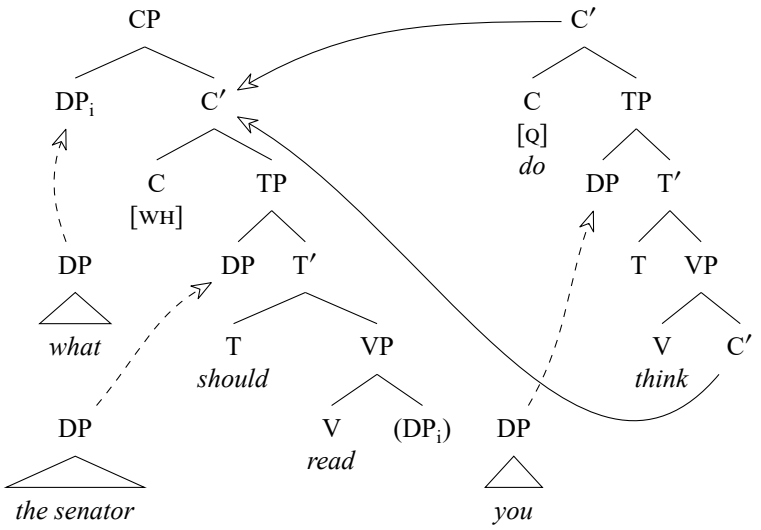
to have this property (Kallmeyer 2009: 59) and the power of TSGs is consequently limited as they contain no other types of elementary trees (Shieber 2014: 74).

A tree rewriting system that are intended to be recursive in the sense just outlined must be endowed with at least an elementary tree that contains a non-terminal frontier node with the same symbol of its root. In this sense, TAGs can be intended as an empowered version of TSGs where another kind of elementary trees is added to the set of initial trees for this very purpose. They are called *auxiliary* trees, and the non-terminal leaf whose symbol is identical to the root is called the *foot node* of that tree (Joshi & Schabes 1997: 71).

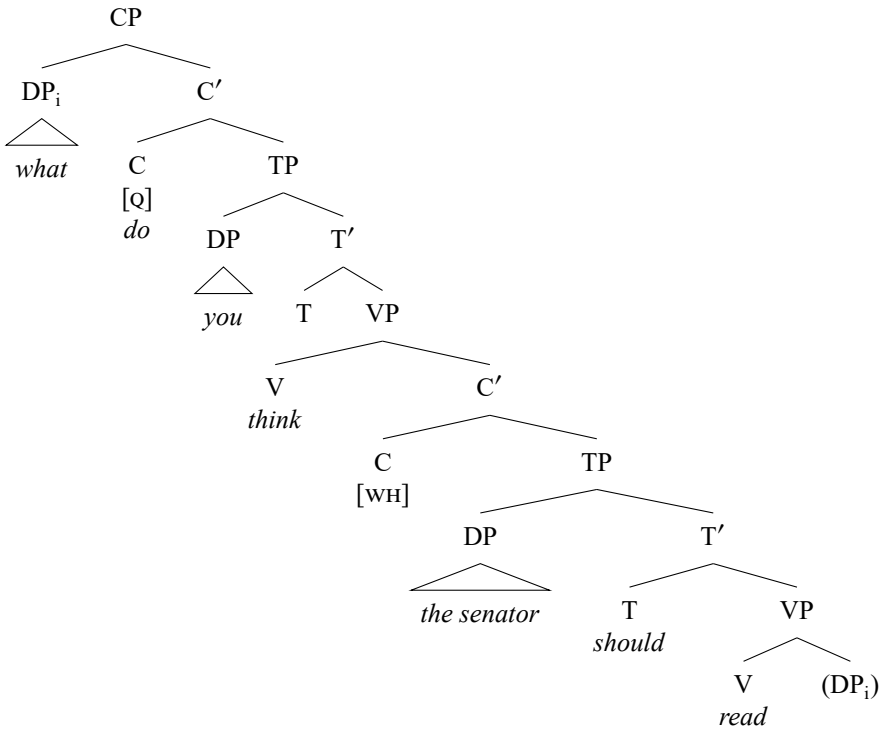
Auxiliary trees are pivotal for the application of the second fundamental TAG operation (which is what TAGs are named after), namely *adjoining*. In this operation, the root and the foot node of an auxiliary tree target a single internal (i.e. non-leaf) node of another elementary tree, but they can do so only if the latter has the same symbol as the first two (Joshi & Schabes 1997: 71, Kallmeyer 2009: 60). The targeted elementary tree is thus expanded by the auxiliary tree, which "shoehorns" the former at the target node (Krivochen & Padovan 2021: 3).

By starting from similar examples provided by Frank (2006: 155-157, 2013: 236), Frank & Hunter (2021: Sec. 3.2), and Krivochen & Padovan (2021: 18), the TAG operation just introduced can be illustrated by means of the elementary trees in (155), where the auxiliary tree that is recursive at C′ adjoins at the C′ of an initial tree that is a slight modification of the one in (150). The adjoining operation in (155) can also be recognized by the solid arrows, which is a choice to be further discussed in section 5.1.3. The structure in (156) shows the derived tree obtained from the combination of the elementary trees in (155). By following (among others) Diercks, Santorini & Kroch (2025: 304), the two [Q] and [WH] features (each appearing on a distinct C head) drive the T-to-C movement of the main clause and the *wh*-fronting, respectively.

(155)



(156)



Adjoining is the key operation of TAGs because it makes them capable of providing an adequate structural description for sentences like the Swiss-German one in (157), adapted from Shieber (1985: 335) and Frank (2004: 675):

(157) *Jan säit das...* ('John says that...')

... mer d'chind em Hans es Huus lönd hälfe aastriiche

... 1PL DET.children.ACC DET.DAT Hans DET.ACC house let help paint

'... we let the children help Hans paint the house.'

What makes (157) relevant is that its subordinate clause is a well-formed case of cross-serial dependencies in natural languages syntax that can be represented as $DP_1DP_2DP_3V_1V_2V_3$, where subscripted integers indicate dependencies between nouns and verbs that are explicitly enforced by case marking.[127] Joshi (1985) shows that TAGs belong to a class of formal grammars that manage to handle the cross-serial dependencies in (157), whereas CFGs fail at assigning the correct structural description (Shieber 1985: Sec. 2-3, Frank 2004: 673-675, Findlay 2023: Sec. 2).

With respect to the Chomsky-Schützenberg hierarchy (see Chomsky & Schützenberger 1963, Jäger & Rogers 2012), TAGs are thus called *mildly context-sensitive* because they share many welcome formal properties of CFGs while having a weak generative power that goes beyond context-freeness but remains below full context-sensitivity (Frank 2004, Lichte 2015: 134-136). Therefore, a well-known claim in favor of TAGs is that they dispose of the minimal power necessary to provide an adequate formalization of natural language syntax (and the same holds for the mildly context-sensitive alternatives discussed in Weir 1988, Kallmeyer 2010: 215 Chen-Main & Joshi 2014: 990, Osterholzer 2017: Ch. 6, Frank & Hunter 2021: Sec. 1, and Findlay 2023: 2091).[128]

Another reason that makes adjoining a central operation in TAGs is that it makes successive cyclicity dispensable with. As (155) shows, a *wh*-element substitutes at Spec,CP of the elementary tree anchored by the lexical verb selecting it (i.e. *read*) as a consequence of the FTH in (152), page 214. The eventual long-distance relation between the *wh*-element and the verb assigning a θ-role to it is not the result of several intermediate movements of the former, but rather of the insertion of in-between local domains defined by auxiliary trees that must also comply with FTH, CETM and the θ-criterion established by Frank (2002, 2004, 2006, 2010, 2013). These premises lead Frank (2002: 22) to state the following:

(158) Non-local Dependency Corollary:
Non-local dependencies always reduce to local ones once recursion is factored away.

[127] For relevant considerations about the syntactic contrast between crossing and nesting dependencies, see a.o. Richards (1997: Ch. 3, 5), Frank (2004: Sec. 2), Heck (2016: 18 ff.), Tollan & Clemens (2022), and Krivochen (2023b: Sec. 4.3.2).

[128] See however Frank (2002: Sec. 1.4), Frank (2004), Stabler (2019), S. Müller (2023b: Sec. 12.6.3, Ch. 17), Frank & Hunter (2021) and Findlay (2023: 2092-2093) for relevant discussion on whether and why one should prefer mildly context-sensitive grammars over much more powerful alternatives.

Another way of putting (158) is the slogan proposed by Bangalore & Joshi (2010: 2), namely "complicate locally, simplify globally". On the one hand, the local complication consists in having "[t]he primitive structures" of TAGs "provide a *local* domain for specifying linguistic constraints" (Bangalore & Joshi 2010: 4, italics in the original). On the other hand, the global simplification commonly consists in dispensing with successive ciclicity (see e.g. Joshi & Schabes 1997: 95 ff., Joshi 2003: 250), so that (158) is exploited in order to eventually avoid making syntactic movement an (almost) all-pervasive property of the whole machinery.

In fact, TAGs rely on a formal tool which has been out of use in the generative tradition since Chomsky (1965), namely *generalized transformations* (Lasnik & Lohndal 2013: 30 ff., Frank 2013: Sec. 8.1). By means of the latter, TAGs diverge from many of the other approaches previously discussed in the present work because they can, but need not, "include rules for singulary transformations (such as wh-movement)" (Krivochen & Padovan 2021: Sec. 4).

What has just been said further underlines that TAGs may look problematic to generative syntacticians. In fact, no movement from Spec,*v*P to Spec,TP is needed (as already pointed out in section 5.1.1) and no *wh*-movement is either, as the bracketed non-terminal DP in (155) and (156) suggests. Although TAGs are not "inherently incompatible with singulary transformations" (Krivochen & Padovan 2021: Sec. 4), their inclusion in the formalism needs some further explicit statements which will be discussed in section 5.1.5.

There are also other potential issues that arise from (155) and (156). Why is the main clause an auxiliary tree expanding the embedded one, rather than the opposite? Why should the C head projecting a CP in the *should-read* tree in (155) eventually fail to project up to CP in the derived tree (156)? And is there a non-stipulative reason to motivate that the C head in the extended projection of the *think* auxiliary tree triggers the T-to-C raising? In order to reply to such questions (which will be ultimately answered in section 5.4), the discussion first needs to turn to other relevant facets of the TAG approach, to be faced in the upcoming subsections of sections 5.1 and 5.3.

5.1.3 Derivation trees

A derived tree like the ones in (151) and (156) cannot represent the only theoretical object of a TAG, because they fail to illustrate how they have been obtained by means of the operations

that have been applied throughout their derivation. In fact, apparently similar derived trees, and their subsequent linearizations, can be the result of rather different processes.[129]

The distinction among apparently similar derived trees can be inferred from what is called the *derivation tree*, whose primary goal is to illustrate the protocol of the derivation and to characterize the derived tree in an unambiguous manner (Lichte 2015: 130; see also Vijay-Shanker 1987: 11 ff.).[130] Kallmeyer (2010: 57 fn. 3) clarifies the relation between derivation trees and derived trees in the following way:

> [T]here can be more than one derivation tree for a single derived tree since there might be more than one combination of adjunctions and substitutions with elementary trees leading to the same result. However, for a given derivation tree, the derived tree is uniquely determined.

Both the derivation and the derived trees are said to be *saturated* once all the necessary substitutions and adjoining operations have been performed (see section 5.1.1 and Kallmeyer 2010: 57). Frank (2013: 244) further defines this concept as follows with respect to derivation trees:

(159) Derivational completeness:

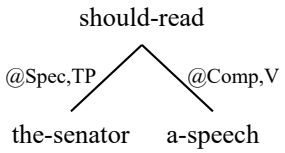
a. The root of a derivation tree cannot be an auxiliary tree.
b. Any node of a derivation tree must have as many substitution daughter nodes as its corresponding elementary tree has substitution sites.

The derivation tree illustrates the protocol of a derivation by representing the elementary trees involved as nodes, so that there is one and only one node per elementary tree. Similarly, one and only edge can link any two different elementary trees T_1 and T_2, because it stands for the single operation that combines them, be it either a substitution or an adjoining operation. If T_2 expands and rewrites a node of T_1 in the derived tree, then T_2 is dominated by T_1 in the derivation tree, in a daughter-mother kinship (Abeillé & Rambow 2000: 10). The derivation tree for the the sentence *The senator should read a speech*is used below. The elementary and derived trees were previously illustrated in (150) and (151) from page 214 on.

[129] As Weir (1988: 21) puts it: "Like the *strings* derived from a CFG [...], the *trees* derived by a TAG do not encode their own derivation history". The italics are in the original text, and they are arguably used to highlight what Abeillé & Rambow (2000: 9 fn. 8) explicitly remark: "It is sometimes said that in a CFG, the derived tree is the same as the derivation tree. This is not strictly speaking correct. As a string-rewriting formalism, a CFG does not produce any derived tree at all, only a derived string".

[130] The main text above is an adaptation of the following German text in Lichte (2015: 130): "Der Ableitungsbaum ist das Protokoll der Ableitung und charakterisiert den abgeleiteten Baum in eindeutiger Weise".

(160)



The structure in (160) illustrates some common practices in the TAG literature. In particular, the edges linking the daughter nodes (indicated as *the-senator* and *a-speech*) with their mother (*should-read*) are represented with solid lines that can be enriched with labels bearing some further information. In (160), such labels indicate "the address of the node in the parent tree to which the adjunction or substitution has been performed" (Joshi & Schabes 1997: 75). Given that CETM in (153) ensures for each node of an elementary tree to be essentially unique, its label (or features, as section 5.4.4 will discuss) can be used to identify where rewriting has occured throughout a derivation like (160). This allows to dispense with the alternate notation that section 5.1.4 will discuss.

In the TAG literature, there are two further kinds of labels that can be found on derivation trees. The first is concerned with the grammatical relation between a mother node and her daughter(s), as exemplified by Abeillé & Rambow (2000: 11), Frank (2013: 243, 251, 253) and Findlay (2023: 2086), while the second specifies the type of operation applied to combine the two trees linked by the edge, as in Chiang & Scheffler (2008: 18, 23). For the present purposes, the former type of labels can be inferred from the derived/elementary trees and is thus superfluous on the derivation trees, while the latter deserves some further discussion.

In fact, labeling the edges of a derivation tree with the type of operation it stands for is mostly relevant to those formalisms where further alternatives to the standard substitution and adjoining operations are assumed, e.g. *reverse adjunction* (see Chiang & Scheffler 2008 and references therein, as well as section 5.4.1). However, in most TAG works the distinction between substitutions and adjoining operations is clarified on the derivation tree by means of a labeling convention that applies to its nodes, which stand for the elementary trees involved in the derivation. In particular, Greek letters are generally used in such a way that "α and its subscripted and primed variants [stand] for initial trees, β and its variants for auxiliary trees, and γ and its variants for elementary trees in general" (Schabes & Shieber 1994: 92 fn. 1).

For most TAG formalisms, this strategy is already sufficient to account for the difference between substitutions and adjoining operations, especially if the initial trees used are also represented in detail. In a derivation tree, if a daughter node α stands for an initial tree, then its relation with the mother node γ must have been mediated by means of substitution. If

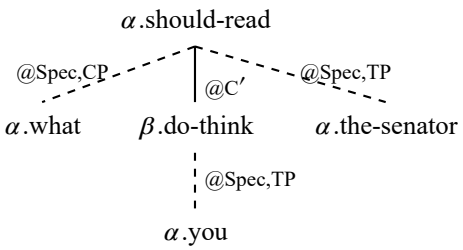
instead a daughter node β stands for an auxiliary tree, it must have adjoined at an inner node of γ.[131]

There are however some exceptions to this trend, e.g. Joshi & Schabes (1997: 76) and Joshi (2014: 151), where the distinction made through Greek letters on the nodes goes along with a different representation of the edges of the derivation tree, i.e. dashed lines for substitutions and solid lines for adjoining operations.[132]

Although this further visual difference is somewhat redundant in TAG approaches like Joshi & Schabes (1997), it will be extremely useful for the remainder of this chapter, for reasons that will eventually become clear in section 5.4.

Besides, the same dichotomy has been adopted to illustrate how elementary trees relate with one another in (150) and (155), and will be pursued further in the remainder of this work. The TAG derivation in (155) on page 218 (which accounts for the sentence *what do you think the senator should read*) is now represented with its corresponding derivation tree in (161), which is an example of what the derivation trees will look like from now on.[133]

(161)



5.1.4 An aside: Two remarks on derivation trees

The present section will discuss two points that bear a marginal role in the general discussion of the present work, but that might interest those who are already used to TAGs. The first point is concerned with the motivations for not adopting a notational convention that sometimes is used in the TAG literature. The second point elaborates instead on Frank (2002) *et seq.* in

131 In the cases illustrated above, the letter γ represents a concise way to indicate that it does not matter if the targeted tree is an initial or an auxiliary tree.

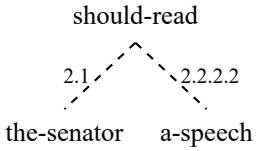
132 Frank (2013: 243, 251, 253) and Gardent & Narayan (2015: 54) adopt the same graphic dichotomy on the edges of their derivation trees, but do not use the Greek letters to identify the nodes. Kallmeyer (2010: 56 fn. 1) and Pogodalla (2017: 534) switch instead the contrast, so that instances of adjoining are represented with dotted arrows.

133 The order of sister nodes in (161) is arbitrary, as nothing in the theoretical premises made here explicitly imposes the one chosen above (see Frank 2006: 160 for an equally good alternative of the same derivation tree). For the task at hand, this is not a problem. However, section 5.2.1 will show that a more precise specification in this sense can improve the descriptive adequacy of TAGs, especially when the sister nodes in the derivation tree are auxiliary trees.

order to provide a novel formal statement related to derivation trees. Readers who are not interested can move to section 5.1.5 on page 228.

Gorn addresses, named after Gorn (1967), represent another way of labeling derivation trees so as to indicate "the address of the node in the parent tree to which the adjunction or substitution has been performed" (Joshi & Schabes 1997: 75). In a Gorn address, "k is the address of the k^{th} child of the root node, and $p.q$ is the address of the q^{th} child of the node at address p" (Joshi & Schabes 1997: 75 fn. 4). This can be illustrated by means of (162), which adapts the derivation tree first shown on page 222 by changing the labels accordingly and by making the edges (that stand for the two occuring substitutions) dashed as in (161).

(162)



The counting of the nodes (illustrated a.o. in Longenbaugh 2014: 23 fn. 6) must be performed as in "the breadth-first level-order traversal" by following "the standard convention" of traversing "children/subtrees from left to right" (Toyoshima 2013: 361 fn. 3, see also Krivochen 2023c: 8-9). Moreover, Findlay (2023: 2086 fn. 8) remarks that in the TAG literature the root of an elementary tree can have the address *0* or ϵ, as can be observed in Joshi & Schabes (1997: 76) and Kallmeyer (2010: 57), respectively.

Gorn addresses will not be used in the present work because they bear two difficulties. The first one is theoretical: in those approaches where the TAG operations are carried out step by step (rather than all at once), the following issue pointed out by Kallmeyer (2010: 57-58) emerges:

> [A]fter having performed an adj[oining operation] on a tree γ, the addresses of nodes in γ below the adj[oining] site change. Consequently, when adjoining (or substituting) later to such a node, instead of the current address of this node, we need to take its address in its original elementary tree as the edge label in the corresponding derivation tree.

The second difficulty related to Gorn addresses is instead merely practical. As the derivation tree in (160) shows, Gorn addresses are not well-suited for binary branching structures. In fact, only three digits are used (i.e., 1, 2 and 0, the latter only for root nodes), the ad-

dress can become rather long (as for the substitution site of the *a-speech* tree) and its manual calculation on the original tree in (150) a rather tedious task.

The second point that this section wants to discuss is that the word *tree* in *derivation tree* must be intended in a graph-theoretic sense.[134] This is a novel contribution that is made by reformulating an argument previously brought by Frank (2002: 29-30, 74-79, 201-214), Frank (2006: 158-160, 171) and Frank (2013: 245, 250-253), where it is stated that TAG derivations must obey a Markovian condition, identified as such because "the possible next steps are determined only by the current state and not on the past" (Frank 2013: 245). The exact formulation changes from Frank (2006: 159) to Frank (2013: 245), but both variants ultimately posit the same restriction. They are reported in (163a) and (163b), respectively.

(163) Markovian condition on TAG derivations:

a. Each step in a TAG derivation must involve the well-formed combination of two elementary trees.
b. The well-formedness of a TAG operation is determined by the pair of elementary trees in a parent–child relationship in the derivation tree.

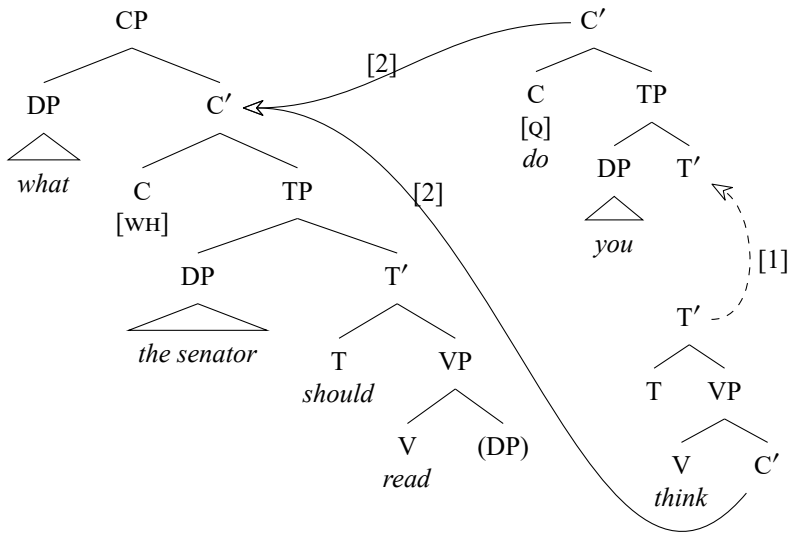
In Frank's view, the derivations that would violate (163) require an initial tree that is adjoined to a target tree even though the former lacks the necessary foot node to run the operation in a proper way. Because of this, Frank's argument in favor of (163) consists in showing that illegitimate derived trees (e.g. super-raising and *wh*-islands violations discussed in Frank 2013: 250-254) imply ill-formed elementary trees, while derivation trees only play an ancillary role in the proof. However, the statement in (163) can also be suppoted by starting from a legitimate derived tree. This is possible because the latter can be obtained through different derivations (as per the quotation on page 221 from Kallmeyer 2010: 57 fn. 3), including those that are not legit for independent reasons.

Such a case can be exemplified by supposing that the derived tree shown in (156) on page 218 is not obtained by means of the auxiliary tree β*.do-think*, as in (155) and (161), but rather by using two initial trees in its stead, i.e. α*.do-you* and α*.think*, that combine with one another by means of substitution and then adjoin at α*.should-read* just like β*.think* would. This alternative, which necessarily subsumes a TAG approach where the substitutions and the adjoining operations take place step by step, is depicted in (164). The substitutions of the

[134] See Krivochen (2023b) for related discussion on the application of graph theory for linguistic purposes.

various arguments are not shown because not relevant for the present concern.

(164)



It is easy to see that if *α.think* substitutes at the T′ node of *α.do-you*, the derived tree would be identical to *β.do-think* in (155), apart from the (non-)susbstituted argument. One can ban this scenario on several independent grounds. For instance, by appealing to FTH, CETM and the theta-criterion respectively in (152), (153), (154), just like Frank (2013: 250-253) does for the trees he discusses. Moreover, one can also observe that *α.do-you* is not properly anchored as per the assumptions made so far. However, (164) would nonetheless give rise to the well-formed derived tree in (156).

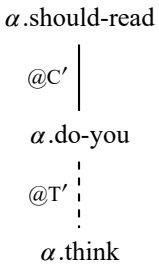
The case depicted in (164) can be generalized as follows. Given two elementary trees γ_1 and γ_2 (previously exemplified by *α.do-you* and *α.think*, respectively), where the latter substitutes at a node in the frontier of the former, one could hypothesize for them to behave as a single auxiliary tree that adjoins to an inner node of a third tree γ_3 (as *α.should-read* in the previous case).[135] If the label of the root of γ_1 matches with the one of a node in the frontier of γ_2, the latter node could work as a foot node, so that γ_1 and γ_2 could be used as a derived auxiliary tree to be adjoined somewhere. The Markovian condition in (163) effectively bans this scenario, and according to Frank it does so because an adjoining operation leads to a well-formed combination only if the source auxiliary tree is "born and not made" (Frank 2002: 78, 2006: 160, 2013: 245).

Frank (2002) pays little attention to derivation trees, as they are only briefly discussed

[135] The scenario where γ_2 adjoins at a node in the frontier of γ_1 can be set aside from the present discussion. In fact, in many works it can apply exclusively at inner nodes of trees (as already pointed out in section 5.1.2). But even if this possibility is allowed, then γ_1 already bears the relevant label at that node in its frontier, so that having a tree γ_2 adjoined there is of little help in forming a new derived auxiliary tree that features a node whose relevant properties are not already present in γ_1.

in Frank (2002: 78, 291 fn. 3). Frank (2006: 160) shows instead a well-formed pair of derived and derivation trees that comply with (163) and then considers an ill-formed derived tree, whose corresponding derivation tree is based on the shape of the well-formed one previously shown. The result of this process (which Frank 2013: 243, 251 replicates) leads to the structure in (165) if applied to our case in (164).

(165)

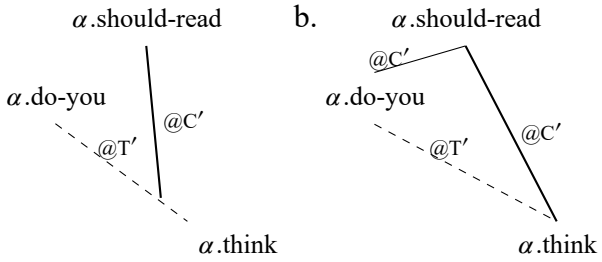


Just like the derivation trees in Frank (2006: 160, 2013: 251), the structure in (165) shows that *α.think* is *α.do-you*'s daughter, because it substitutes at its T′ node, while *α.do-you* is *α.should-read*'s daughter because it shoehorns its C′ node. Frank's discussion goes on to argue that the latter parent-child relation is not well-formed, and thus violates (163), because it would be possible only if the necessary foot node (C′) had been part of *α.do-you* from the beginning.

Although the conclusion that Frank (2006: 160, 2013: 251) draws is sound, (165) and the derivation trees he shows are *not* the actual representations of the respective ill-formed derived trees. Taking (165) as an example that fits all, it is the «incestuous» aspect of (164) that is overlooked. In fact, *α.think* is also a daughter of *α.should-read*, because the former directly adjoins to the latter just like *α.do-you* does. The combination of *α.do-you* and *α.think* that is required to carry out the adjoining operation is what makes *α.do-you* both a mother and a sister of *α.think*.

In the *actual* protocol of the derivation accounting for (164), one single edge is supposed to connect both *α.do-you* and *α.think* to *α.should-read*, because they adjoin to the latter as a single unit. The representation in (166a) tries to capture these facts, but it must do so by making *α.should-read* connect not with a node, as in standard assumptions, but rather with the edge that connects *α.do-you* and *α.think*. The alternative in (166b) tries to fix this by having all nodes connected as expected, but this violates another standard assumption on derivation trees, namely the bijective relation between the set of edges and the set of TAG operations in a derivation tree.

(166) a. b.



The discussion just made shows that the Markovian condition on TAG derivations in (163) has a more solid foundation by focusing on the limits imposed by the derivation tree in its graph-theoretic sense. In fact, derived trees that violate (163) are ill-formed because the protocols of their derivations are not trees. In (165) and (166a) at least a pair of vertices is not connected as expected, while in (166b) for any daughter of the root node there are two walks (instead of one) connecting the former with the latter. The formulations in (163) can thus be replaced with something like the following:

(167) Revised Markovian condition on TAG derivations: Since TAG derivation graphs are trees (i.e. connected and acyclic), any pair of daugthers of the same node cannot be linked by means of an edge.

5.1.5 On the interaction between Merge and TAG operations, pt. 1

It is now time to introduce the possible ways to reconcile a TAG approach with apparently diverging tenets adopted in Minimalism, in particular in Chomsky (1995, 2000, 2001, 2008). One major problem is represented by the non-terminal nodes in the frontier that are targeted by substitution, as in (150) and (155) on page 214 and on page 218, respectively. As Frank (2006: 152) remarks, such substitution nodes "cannot come about in the way non-terminals are usually thought to arise, namely via Merge, as they do not include any projected elements. Consequently, the non-terminals must be included in the lexical array that feeds the elementary tree's derivation".

In the quotation above, *lexical array* is understood as the often assumed set of lexical items (each of which is actually a bundle of features) that are manipulated in the course of the derivation driven by Merge and Agree (Citko 2014: 8, Chesi 2015: 77). The inclusion of non-terminals in the lexical array seemingly violates the part of the inclusiveness condition which dictates that "any structure formed by the computation [...] is constituted of elements already present in the lexical items selected for" the lexical array (Chomsky 1995: 228). However,

Frank (2002, 2004, 2006, 2010, 2013) claims that the two views can be reconciled as soon as two factors are considered. The first is the theta-criterion in (154) on page 215, which allows Frank (2006: 152-153) to solve the incompatibility with the inclusiveness condition by supposing the following:

> [S]election of a lexical head for inclusion into a lexical array also necessitates the inclusion of the features determining the syntactic realization of that head's arguments, specifically in the form of non-projected non-terminals of appropriate syntactic categories. [... Then, no problem arises] in a derivationally constructed elementary tree so long as all of the non-projected non-terminals in a lexical array are incorporated in a derivation.

To understand why Frank requires a lexical array to contain non-terminals in place of the fully projected arguments, the second anticipated crucial factor must be considered. Whenever a Minimalist approach assumes lexical arrays (see a.o. Chomsky et al. 2023 for an example where this notion is dispensed with), a certain sentence must be accounted for by making the required ones become numerations. The term *numeration* indicates a lexical array which is "augmented by information on how many times each lexical item is selected from the lexicon" (Citko 2014: 8-9).

In a TAG approach *à la* Frank, however, there is not one numeration per phase as in Chomsky (2000), but rather one numeration per elementary tree (see Frank 2002: 260 en. 21). On the one hand, this implies that any two fully projected arguments cannot be part of the same numeration, because each of them constitutes its own extended projection that must comply with the CETM in (153) on page 214. As a result, Frank (2002: 124) argues that in his approach "there is a strict limit on how many lexical elements may appear within a numeration".

On the other hand, many types of lexical verbs anchoring their own elementary trees (though not all, as will be shown in due course) are assumed to project up to CP. This implies that, although "there is a natural affinity between the idea that economy is computed over separate derivational phases and the idea that well-formedness [...] is computed over separate elementary trees" (Frank 2002: 260 en. 21), the TAG approach being discussed does not have to assume the phasality of *v*P and its related underpinnings. Frank (2006: 154 fn. 6) clarifies this well (italics in the original):

> [W]e would not necessarily want to assume that wh-movement of an object

> passes through specifier of *v*P. Since elementary trees are larger than phases, the requirement that movement take place within a local domain does not demand such intermediate steps of movement. It is worth noting that even if we were to require such intermediate steps of movement within an elementary tree, there would be no corresponding movement through specifier of *v*P in higher clauses through a long-distance-moved wh-element passes.

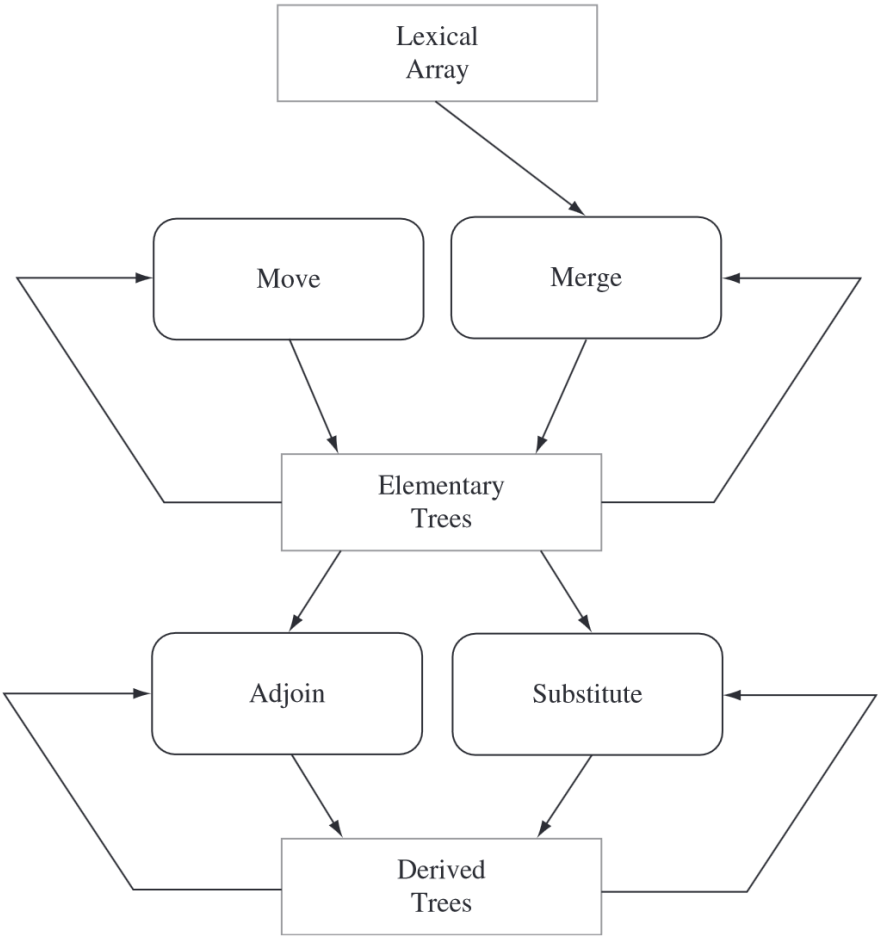
The last possibility in the quotation above must be ruled out because the FTH in (152) on page 214 prevents any syntactic object to be remerged across different elementary trees. As far as intermediate steps of (the same type of successive-cyclic) movement are concerned, in section 5.4.4 it will be argued that, in TAGs, they represent a problem one can do without. Still, this does not mean the complete ban of movement within a single elementary tree. On the contrary, it suggests that each type of movement (say, A- as opposed to A′-movement) can occur only once in the elementary tree.

This in turn hints at an understanding of *singulary transformations* where the singularity is not only related to the structure whose elements are manipulated (Frank 2002: 4, Krivochen 2019: 53 and references therein), but is also concerned with the admitted occurrence of the manipulation itself. Within an elementary tree, an argument can thus be displaced because its non-projected non-terminal DP undergoes Internal Merge and reaches the target position in one fell swoop, as in Frank (2006: 154-155). With these premises in mind, each substitution illustrated in the previous sections simulates the effects of the Merge operation(s) by filing in "one of these argument slots" (Frank 2006: 153).

To summarize, the reconciliation between Minimalism and a TAG approach *à la* Frank (2002, 2004, 2006, 2010, 2013) proceeds by assuming that the Merge and Agree operations take care of the construction of elementary trees, whereas the combination among such obtained objects is handled by means of substitutions and adjoining operations (illustrated in sections 5.1.1 and 5.1.2, respectively). This gives rise to the schema in (168) on the facing page, which is taken from Frank (2002: 24, 2006: 151, 2013: 241) and boils down to "two distinct stages" (Frank 2013: 240). As Frank (2002: 12) observes, the first one corresponds to the derivational path assumed in Chomsky (1993, 1995). As section 5.3.2 will discuss, works like Uriagereka (1999, 2012) and Heck (2016, 2023) provide the basis for arguing that the various elementary trees are derived in a parallel way in the first stage of the schema in (168), so that they can be combined with one another in a subsequent step by means of the

TAG operations in the second stage.

(168)



Given that the two stages in (168) are pretty clearly distinct, one may wonder if it is more appropriate not to include TAG operations in the narrow syntax. The answer to this question must be negative. As will be shown from section 5.3 on, TAG operations aim to provide an adequate account for facts that are concerned with both the phonetic realization of grammatical sentences, as well as with their related semantic interpretation. As a consequence, it would be wrong to say that TAG operations take place after Spell-Out and are thus not narrow-syntactic. Thus, TAG operations neatly fit in a view of narrow syntax that broadly constists "of those operations that build trees" (Safir 2008: 331).

One may still argue that yes, TAG operations are narrow-syntactic, but as soon as one infers from (168) that they compactly follow all the Merge operations, substitutions and adjoining operations can be assumed to occur in a separate component of their own. In order to assess whether it would be profitable to claim something like this, one can formulate the

following thought experiment. Let us assume a hypothetical scenario where External Merge (EM) and Internal Merge (IM) are understood as they are in real contexts, but that there is constantly growing evidence that in each derivation all IM operations take place after all EM operations have been performed.

As a consequence, no debate would emerge with respect to Merge-over-Move (Chomsky 2014–1995: 320 ff., Castillo, Drury & Grohmann 2009–1999), but it is likely that someone would propose that, if that is the case, then it is better to argue that EM and IM belong to distinct modules, although they have many common properties that make them likely to be both narrow-syntactic. One may further complicate this hypothetical scenario by imagining that such an alternative is not completely adequate either, as some scholars observe that there are some (perhaps fringe) cases where some IM operation is mingled with EM operations, so that settling the matter becomes a difficult task.

As absurd as it may sound, the scenario just envisioned is not that far away from some real debates, as soon as the appropriate conversions are performed with due care. In fact, the EM and IM operations of the thought experiment above can be taken as an analogy not only of the Merge and TAG operations in (168) on the previous page, but also of the debate on the morphology-syntax interface, of which Siddiqi (2014) provides a good summary (though see Wiltschko 2014: Sec. 2.4 for related discussion). For a start, Siddiqi (2014: 346) dubs as "Anti-Lexicalist" those approaches which "assume that the syntax, in one way or another, is responsible for building complex words as well as phrases". As long as the Anti-Lexicalist position is limited to this statement, there is no doubt that the schema in (168) on the preceding page adheres to a transposed version of it, as both elementary and derived trees are deemed to be built by the syntax.

However, the original Anti-Lexicalist position is often associated with two other assumptions which are at odds with TAGs. The first is that TAGs continue the "lexicalist tradition of the Principles and Parameters model" (Putnam 2020: 3), whereas the Anti-Lexicalist approaches tend to be *exo-skeletal*. As Putnam (2020: 4) clarifies, this means that "features are attributed to functional heads in the syntax rather than individual lexical entries". Consequently, the lexicon is seen as "(largely) void of semantic and phonological content" and the lexical items "are stored external to the syntax, resulting in a lexicon that is post-syntactic (i.e., derived by means of computation rather than being previously established)" (Putnam 2020: 4).

The second problematic assumption of many Anti-Lexicalist models is that the whole unique morphosyntactic component is ruled by a single fundamental process, i.e. Merge (Siddiqi 2014: 356). The application of this view to (168) appears to be in conflict with what Frank & Hunter (2021: Sec. 1) point out. On the one hand, Merge obeys the Extension Condition (see Chomsky 2014–1995: 227 ff.), which states that the structure is always built at its root by involving the nodes that the latter immediately dominates, while on the other hand both substitution and adjoining violate such a statement by being capable of also rewriting other nodes of the same structure. As long as one "understand[s] the intent of the E[xtension] C[ondition] to be that of limiting changes to already constructed structures" (Kobele 2023: 80), any TAG formalism is thus inherently counter-cyclic overall, even though it can encompass a derivational stage like the upper half of (168) which, for its part, complies with the Extension Condition.

The opposite opinion to Frank & Hunter (2021: Sec. 1) would consist in asserting that the Merge and the TAG operations can be regarded as the same syntactic process applying at different levels of complexity. This can be done by following Kobele (2023), where it is argued that cyclicity (as a formal mechanism) and the application of the Extension Condition are two separate aspects that do not have to overlap.

As counterintuitive as it may seem, this assertion can be put to test by wondering if the two stages proposed by Frank (2002: 24, 2006: 151, 2013: 241) in the schema in (168) are strongly or weakly compartimentalized. Such a doubt is more akin to the divide between the two major lexicalist views (Siddiqi 2014: 346), and it boils down to decide whether the TAG operations can be interspersed among the Merge operations or not. Frank supports a strong compartimentalization in his works, as the elementary trees are only combined with one another after their formation is completed in accordance with the following rule, whose formulations in (169a) and (169b) can be found in Frank (2002: 126) and Frank (2006: 176), respectively.

(169) Maximal Checking Principle (MCP):

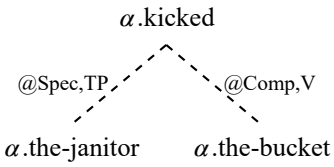
a. The output of an elementary tree derivation from numeration N must contain as few uninterpretable features as possible. [In other words:]
b. Check as many features as possible within an elementary tree.

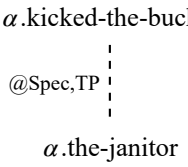
The interaction between the MCP in (169) and strong compartimentalization leads to the

following two consequences. On the one hand, wherenever the circumstances allow some element to be remerged within an elementary tree, movement should take place; ditto for agreement operations. On the other hand, an initial tree can safely rewrite the head of the movement chain because the latter was formed in a previous derivational step, so that there is a basis to prevent the initial tree from substituting elsewhere.

On the contrary, DPs are always combined in their *in situ* positions if one assumes a weak compartimentalization that allows initial trees to substitute as soon as they can in the usually assumed bottom-up derivation of elementary structures (Brattico & Chesi 2020: 2, Kobele 2023: Sec. 3). However, this option raises a problem to any to-be-displaced non-terminal DP, as it would not be possible to postpone the substitution, unless one assumes that non-terminal DPs somehow know whether they will be left *in situ* or not. Given that the latter option raises in turn a look-ahead problem, a weak compartimentalization approach should thus be forced to argue that, once an initial tree has been substituted, it can be remerged within the targeted elementary tree in the same way in which the DP non-terminal would in the strong counterpart illustrated above following Frank (2006: 154-155). As a result, a weakly compartimentalized TAG derivation would inherit the common pros and cons related to the Minimalist application of IM (see a.o. Witkoś 2021, Gärtner 2022, Krivochen 2023d).

The adoption of the aforementioned kind of weak compartimentalization would also compromise the usual treatment of idioms in TAGs, discussed in Frank (2013: 242) and Abeillé & Rambow (2000: 26-27), among others. In fact, the twofold interpretation of a sentence like *The janitor kicked the bucket* is accounted for in the following way. On the one hand, the literal reading emerges when "an elementary tree headed by the verb *kicked* is targeted by substitution of two DP elementary trees, associated with the lexical material *the janitor* and *the bucket* (Frank 2013: 242, italics in the original). On the other hand, the idiomatic reading follows from "a single substitution operation: the DP *the janitor* substitutes into a complex elementary tree headed by the non-compositionally interpreted material *kicked the bucket*" (Frank 2013: 242, italics in the original). This observation anticipates something that would be discussed in section 5.2.1, namely that in TAGs it is often argued that different semantic interpretations are the result of different derivation trees. The literal and the idiomatic meanings are thus deemed to follow from (170a) and (170b), respectively.

(170) a. 

b. 

It can be argued that the dichotomous account just shown in (170) should be given up by a weak compartimentalization that allows initial trees to substitute as soon as they can. In fact, under such an assumption it would become less clear how to distinguish between a built-in (and hence idiomatic) argument from a substituted one (whose interpretation is consequently literal). This does not mean, however, that one is left with no possible alternate explanation for idioms. One could decide, for instance, to benefit from the analysis given by Bruening, Dinh & Kim (2018) in this respect (and discussed in section 2.2.4), or to rely on the one provided by Lichte & Kallmeyer (2016).

All things considered, the two aforementioned problems for weak compartimentalization do not seem to be insurmountable (even though the strong alternative appears to be preferable). From now on, it is possible that the readers will have the feeling that the derivations to come hint at one view or the other in a (perhaps) indecisive manner, with the exception of those representations that explicitly follow the assumptions of the authors contextually mentioned. As was already the case for (164) on page 226, the choice of (not) showing certain substitutions must simply be understood as a consequence of expository purposes, and not as some committed endorsement of either one of the two options. In absence of compelling arguments with respect to either a strong or weak compartimentalization between the Merge and the substitution operations, the ultimate decision in this respect is left to the readers themselves.

As far as substitution is concerned, both the strong and the weak compartimentalization of the schema in (168) on page 231 must agree on one point, namely that the derivation of the targeting initial tree must have been completed before it substitutes. Otherwise, its root could not properly match with the node in the targeted elementary tree. Interestingly, the same assumption must be made with respect to the derivation of each auxiliary tree, as the alternate take would lead to the violations of the Markovian property discussed in section 5.1.4. These remarks thus bear a twofold consequence. First, the burden of the comparison being made between strong and weak compartimentalization of (168) depends entirely on the derivation of the elementary tree to be targeted by some TAG operation(s). Under the strong alternative, its derivation by means of Merge and Agree operations must be completed before it can access

the second stage of the schema in (168), whereas under the weak alternative it does not have to be.

The latter point leads to the second aforementioned consequence, which arguably makes the strong counterpart less appealing than the weak one. In fact, a first advantage of the latter alternative is that it allows to build the derivation tree from the bottom up, just like each single structure is commonly assumed to be in Minimalism (see Brattico & Chesi 2020: 2, Kobele 2023: Sec. 3 and references therein). This can be seen by taking as an example a derivation tree where two unspecified elementary trees γ_1 and γ_2 combine to a root α.[136]

A twofold implication descends from the assumption of a weak compartimentalization where both γ_1 and γ_2 must be derived completely (from the bottom up) in order to effectively combine with the structure under construction that will later become α. On the one hand, this means that there is a point of the derivation where γ_1 and γ_2 are legitimate nodes of a derivation tree while the structure that will become α is not (whence the bottom-up building of the derivation tree).[137]

On the other hand, if γ_1 or γ_2 are (well-formed) auxiliary trees that must adjoin as soon as they can, then their adjoining also complies with the Extension Condition mentioned above (see Chomsky 2014–1995: 227 ff. Safir 2008: 331, 337-338, Kobele 2023 and Frank & Hunter 2021: Sec. 1). In fact, under such an approach each auxiliary tree would target the current root of a structure under construction, so that rule ordering can "be abandoned: rules apply in a (strict) cyclic manner whenever their structural descriptions are met" (see Kobele 2023: 74 and references therein). Given however the still ongoing first-stage derivation of the latter, the said targeted node eventually becomes an internal non-terminal node different from the root of the ultimately obtained elementary tree.

In order to assess if and how the scenario just envisioned would be problematic for a TAG approach *à la* Frank, it is better to shift the focus towards his implementations of feature checking. In fact, although the Maximal Checking Principle in (169) on page 233 forces agreement to be "local to an elementary tree" (Frank 2006: 190), it does not prevent certain features from remaining "unchecked within the confines of an elementary tree" Frank (2002:

136 The readers are invited to recall that the latter must be an initial tree because of (159a) on page 221, i.e. the derivational completeness assumed by Frank (2013: 244).

137 Under such a view, the nodes of the derivation tree would be comparable to the result of Spell-Out as per Uriagereka (1999), which "is no longer a syntactic structure, but rather an object analogous to a lexical item. The internal properties of such an object are inaccessible to subsequent syntactic derivation, but this object may itself be incorporated into a subsequent syntactic derivation in a manner analogous to other lexical items" (Frank 2010: 374).

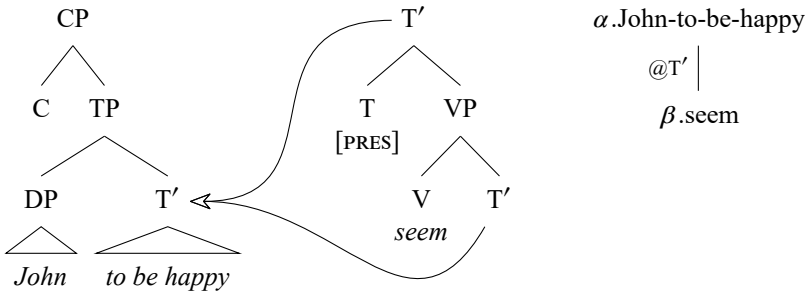
153). As a consequence, such features must eventually be checked by having the elementary tree bearing them combined with another one.

By following Frank (2002: Sec. 4.6, 2006: 195 ff.), one can take the sentence *John seems to be happy*, i.e. an example of subject-to-subject raising, to illustrate how feature checking can work in a TAG approach, and which insights it can offer with respect to the schema in (168). Instead of assuming the analysis in (171), where the subject selected by the embedded infinitive A-moves to the subject position (i.e. Spec,TP) of the matrix clause (see Heck 2016: Ch. 6, Diercks, Santorini & Kroch 2025: Ch. 11 and references therein), Frank (2002, 2006) proposes a TAG account where the overt realization of the subject occupies Spec,TP within the elementary tree of the embedded infinitive (see also Longenbaugh 2014: Sec. 1.6). The latter is in turn shoehorned by an auxiliary tree recursive at T′ and anchored by the tensed raising verb *seem*. The structure of the latter tree is determined by the fact that "raising predicates like *seem* are distinguished by their failure to assign a θ-role to their subject.

Hence, the θ-Criterion does not require the existence of a subject position in an elementary tree headed by such a predicate" (Frank 2002: 70-71, italics in the original). As a result, the crucial difference between the purely derivational analysis in (171) and the TAG one in (172) is that (A-)movement is an indispensable assumption in the former, whereas it is entirely optional in the latter.

(171) John seems ~~John~~ to be happy.

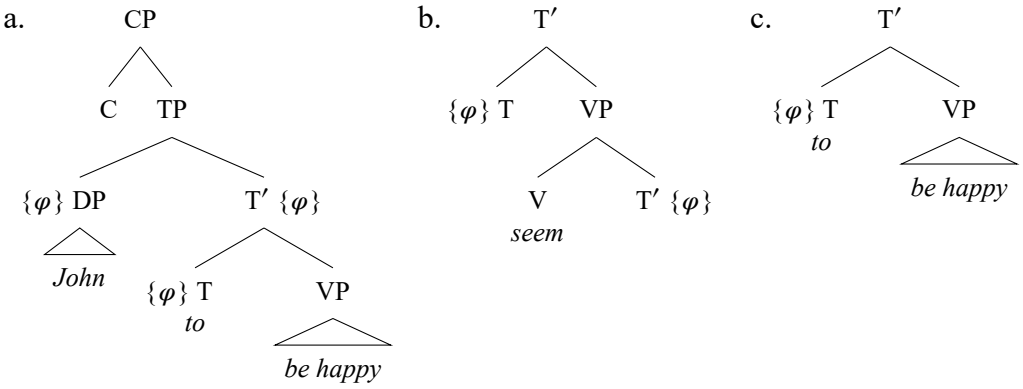
(172)



For the present concerns, it is important to evaluate whether *β.seem* could legitimately adjoin as soon as T′ is projected in the targeted structure, prior to the merger of Spec,TP and the C head. As long as agreement operates in the way discussed in Frank (2006: Sec. 5), the answer must be negative. In fact, the core of such a proposal is that "what appears to be agreement between" two elements belonging to distinct elementary trees "is in fact an

indirect relation, consisting of two separate instances of agreement, each encoded within the local domain of an elementary tree" (Frank 2006: 186).[138] Thus, according to Frank (2006: 195) the two trees in (172) can combine because in *α.John-to-be-happy* "the specifier-head relation holding between the DP and the T head" makes the pertinent features percolate to the T′ node in the same projection, whereas in *β.seem* "there is an agreement relation holding between the T head and the raising predicate's T′ complement" involving the same features. These details of the two trees are respectively reported in (173a) and (173b) by adapting what Frank (2006: 195) shows (see also Longenbaugh 2014: 16).

(173)



In order for agreement to be passed on when an auxiliary tree adjoins to another elementary tree, the relevant feature(s) must be present on the node that is shoehorned, either because it is a probe or a goal in a pertinent Agree relation or because it is a member of the projection within which feature percolation can take place. By following these indications from Frank (2006: 190), it can be inferred that (173b) can only adjoin to (173a), thus favoring the strong compartimentalization of the schema in (168) on page 231. On the contrary, the assumptions in Frank (2006: Sec. 5) are not compatible with a weak compartimentalization that would account for (172) by having (173b) combined to (173c). In fact, the latter bears no adequate target node because the T head in (173c) has not entered an Agree relation yet, and hence no feature percolates at T′ as required.

However, Frank (2006: 195) admits that the aforementioned mechanism differs from how Frank (2002: Sec. 4.6) accounts for raising constructions like (172). The latter alternative consists in two main ingredients. The first is that adjoining is prohibited if it "does not achieve any reduction in the number of uninterpretable features present in an elementary tree" (Frank

[138] Actually, such a mechanism can further span across several elementary trees, provided that each one of the elementary trees involved "include the trigger of agreement [...] and a head on which agreement is borne" (Frank 2006: 190).

2002: 154), whereas the second consists in arguing that "inter-elementary-tree feature checking may take place only under a specifier-head configuration" (Frank 2002: 161). If such a configuration (discussed in section 3.1.3) were really sufficient for checking the relevant EPP- and φ-features across the two elementary trees in (172), then there would be no compelling reason for rejecting the adjoining of (173b) at the T′ node of (173c). The choice of this alternative would not only make adjoining compatible with both the Extension Condition and the weak compartimentalization of (168), but it would also provide an independent explanation for the two assumptions that Frank (2002: 154, 161) must stipulate. In fact, both the reduction in the number of uninterpretable features and their checking under a Spec-Head configuration would be nothing but a consequence of the subsequent merger of the DP *John* together with the other typical selectional requirements assumed in Chomsky (2000) for the first stage of the schema in (168), as acknowledged by Frank (2002: 159) himself.

Two remarks must be made before concluding this section. On the one hand, it was no coincidence to choose a raising construction as an illustratory example. In fact, raising is both problematic and challenging for a TAG approach *à la* Frank, and many of such issues have not been mentioned yet, but will be displayed in due course. On the other hand, the discussion conducted so far wanted to illustrate that the model for TAGs shown in (168) following Frank (2002: 24, 2006: 151, 2013: 241) is both restrictive and open to different perspectives on the nature of syntactic derivations.

5.2 L-TAGs and U20

The TAGs discussed so far can be addressed more specifically as *L-TAGs* (or, more rarely, *pure TAGs*, as in Bleam 2000, Longenbaugh 2014). This is done not only because of their lexicalization property, but also in order to distinguish them from other variants aimed at increasing their weak or strong generative powers (see a.o. Joshi 2003: 245 ff.). Therefore, *TAG* will be used from now on whenever a certain matter concerns any kind of TAG, while *L-TAG* will specifically address the formalism as it has been presented in the previous sections.

Section 5.2.1 will show a major problem of L-TAGs *à la* Frank (2002, 2004, 2006, 2010, 2013), namely that the latter produce structural descriptions that are syntactically adequate but that fail at capturing some desiderata with respect to the semantics. The related discussion will however introduce some theoretical tools that are nonetheless highly relevant. In fact,

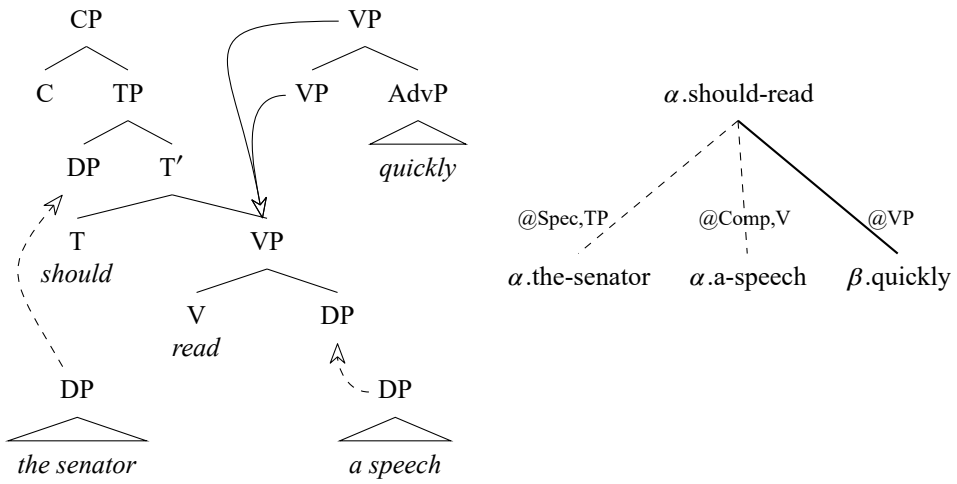
the upcoming distinctions among auxiliary trees are not only sufficient to provide a novel way of analyzing multiple adjectives modifying a single noun, but also foster two L-TAG adaptations of previous analyses of U20, which will be respectively discussed in section 5.2.2 and section 5.2.3.

As opposed to the original analyses illustrated in section 3.2, the L-TAG adaptations either dispense with syntactic movement or require it in a way that is not successive-cyclic. However, the aforementioned semantic inconsistency of L-TAGs becomes even more evident when dealing with a matter like Greenberg's U20. Besides, such two L-TAG analyses also fail to capture the insights that emerged in chapter 4.

5.2.1 The interface between TAGs and semantics

The discussion in section 5.1.2 presented adjoining as the key operation to generate recursively embedded structures in TAGs, so that it becomes possible to derive a sentence like *What do you think the senator should read?*. Although this is a fundamental aspect of the approach, one should also consider that the same operation is usually adopted to account for adjuncts as they are often understood in the generative tradition, i.e. as some piece of structure that can, but need not be, part of a derivation. Thus, one can elaborate on the derivation in (150) on page 214 and represent the sentence *The senator should read a speech quickly* as shown in (174). The interested readers can further check the examples in Frank (2006: 154, 2013: 231).

(174)



There are a few pitfalls descending from the fact that adjuncts are auxiliary trees adjoined to another tree, starting from a terminological one. In the TAG literature, adjoining is al-

ternatively referred to as *adjunction* (as in Joshi, Levy & Takahashi 1975, Chiang & Scheffler 2008, Kallmeyer 2009, Kuhlmann & Satta 2012), even though "the range of structures derivable with adjoining is broader than those involving traditional adjunction" (Frank 2006: 156), like the main clause in (155) and (156).[139] This fact might lead to the formulation of a desideratum that is seemingly both intuitive and appealing, which consists in having a bijective relation between substitution and argument selection on the one hand and between adjoining and adjuncts on the other hand.

However, TAGs akin to the one proposed by Frank (2002, 2004, 2006, 2010, 2013) are notorious for not being able to establish it (Kallmeyer & Kuhlmann 2012: 108, Lichte 2015: Sec. 5.3.4). As the adjunction discussed in section 5.1.2 suggests, this is due to the fact that "not all complementation can be covered by substitution, so that some auxiliary trees do not modify the trees they adjoin to, but take them as their complements" (Scheffler 2003: Sec. 3.1). By hinging on the technical aspects to be introduced from section 5.3.1 on, a path towards a new solution that is closely related to what has just been clarified will be sketched in section 5.4.

The aforementioned issue has been tackled in the TAG literature by means of the introduction of a dichotomy between *modifier* and *predicative* trees, which Schabes & Shieber (1994: 96 fn. 8) formulate by following the similar distinction between *complement* and *athematic* trees first proposed by Kroch (1989: 70 ff.). Such a divide proves itself capable of cutting across the contrast between initial and auxiliary trees by establishing two subsets for the latter type of trees (Candito & Kahane 1998: Sec. 2). On the one hand, it becomes possible to posit modifier auxiliary trees like β*.quickly* in (174), which lexicalize adjuncts. On the other hand, each predicative auxiliary tree can be recognized for the fact that "the foot node, like the substitution nodes, corresponds to an argument node selected by its lexical anchor" (Gardent & Narayan 2015: 45).

Candito & Kahane (1998: Sec. 2) further observe that all initial trees, like α*.the-senator*, α*.should-read* and α*.a speech* in (174), can be regarded as predicative. In fact, in both the initial trees and the predicative auxiliary ones the root belongs to the extended projection of a lexical anchor, in compliance with CETM in (153) on page 214. On the contrary, in any mod-

[139] Readers who are familiar with Government and Binding (Chomsky 1981) may observe that the substitution operation assumed by TAGs could have raised similar concerns back then, but that now it does not because the term *substitution* "is no longer in general use" among generative approaches (Barrie & Mathieu 2014: 138). For related observations, see Barrie & Mathieu (2014: 138-139, 143-145), Wiltschko (2014: 86-87) and references therein.

ifier tree both the root and the foot nodes are nothing but "an open slot corresponding to the phrase being modified" (Gardent & Narayan 2015: 45). The tree *β.quickly* in (174) illustrates this clearly, as it is labeled VP because its content modifies the main lexical verb. In order to make this insight compatible with a Minimalist approach, Frank (2006: 154) argues that Merge can handle *θ*-marking (see again section 5.1.5) just as easily as *θ-identification*, which is a second semantic combination posited by Higginbotham (1985: 564-566) to account for instances of modification.

The present section aims to discuss two aspects that are fostered by the dichotomy that Schabes & Shieber (1992, 1994) propose, namely the structural relations that the foot node entertains with the other non-terminals of the auxiliary tree it belongs to on the one hand, and the theoretical significance of derivation trees as they have just been presented in section 5.1.3 on the other hand. By starting from the former, one can follow Nesson (2009: 4) and Kuhlmann & Satta (2012: 623) and use *spine* to refer to the unique path from root to foot in an auxiliary tree (bearing in mind, however, that the meaning of the term in this context has nothing to do with what has been said in sections 2.3.2 and 2.3.3 while discussing FOFC). One can thus follow Krivochen & Padovan (2021: 3) in observing that, in (174), *β.quickly* has a spine formed by only two nodes in a parent-child relationship, whereas the auxiliary trees discussed in sections 5.1.2 and 5.1.5 have longer spines that also feature other non-terminals.[140]

By starting from the properties of a spine, a further distinction can be formulated. If an auxiliary tree features overt lexical material on both sides of the foot node, then it can be called a *wrapping tree* (Kuhlmann & Satta 2012: 628). Although each member in the complement of the set of wrapping trees can be identified as *non-wrapping*, such a term is hard to find in the TAG literature. This is due to Schabes & Waters (1995: 481) and related works, where non-wrapping trees are further split into two subsets. On the one hand, in any *left* auxiliary tree "every nonempty frontier node is to the left of the foot" (Schabes & Waters 1995: 481), so that the latter must be the rightmost non-terminal in the related spine.

Similarly, each auxiliary tree whose foot is leftmost member of the spine is defined as a *right* auxiliary tree (Schabes & Waters 1995: 481).[141] Thus, an auxiliary tree like *β.quickly* in (174) is a (non-wrapping) right auxiliary tree, and its binary branching nature forces its

[140] Moreover, the auxiliary tree in (174) does not show the shortest spine possible, as section 5.3.1 will introduce a type of auxiliary trees whose spines collapse into single nodes.

[141] Actually, Joshi (2003: 254) points out that there can be nodes on the right (or the left) of the foot node in a left (right) auxiliary tree, "as long as their yield is empty".

spine to contain no non-terminal which is distinct from its root and foot nodes (the latter being the leftmost node of the tree).[142]

An instance of (non-wrapping) left auxiliary tree has been previously shown in (172) on page 237 when illustrating a TAG account of subject-to-subject raising, and such a phenomenon can be reproposed here in order to show the distinction between wrapping and non-wrapping trees. In fact, Storoshenko (2006) argues that *β.seem* in (172) is an oversimplification that does not manage to provide a fully satisfactory account of raising verbs in English, as the matter becomes more complex whenever a raising verb selects an experiencer, as in *John seems* to Mary *to be happy*. Storoshenko (2006: 162) claims that English raising verbs require a VP-shell analysis *à la* Larson (1988, 2014) by highlighting the three contrasts in (175)-(177) that are consistent with Larson's c-command tests.[143]

(175) a. John seems to nobody to like anything. [NPI licensing]
b. * John seems to like anything to nobody.

(176) a. John seems to every boy_i to like him_i. [Variable binding]
b. John seems to like him_{*i} to every boy_i.

(177) a. * What does John seem to whom to like ~~what~~? [Superiority]
b. What does John seem to like ~~what~~ to whom?

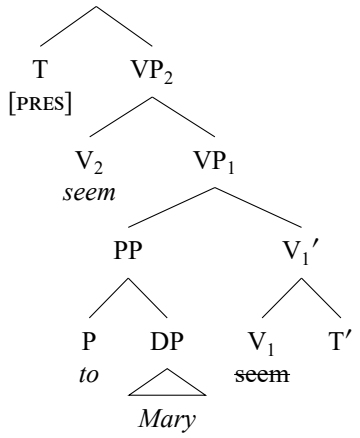
For all pairs in (175)-(177), the first alternatives show that the experiencer θ-marked by the raising verb c-commands the object of the mono-transitive verb *like*, whereas the second options illustrate that the same experiencer fails to do so as long as it is in the final position. Storoshenko (2006: 162) thus concludes that the VP-shell in (178a) is a more detailed counterpart of the non-wrapping left auxiliary tree in (172), because it manages to account

[142] If instead an auxiliary tree is n-ary branching with $n \geq 3$, it can be either wrapping or non-wrapping as soon as it is not zero-dimensional, namely whenever its root and foot nodes do not coincide (see section 5.3.1). Thus, one can argue that zero-dimensional auxiliary trees (where root and foot nodes are necessarily coincident) are the only ones that must always be of the non-wrapping type.

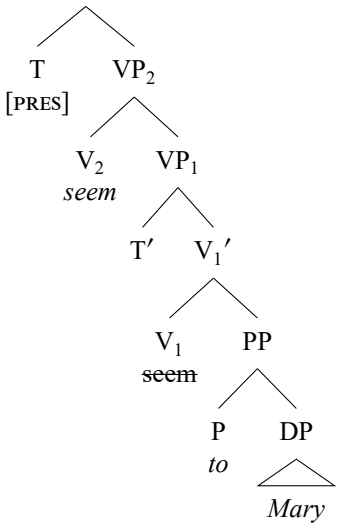
[143] The contrast in (175) focuses on the capacity of *nobody*, i.e. an example of negative quantifier, to license under c-command *anybody*, i.e. an instance of Negative Polarity Item (whence NPI—and its licensing. See a.o. Zeijlstra 2007: Sec. 3). On the variable binding in (176), see Bruening (2021a: Sec. 2.1) and references therein. The contrast in (177) is instead determined by the Superiority Condition, which establishes "that when a rule can apply to more than one element in a given structure (because those elements are similar in some respect), it must pick the element that is structurally higher [...] (that is to say, the one that is closer [to the c-commanding structural position that enforces said rule])" (Zanuttini, Leu & Kayne 2013: Sec. 4.1). In Minimalism, this phenomenon can be accounted for by means of two closely related principles that nonetheless have different theoretical motivations and scopes of application, namely the Minimal Link Condition (MLC, see a.o. Heck 2016: Sec. 2.2 and references therein) or the Attract Closest Condition (ACC, see a.o. Frank 2006: 147 ff. and references therein).

for (175a), (176a), (177a). The opposing behavior of (175b), (176b), (177b) can instead be captured by assuming the wrapping auxiliary in (178b), where the experiencer does not c-command the T′ foot in the same structure. As a consequence, the experiencer also fails to c-command whatever is dominated by T′ once the auxiliary tree is adjoined to *α*.*John-to-be-happy*.

(178) a.



b.



An interesting fact to observe is that the analysis given by Storoshenko (2006) is also compatible with an alternate principle to c-command, namely the "precede-and-command" approach favored by Bruening (2014, 2018b) and references therein. By following Bruening (2014: 343), the experiencer in the sentences in (175)-(177) manages to precede-and-command the object of *like* only if the conjunction of two requirements is verified, namely that the experiencer both linearly precedes and phase-commands such an object. *Phase-command* entails that "there is no ZP, ZP a phasal node, such that ZP dominates" the experiencer but does not dominate the object (Bruening 2014: 343).

Although Bruening (2014: 343) takes *v*P as one of the three possible phases, i.e. CP, *v*P and NP (the last one representing an endoresement of the NP-hypothesis—see sections 2.2.3 and 2.2.4), Frank (2006: 154-155 fn. 6) argues that *v*P cannot be phasal in a TAG approach (see again page 230) and therefore no problem for the trees in (178) arises even if there is a *v*P in the projection of *like*.[144] It is rather the linear precedence requirement which makes

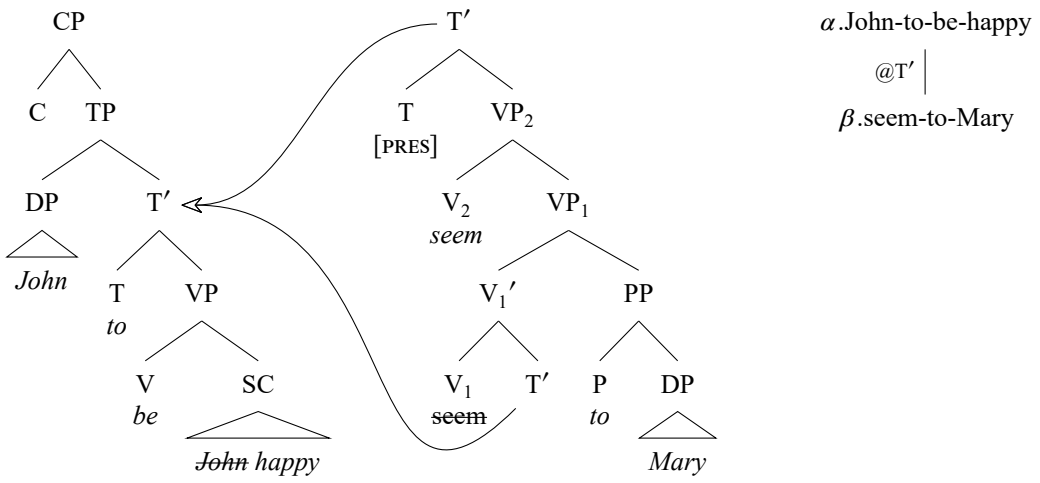
[144] It is also worth noting that Heck (2016: 117) further claims that "the little v accompanying raising verbs [...] is a phase head". The observations made in the main text with respect to Bruening (2014: 343) also apply to Heck (2016: 117), as the latter needs to add this assumption to provide an adequate account for raising under his alternate approach, which will be discussed in section 5.3.2.

the experiencer in (178a) succeed, and the one in (178b) fail, to precede-and-command the object of *like* in (175)-(177).

Thus, a TAG approach that follows Storoshenko (2006) has the potential to make linear precedence correlate with a structural distinction between the two auxiliary trees in (178), namely the non-wrapping nature of (178a) against the wrapping one of (178b). The difference between (178a) and (178b) ultimately consists in swapping PP and the T′ foot as specifier and complement of V_1. Although this allows to capture the facts in (175)-(177), it remains entirely stipulative and counters the Uniform Theta Assignment Hypothesis (see a.o. Diercks, Santorini & Kroch 2025: Sec. 7.6.2 and references therein).

A precede-and-command approach thus allows to posit an alternative that arguably performs better in these respects, namely a wrapping counterpart of (178a) that maintains PP and T′ as specifier and complement of V_1, respectively. This is shown below in (179), which derives *John seems to be happy to Mary*. In (179), the A-movement of *John* is assumed even though not strictly required by the TAG formalism (see section 5.1.5). Furthermore, the latter DP originates within SC, namely a small clause that Moro (2000: 42) and related works posit to better analyze copular constructions.

(179)



The discussion of subject-to-subject raising can go on by following Heck (2016: Sec. 6.1), who observes that raising across an experiencer is permitted in English (as shown by the aforementioned sentences), but that the same is not allowed in Italian, French and Icelandic. The examples in (180) and (181) sharpen the contrasts provided by Heck (2016: 113-115) and references therein.

(180) a. *Gianni sembra essere intelligente.* [Italian]
Gianni seems to.be intelligent

b. *Jean semble être intelligent.* [French]
Jean seems to.be intelligent

c. *Ólafur virðist vera gáfaður.* [Icelandic]
Olaf.NOM seems to.be intelligent
'John seems to be intelligent.'

(181) a. **Gianni sembra a me essere intelligente.* [Italian]
Gianni seems to me to.be intelligent

b. **Jean semble à moi être intelligent.* [French]
Jean seems to me to.be intelligent

c. **Ólafur virðist mér vera gáfaður.* [Icelandic]
Olaf.NOM seems 1SG.DAT to.be intelligent
'John seems to me to be intelligent.'

TAGs can capture the different behavior of English as opposed to Italian, French and Icelandic with respect to (181) by assuming that the auxiliary tree proposed in (178a) does not have a comparable counterpart in the latter three languages. This goes to illustrate what has been anticipated in the introduction of the present chapter, namely that in TAGs "the set of elementary trees in the grammar of a language is the only locus of cross-linguistic differences" (Frank 2013: 238). However, an exact illustration of the thus modified auxiliary trees goes beyond the scope of the present work. This is due to the fact that they should prove to adequately account for all the data that Heck (2016: Ch. 6) illustrates and captures by means of a derivational analysis that section 5.3.2 will show to deeply affect the TAG approach to be proposed in section 5.4.

As anticipated while discussing idioms in section 5.1.5, derivation trees play an important role when it comes to semantic accounts in TAGs. To understand why and how, it is advisable to repeat that in Vijay-Shanker (1987: Ch. 1) and many subsequent works each non-terminal node of an elementary tree γ can be targeted by at most one elementary tree γ' during a derivation. More specifically, non-terminal leaves can be targeted by initial trees substituting there, whereas internal non-terminals can be shoehorned by auxiliary trees (see again sections 5.1.1 and 5.1.2). Foot nodes stand out for not being a suitable target for any tree (see

however Chen-Main & Joshi 2008 and section 5.4.2 for proposals countering this claim). On the one hand, no auxiliary tree can adjoin there because foot nodes are non-terminal leaves, and "adjunction at leaves is not possible" (Kallmeyer 2010: 55). On the other hand, foot nodes are often singled out for not being substitution nodes by definition (Kallmeyer 2010: 54, Joshi & Schabes 1997: 71).

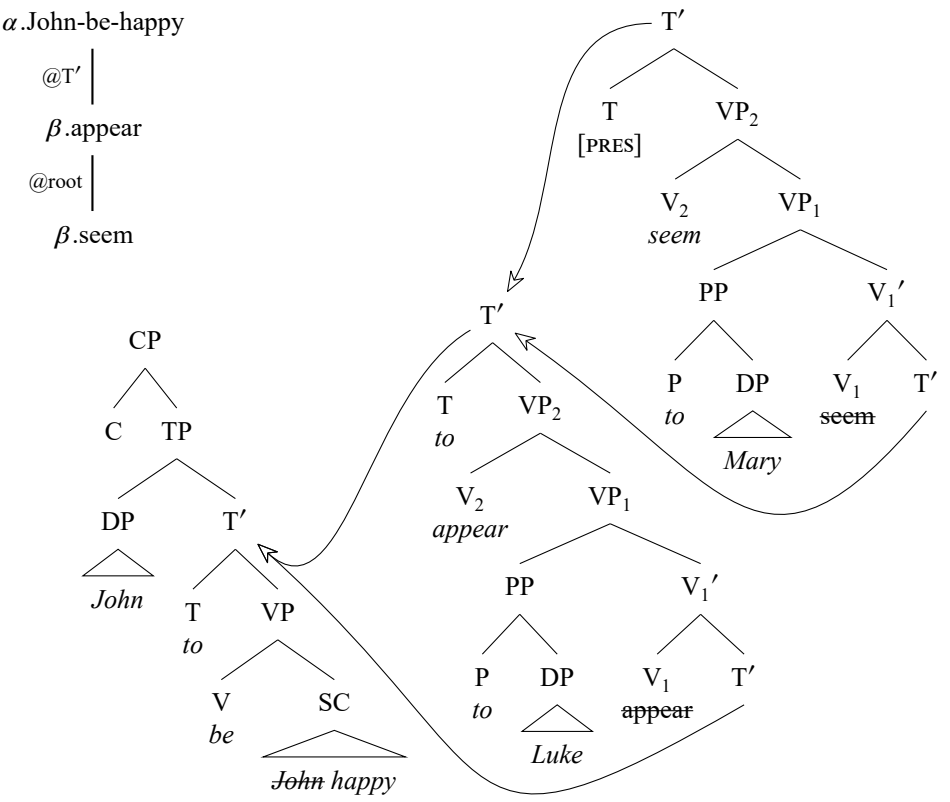
Since the present section focuses on auxiliary trees, one first trivial type of legitimate derivation is the one where, given any two auxiliary trees β_1 and β_2, they adjoin to the same tree γ as long as they shoehorn different nodes of the latter, so that they are both daughters of γ in the corresponding derivation tree. An example of this case (in addition to the one given by Frank 2013: 231-232) would be the sentence *Actually, Mary really thinks you should leave*, with the adverbs *actually* and *really* being contained in β_1 and β_2, respectively.

Another somewhat less trivial consequence of the aforementioned standard notion of derivation is that two auxiliary trees β_1 and β_2 are still allowed to first combine with one another and then target one single node of a third elementary tree γ. To illustrate this other case, it would be sufficient to sketch a derivation where an adverb adjoins to a verb, whose clause then shoehorns another one. However, a more interesting circumstance is the "special case of adjoining" where β_2 adjoins at the root of β_1 (Kallmeyer & Joshi 2003: 13), which in turn adjoins to γ. Such a scenario stands out because it requires β_1 and β_2 to be recursive at the same node on the one hand, while the internal structure of both β_2 and β_1 remains substantially unchanged in the subsequently derived tree on the other hand, even though the former has been combined to the latter. This derivation can be exemplified by the sentence *John seems to Mary to appear to Luke to be happy*, where two raising verbs combine to the same non-finite clause (see also Frank 2002: 72, 79).

As the readers can see, (182) on the next page benefits from what has been discussed so far by featuring the auxiliary tree in (178a) as well as the initial tree in (179). The derivation tree in (182) differs from the ones illustrated before because it does not tell the label of the node at which β.*seem* adjoins, but rather that such a node is the root of the targeted tree. Although the two alternatives can be regarded as substantially equivalent, section 5.4.2 will provide arguments in favor of the latter choice, due to the aforementioned peculiar properties of this type of derivation.

It is also worth discussing that having two (or more) left auxiliary trees bears an interesting consequence with respect to the correspondence between the derivation tree and the derived

(182)



string in (182). In fact, the content of *β.seem* precedes that of *β.appear* in the corresponding string because the former is a daughter of the latter in the derivation tree in (182). Had such auxiliary trees consistently been of the (non-wrapping) right type, the content of *β.seem* would have followed that of *β.appear* in the related string. See page 297 for discussion of such an example.

All the auxiliary trees presented so far with respect to raising constructions suffer from a problem already pointed out at the end of section 5.1.2, namely that they are recursive at a bar-node, which is problematic under Minimalist assumptions. This situation is worsened in (182), where the T heads in the two auxiliary trees have completely different features, as the one in *β.appear* is non-finite whereas the other in *β.seem* is finite. Section 5.4.4 will sketch some refinements that arguably manage to overcome such difficulties.

A couple of questions can be asked in order to introduce the next relevant issue concerning (182). Does the T′ foot indicate that *β.seem* modifies the tense of *β.appear* in a crucial way so as to obtain a grammatical derived tree? Or does it indicate that the anchor *seem* selects a non-finite sentential complement that does not project beyond T? While approaches that are entirely derivational have no choice but to argue for the latter (Diercks, Santorini &

Kroch 2025: Sec. 11.3.3), TAGs allow to choose between the aforementioned alternatives, especially in light of the distinction between modifier and predicative trees introduced by Schabes & Shieber (1992, 1994).

Assuming for a moment that both β*.seem* and β*.appear* in (182) can be considered as of the modifier type, one may subsequently wonder whether they do modify α*.John-be-happy* independently of one another. Although the similarity between the verbs *seem* and *appear* is arguably insidious so as to assess this, the answer to such a question must be eventually negative. This can be better seen by replacing β*.appear* in (182) with another raising verb, namely *be likely* (Gelderen 2017: 95-96, Krivochen 2023a: 142). If the two auxiliary trees did modify the targeted initial tree in an independent manner, the semantic interpretation of the whole sentence should remain unaffected by their structural (and subsequently linear) swapping, contrary to fact. This can be witnessed in (183), where such an operation from (183a) to (183b) must be accompanied by a similar exchange of the T heads in the two auxiliary trees, so as to always have the highest T head in the derived tree be the only finite one. This additional refinement, done to avoid the generation of the ungrammatical sentence in (183c), highlights another shortcoming that is due to the assumption of auxiliary trees that are recursive at T′, i.e. they hinder the convenient swapping of two trees even when the latter is an arguably legitimate operation.

(183) a. John seems to Mary to be likely to be happy.
b. John is likely to seem to Mary to be happy.
c. *John to be likely seems to Mary to be happy.

The contrast between (183a) and (183b) shows that the order of the adjoining operations illustrated in (182) does matter, so that changing the kinship relation among the nodes in the corresponding derivation tree affects the subsequent semantic interpretation of the whole sentence. It is from observations of this kind that, starting from Candito & Kahane (1998), scholars working with TAGs have argued that "[c]ompositional semantics is defined on the derivation tree" (Joshi 2003: 252) and have attempted to compare TAG derivation trees with dependency structures (see Joshi & Rambow 2003, Joshi, Kallmeyer & Romero 2003, Chen-Main & Joshi 2007, 2014, Chen-Main 2012, Kallmeyer & Kuhlmann 2012; see also Kallmeyer & Joshi 2003: 4, Findlay 2023: 2088-2089 for further references). However, the beginning of the present section has already pointed out that such a correspondence is flawed,

as TAGs exhibit cases "where predicates adjoin into their arguments, an analysis that is chosen whenever an argument allows for long extractions" (Kallmeyer & Kuhlmann 2012: 108), as section 5.1.2 did show. See Candito & Kahane (1998: Sec. 4) for further issues that defy this desirable correspondence, which will be a central concern in section 5.4.

Another related problem can be introduced by focusing again on the derivation in (182). The pair of sentences in (183a)-(183b) has proven that any two raising verbs cannot independently modify a main clause, because the order would not matter if they did. However, Schabes & Shieber (1992, 1994) remark that the standardly assumed derivation trees leave no room for such an alternate interpretation. This observation arguably puts a TAG approach *à la* Frank on a par with the cartographic-style syntax (Rizzi & Cinque 2016, Rizzi & Bocci 2017). As Krivochen & Padovan (2021: Sec. 3.1) observe, having a portion of structure dominated by some other portion forces the latter to scope over the former (as per the definition of scope standardly adopted since May 1985), so that comparable issues with respect to semantic interpretation arise in both approaches.

In an effort to solve this kind of problem, Schabes & Shieber (1992, 1994) propose a different TAG formalism (also in comparison to Krivochen & Padovan 2021) which is weakly equivalent to the standard one (Schabes & Shieber 1994: 108) by introducing the aforementioned divide between predicative and modifier auxiliary trees. Schabes & Shieber (1994: 94) establish that trees of the former type always give rise to *dependent* derivations, which essentially correspond to those that are standardly assumed in TAGs. This further implies that, under their approach, the auxiliary trees in (182) and (183) must be of the predicative type because they participate in a derivation of this kind. On the contrary, modifier (auxiliary) trees must be involved in some *independent* derivation and are allowed to target the same node in a different tree γ. Since such trees can adjoin there independently of one another, Chen-Main & Joshi (2008, 2014) call this operation *multiple adjoining*. However, another consequence of this possibility is that the relation between such auxiliary trees is left unspecified in the derivation tree, so that they appear as sister nodes (Gardent & Narayan 2015: 42). It is for the latter reason that Kallmeyer, Osswald & Van Valin (2013: 182) talk about *sister adjunction*, as it "allows to add a new leftmost or rightmost subtree [in the derivation tree] below the node that is the adjunction site".

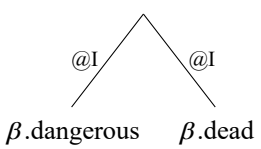
The difference between dependent and independent derivations can be shown by means of an example that has been anticipated in section 3.1.3 and that is often quoted when deal-

ing with the hierarchical relation of multiple adjectives (see a.o. Pereltsvaig 2006: Sec. 2, Kotowski 2016: 129-130, Morzycki 2016: 52-53), i.e. the phrase *a dead dangerous animal*, originally provided by Svenonius (1994: 450-451). As Morzycki (2016: Sec. 2.1) illustrates, such adjectives provide an intersective reading, because the animal they refer to must belong to both the set of dead entities and the set of dangerous ones. Crucially, the inclusion of the animal in either set happens independently, so that the mentioned phrase is felicitous when referring to a lion (that is considered dangerous also when alive) but not to a sheep. Thus, the phrase in (184a) is expected to be the quintessential case of independent derivation by following Schabes & Shieber (1992, 1994).

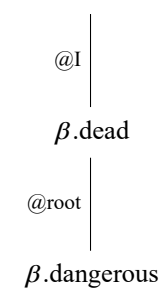
The issue that makes this phrase worth the debate is that the swapping of the two adjectives gives rise to an interpretation that can felicitously apply to a sheep riddled with disease, namely (184b). The observation that in (184b) the danger is linked to the current dead state of the animal suggests that such an option must be the result of a dependent derivation. However, an often marginalized fact is that *a dead dangerous animal* can still refer to a lion, either as a consequence of an intersective reading (184c) or because of the interpretation that can also apply to a sheep in (184b). Thus, the following asymmetry emerges. Under a certain linear order of the adjectives, only the intersective reading arising from an independent derivation is allowed, whereas in the alternate order adjectives can be adjoined in either way and an ambiguous interpretation is possible. The legitimate derivations are shown in (185a)-(185c), where the targeted node addressed as *I* on certain edges refers to an internal non-terminal node of α*.a-animal* that must be c-commanded by the node in which the indefinite determiner is overtly realized.

(184) a. A dead dangerous animal (i.e. a stuffed lion / #sheep)
b. A dangerous dead animal (i.e. an infected sheep / lion)
c. A dangerous dead animal (i.e. a stuffed lion / #sheep)

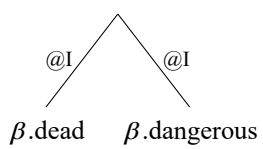
(185) a.



b.



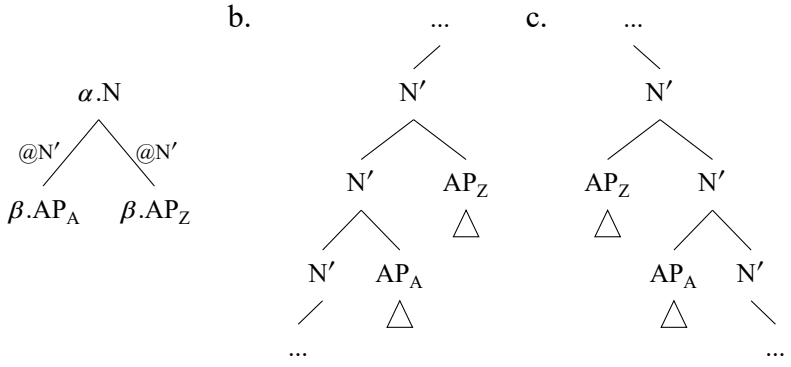
c.



The two structures in (185a) and (185c) illustrate that in independent derivations the order among sisters adjoined to the same node must matter. Otherwise, it would not be possible to ensure the "unique mapping from derivation tree to derived tree" (Gardent & Narayan 2015: 46) which is always guaranteed by dependent derivations (as discussed in section 5.1.3). Therefore, Schabes & Shieber (1994: Sec. 4) posit that trees like (185a) and (185c) are *ordered* derivation trees, so that a total order among sister nodes is established even if they adjoin at the same address in the targeted tree. This rule works in such a way that the leftmost sister in the derivation tree is the lowermost in the derived tree (see Schabes & Shieber 1994: 101, Gardent & Narayan 2015: 46 for detailed illustrations).

Whenever the two or more auxiliary trees involved in an independent derivation are non-wrapping structures, the aforementioned ordering principle bears interesting (but often overlooked) consequences for the eventual mapping to linear order, which will be explored further in section 5.2.2 with respect to Greenberg's Universal 20. If in a certain derivation tree the rightmost sister is of the (non-wrapping) right type, then linear subsequence is determined for its terminal(s) with respect to the one(s) lexicalized by the other sister-adjoined tree(s). Vice versa, if the rightmost sister is a left auxiliary tree, its terminals will eventually precede those mapped from the other sister(s). Thus, it becomes possible for two different grammars to map the two alternate portions of a derived tree represented in (186b) and (186c) from the same independent derivation exemplified in (186a). Even though not explicitly illustrated, it is easy to acknowledge that the order between AP_A and AP_Z would remain the same even if $\beta.AP_A$ were a left auxiliary tree in (186b) or a right auxiliary tree in (186c).

(186) a. b. c.



Since the recursive root and foot nodes must coincide with a projection of the noun (Candito & Kahane 1998: Sec. 2) and all the adjectives in (184) are prenominal, both *β.dangerous* and *β.dead* in (185) are left auxiliary trees, as in (186c). As a consequence, the string which (185a) eventually maps is (184a), where *β.dangerous* is adjoined first and is subsequently

dominated by $\beta.dead$. On the contrary, the independent derivation in (185c) is the unique mapping for (184c).

Although there may well be some combination of intersective adjectives where one order is strongly (if not utterly) preferred (see again section 3.1.3), the alternatives just analyzed show that the aforementioned approach does not face problems of undergeneration. Further implementations can be stated so as to prevent possible overgenerations (see for instance indications in Nesson 2009: 130, Gardent & Narayan 2015: 43), but the details in this respect must be left for future research.

Given that the felicitous reading in (184b) is accounted for by means of the dependent derivation in (185b), two matters must be settled. On the one hand, a rule must be established that bans the alternative where the same auxiliary trees are swapped. Such a possibility must be excluded because it would give rise to the odd interpretation where the dead state of an entity is due to its own dangerous nature. On the other hand, to argue that the elementary trees involved in (185b) are the same of those combined in (185a) and (185c) would contravene a fundamental assumption that Schabes & Shieber (1992, 1994) and Gardent & Narayan (2015) follow, namely that dependent derivations are better suited to predicative trees.

One possible solution to both problems can be introduced by discussing the last type of derivation that Schabes & Shieber (1992, 1994) posit. In dependent derivations standardly assumed for TAGs, a single predicative auxiliary tree can adjoin to an internal node *I*, whereas in the independent derivations that they propose several modifier auxiliary trees can adjoin to a single internal node I'. What if I coincides with I', so that a single node is targeted by a predicative tree as well as some modifier ones? Schabes & Shieber (1992, 1994) call this possibility a *mixed* derivation. If only one modifier tree is involved, a mixed derivation looks like (186a), albeit with one major difference, namely that AP_Z is no longer a modifier tree, but rather a predicative tree.

In fact, Schabes & Shieber (1994: 102) posit the *outermost predication* rule, which establishes that in mixed derivations "predicative trees are always adjoined after any modifier trees at the same address, so that they appear above the modifier trees in the derived tree".[145]

[145] Although both Schabes & Shieber (1994: 103) and Gardent & Narayan (2015: Sec. 2.2) observe that certain TAG analyses can provide counterexamples where a modifier tree arguably adjoins above a predicative tree, sections 5.4.3 and 5.4.4 will show that such cases largely follow from a controversial premise that will be tackled in section 5.4.1, namely the internal structure of the elementary trees which is assumed to be projected. Still, it is worth observing that neither Schabes & Shieber (1994) nor Gardent & Narayan (2015) discuss examples where the predicative auxiliary tree is interspersed among two or more modifier trees in a single mixed derivation.

The latter assumption implies that an approach *à la* Schabes & Shieber (1992, 1994) and Gardent & Narayan (2015) must posit that there exist some adjectives (and, arguably, other kinds of modifiers) which can alternatively lexicalize a modifier tree or anchor a predicative auxiliary tree, while some others belong to either of the two types.

In order to capture the derivations in (184) in a consistent way, it is thus sufficient to say that $\beta.dangerous$ must always be predicative, so that both (185a) and (185c) are mixed derivations. The structure in (184b) is instead a dependent derivation where $\beta.dead$ can (but need not) be the same (type of) elementary tree used for the other two alternate derivations.

The aforementioned issues that (185b) brings to the table could also be tackled by adopting a novel solution, which consists in saying that modifier auxiliary trees can also participate in a dependent derivation. Although discouraged, such an option is not conceptually ruled out Schabes & Shieber (1992, 1994), as they acknowledge that "grammars should forbid adjunction of a modifier tree β_1 at the root of a modifier tree β_2 except where β_1 is intended to modify β_2 directly" (Schabes & Shieber 1994: 108). Given that the latter exception arguably corresponds to the case in (185b), what matters is that the distinction between predicative and modifier auxiliary trees be not undermined. One can obtain this desideratum as soon as the dependent derivation involving the former type of trees is proven to be substantially different from the one involving the latter.

Such a conclusion can be reached by starting from the *strong co-occurrence constraint* discussed by Joshi, Becker & Rambow (2000: 175-177), Chen-Main & Joshi (2008: Sec. 2), and Lichte (2015: 149 ff.), where numbered indexes are used to restrict the possible combinations among auxiliary trees in any dependent derivation (for the formal properties of indexed languages, see references in Lichte 2015: 149). It can be formulated as follows:

(187) Strong co-occurrence constraint (for dependent derivations where all the auxiliary trees involved are of the predicative type): A tree $\beta.A_i$ must select $\beta.A'_{i+1}$ as its argument and adjoin to it, provided that both lexical anchors A of such trees belong to the same syntactic category.

The aforementioned procedure follows the properties that Candito & Kahane (1998: Sec. 2) and Gardent & Narayan (2015: 45) establish for predicative trees, and can safely apply once the following threefold assumption is further made. First, the index of each tree arguably percolates from the lexical anchor to the root. Second, (187) applies by imposing $[i + 1]$ as the index on the foot node of any $\beta.A_i$. Third, such indexes do not affect the matching

between the root and the foot that guarantees the recursion of the auxiliary tree they belong to. In fact, labeling established by projection (Fukui & Narita 2014) must be kept separate from indexing, as the latter can spread across projections, as will be shown shortly (see also footnote 146).

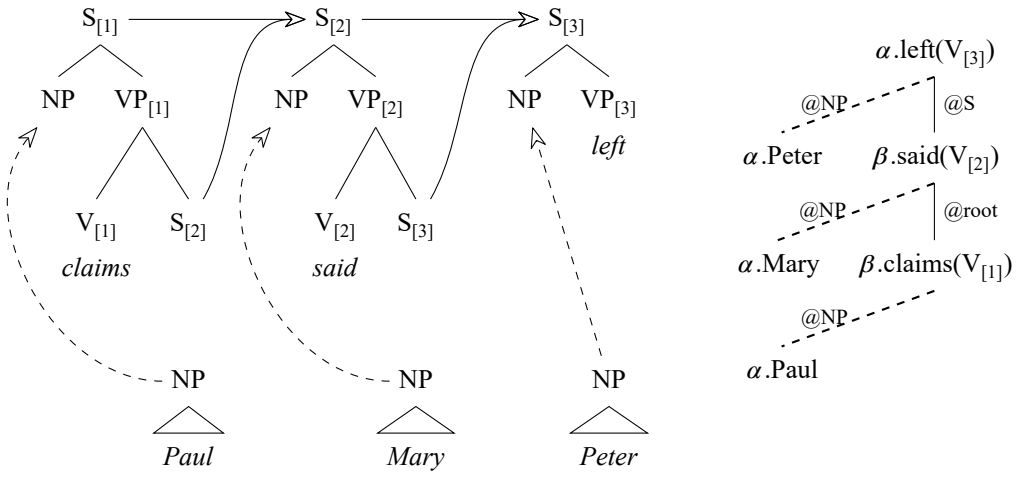
Joshi, Becker & Rambow (2000: 175-177), Chen-Main & Joshi (2008: Sec. 2), and Lichte (2015: 149 ff.) focus on scrambling in German (see also Abels 2015 for related discussion) to illustrate derivations where predicative auxiliary trees are involved. Although one could use one of the structures they discuss as an instance of (187), they will be tackled in section 5.4.5 for reasons that will be made clear at that point. For the time being, (187) will be exemplified by resorting to an English sentence whose analysis is briefly sketched by Candito & Kahane (1998: 23), namely *Paul claims Mary said Peter left*. All the works just mentioned adopt neither the C-T-v-V system nor the DP hypothesis discussed in sections 2.2.3 and 2.2.4, but rather the projections predating it. Thus, Joshi, Becker & Rambow (2000: 175, italics in the original) point out that "all verbs except V_n" (where *n* corresponds to the maximal value that an index can have in the dependent derivation at hand) anchor an auxiliary tree and subcategorize for an *S*, which labels both the root and the foot in each of such trees. All verbs (including V_n) further subcategorize for the nominal arguments they require, which substitute to the positions where they are overtly realized, as illustrated in section 5.1.1.[146] Leaving for section 5.4.1 the discussion that the contrast between the C-T-v-V system and some S projection deserves, one can acknowledge that the representation in (188) manages to provide the desired illustration nonetheless. It is also worth noting that all the auxiliary trees involved in (188) are left auxiliary trees, just like in (182).

The disjoint between labeling and indexing still persists when adapting the strong co-occurrence constraint to a different scenario like (185b), namely a dependent derivation where modifier auxiliary trees are involved. However, (187) must be tweaked so at to effectively apply among modifier trees, as follows:

(189) Strong co-occurrence constraint (for dependent derivations where all the auxiliary trees involved are of the modifier type): An auxiliary tree $\beta.M_i$ must adjoin to $\beta.M'_{i-1}$, provided that both trees eventually modify the same lexical anchor A.

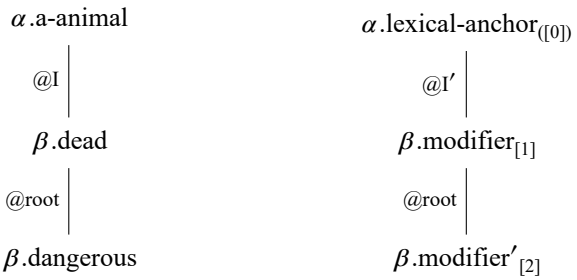
[146] To further ensure that the argument(s) substitute into the desired tree (and not into a wrong one), the index which the latter bears can be assigned to the former, so that each $NP_{[i]}$ combines to the elementary tree anchored by $V_{[i]}$. See Joshi, Becker & Rambow (2000: 175) and Lichte (2015: 148-149), where this principle is called the *weak co-occurrence constraint*.

(188)



The newly proposed strong co-occurrence constraint in (189) is formulated by positing that the modified anchor bears either no relevant index or one whose value corresponds to the one of the closest modifier, but diminished by 1. More generally, the closer a modifier is to the anchor in structural terms, the smaller its indexical value. The dependent derivation in (185b) can therefore be generalized as shown in (190), where I' indicates an unspecified internal node which must be determined on a case-by-case basis, as it was done for (185).

(190)



What has just been pointed out satisfies the aforementioned desiderata. In fact, (189) departs from (187) in a way which is not so *ad hoc*, as the differences between them allow to distinguish between predicative and modifier auxiliary trees in a novel manner. If any β_i belongs to the former type, then its foot must be indexed as $[i + 1]$. If instead β_i is of the latter type, then $[i - 1]$ must be the index of its foot.

A further consideration can be made, namely that the original strong co-occurrence constraint in (187) is established in such a way that the more embedded a lexical verb is in the derived structure, the higher its hierarchical position in the derivation tree and the value of its index. Since this mechanism refers to the absolute structural depths (in both the derived and

the derivation tree) of the anchors involved, and not to the relative structural distances from the only lexical head being modified as for (189), one can argue that the indexing among predicative trees shown in (188) can be reversed, i.e. the more embedded a lexical verb is, the lower the the value of its index. However, this alternate view would be problematic, because both (187) and (189) would posit that an auxiliary tree β_i must adjoin to β_{i-1}, thus making the two dependent derivations hardly distinguishable on the derivation tree. Still, section 5.4.2 and the subsequent sections will pick up on the present observation, as a similar issue will arise there due to unrelated reasons.

5.2.2 An L-TAG adaptation of Bruening's analysis of U20

Bernardi (2023) sketched a possible way to account for the crosslinguistic variation with respect to U20 by using an L-TAG that adheres to what has been discuss so far in the present chapter. Such a proposal turns out to be a TAG counterpart of what Bruening (2018a) proposes, i.e. an account of sixteen orders (rather than fourteen as in Cinque 2005 and related works) that assumes that the rightward reordering of nominal modifiers is totally free. The crucial difference with respect to the original analysis in section 3.2.5 is that no remerger of the nominal modifiers is necessary in the TAG reformulation that will be illustrated shortly.

This TAG account of U20 essentially hinges on the intersection of two divides introduced in section 5.2.1. On the one hand, nominal modifiers can be argued to anchor elementary trees that can be either predicative or modifier auxiliary trees as per Schabes & Shieber (1992, 1994). On the other hand, one can further assume that each of such trees are of the non-wrapping type. This implies following the divide provided by Schabes & Waters (1995: 481), so that prenominal modifiers anchor left auxiliary trees, while postnominal modifiers anchor right auxiliary trees.

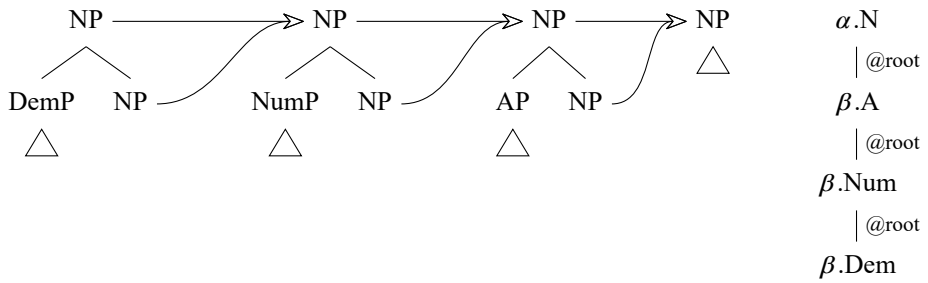
These premises already provide the tools to formulate a coarse-grained account of U20 that captures the well-known rigidity of the order of prenominal modifiers as opposed to the high flexibility of the postnominal ones. On the one hand, the rigidity of the former can be argued to follow from the the following:

(191) *Dem* > *Num* > *A* > *N*

On first sight, (191) is reminiscent of a consensus among the various works discussed in section 3.2, namely the hierarchical relation that is deemed to apply on any derived tree that

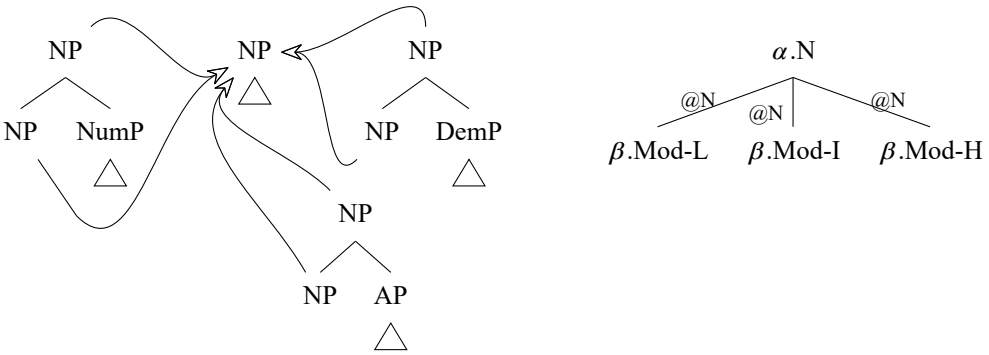
contains the elements of U20. However, in the present TAG context (191) does not apply on the derived trees, but rather constrains the combination among predicative auxiliary trees. Thus, what (191) shows is an *independent principle* that regulates the combinations among auxiliary trees participating in a dependent derivation like the one illustrated in (192). If a TAG derivation is somehow ruled by (191) and the derived and derivation trees respectively encode syntactic and semantic information, then (191) must either have an extragrammatical origin (Ramchand & Svenonius 2014: Sec. 6) or predate both syntax and semantics in their language-specific dimensions (as per Wiltschko 2014: Sec. 8.6.2.4; see also footnote 177 on page 331); whence its independent nature.

(192)



On the other hand, right modifier auxiliary trees arguably override the independent principle in (191) by being modifier auxiliary trees. As a consequence, all the postnominal modifiers of U20 involved in the derivation target the root of $\alpha.N$ at once. The cases where all the nominal modifiers of U20 are postnominal can be accounted for by means of the scenario in (193). Unfortunately, at first sight the auxiliary trees in (192) appear to be of the same nature of those reported in (193). This issue derives from the fact that both representations simplify the internal structure of the auxiliary trees. In fact, in (192) each foot node would be a sister of the anchor (or of a syncategorematic element selected by it, as will be discussed further in section 5.5.2) even if the internal structure of the corresponding auxiliary tree were more complex. On the contrary, in (193) each foot node would always be a daugther of the corresponding root, no matter how complex the internal structure of the modifier is.

(193)



Still, a sharp contrast between (192) and (193) effectively emerges by inspecting the different nature of the protocols shown for their respective derivations. In fact, the one reported in (192) is a derivation tree like the others presented so far, i.e. it maps a uniquely determined derived tree as per the quotation on page 221 of Kallmeyer (2010: 57 fn. 3). On the contrary, the protocol of the derivation in (193) abstracts over distinct derivation trees and their corresponding derived trees. Therefore, it cannot be a derivation tree in the strict sense because it does not illustrate a specific independent derivation, but rather shows how to obtain six different ones where all the nominal modifers are still consistently postnominal.

By taking inspiration from the definition of *tree-family* given by Abeillé et al. (1990: 11), one can argue that the protocol shown in (193) is a derivation-tree-family, as it abstracts over a set of six derivation trees that share the same dependency structure (see also Lichte & Kallmeyer 2017: 208 ff. for a different implementation of a similar idea).

Based on the discussion on page 252, what remains constant among the alternate derivation trees encompassed by (193) is that $\alpha.N$ is the targeted tree and that its rightmost daughter in the protocol always ends up being the highest modifier in the corresponding derived tree, whence its label $\beta.Mod\text{-}H$. On the same ground, the leftmost sister is indicated as $\beta.Mod\text{-}L$ because it ends up being the lowest modifier on the derived tree. Lastly, the daughter of $\alpha.N$ that stands between $\beta.Mod\text{-}L$ and $\beta.Mod\text{-}H$ is indicated as $\beta.Mod\text{-}I$ because it ends up in an intermediate position, i.e. structurally higher than the former but lower than the latter in the derived tree.

Each derivation tree belonging to the derivation-tree-family in (193) eventually ends up being different from the others depending on which modifier of U20 is assigned to a certain daughter of $\alpha.N$. Given that all the auxiliary trees involved are of the right type, the further to the right a daugther is in the protocol of the derivation, the further to the right it is in the

eventual linear order as well (see again related discussion on page 252). By taking the order N-Num-A-Dem as an example, it can be derived by means of (193) by positing that Num, A and Dem anchor *β.Mod-L*, *β.Mod-I* and *β.Mod-H*, respectively.

Based on (192), Dem-Num-A-N should resist diachronic change and be highly frequent crosslinguistically. Although the latter fact is undeniably true, one first problem that the account presented here raises is that it generates the order N-Num-Dem-A, which ruled out in Cinque (2005) and related works. Although one can follow Bruening (2018a) and argue that at least this account does not undergenerate by predicting orders that are yet to be attested, there are even more serious issues to take care of. First of all, the six orders that can be accounted for by means of (193), i.e. those where the noun always comes first, should be prone to greater variability than predicted, on a both intra- and cross-linguistic standpoint.

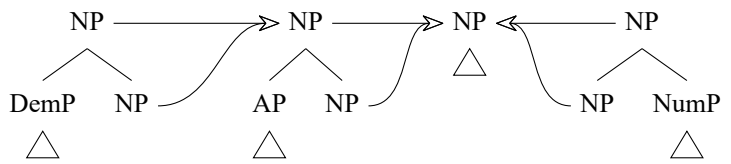
Moreover, in (193) N-A-Num-Dem bears no peculiarity as opposed to the other five N-first alternatives that are much rarer, even though it is the most frequent order in all the samples discussed in chapter 4. One could try to propose that N-A-Num-Dem is preferable because its derived tree complies with the *Dem* > *Num* > *A* > *N* hierarchy in (191), but in the present L-TAG account the latter principle rules the way in which distinct (predicative) auxiliary trees combine with one another, without imposing any restriction on the derived tree.

If such a hierarchy were of any relevance in the derived tree, the other six homomorphic orders that comply with it (i.e. Dem-N-A-Num, Dem-A-N-Num, Num-A-N-Dem, Num-N-A-Dem, A-N-Num-Dem, Dem-Num-N-A) should be more common than the empirical data suggest (an issue already pointed out in section 3.2 for proposals like Rijkhoff 2002, 2015, Abels & Neeleman 2009, 2012), as their derivational history should be overridden by the properties of the derived tree, as just stated for N-A-Num-Dem.

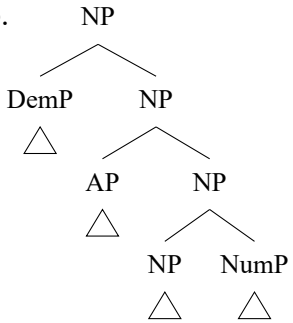
In the present account, the six homomorphic orders can be derived by assuming that the spines of the elementary trees anchored by the modifiers do not bear a single consistent directionality (i.e. some are left auxiliaries whereas others are right auxiliaries). In order to derive one order like Dem-A-N-Num, it is thus necessary to resort to the combination of trees shown in (194a), where the left auxiliary trees *β.Dem* and *β.A* are of the predicative type whereas *β.Num* is a right modifier auxiliary tree. As a consequence, the combination of such trees entails a mixed derivations *à la* Schabes & Shieber (1992, 1994). By following the outermost predication rule posited by Schabes & Shieber (1994: 102), the derivation tree corresponding

to (194a) is thus (194c), and it leads to the derived tree in (194b).[147]

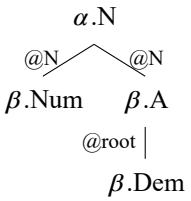
(194) a.



b.



c.



The illustration in (194) shows that the six orders that Rijkhoff (2002, 2015) and Abels & Neeleman (2009, 2012) argue to be base-generated (together with Dem-Num-A-N and N-A-Num-Dem) can be obtained by means of theoretically marked premises, namely the mingling of both predicative and modifier auxiliary trees, with the subsequent mixing of a dependent and an independent derivation targeting the same node of *α.N*. This way of accounting for such orders is arguably problematic, especially because it also allows to generate the non-homomorphic orders Dem-N-Num-A and Num-N-Dem-A, the latter being an order whose attestation is a further matter of debate (see discussion in sections 3.2.5 and 3.2.6).

Finally, the TAG analysis of U20 discussed so far counters one of the major desiderata mentioned in section 5.2.1, namely that derivation trees establish dependency structures that may also account for semantic composition. If the latter idea is on the right track, the (families of) derivation trees presented in (192), (193) and (194) determine interpretations that should greatly differ both intra- and cross-linguistically, even though the same elements keep being involved among the various derivations.

Although such a possibility may deserve further theoretical exploration in the future, it would be at odds with all the works presented in section 3.2, which agree in assuming that

[147] It is worth recalling that Gardent & Narayan (2015) question the outermost predication rule, but they do so by providing counterexamples that arguably hold only if one follows certain assumptions that are by no means cogent (see again discussion in sections 5.4.3 and 5.4.4).

Greenberg's U20 is a typological generalization that applies crosslinguistically among the nominal modifiers by establishing hierarchical and/or scope relations that are always constant. The derived tree in (194b) further shows that the present account cannot pass this assumption on to the derived tree either, as the hierarchical relations among the modifiers of U20 also vary in that kind of structure depending on the permutation that needs to be accounted for.

5.2.3 An L-TAG adaptation of Abels and Neeleman's analysis of U20

Contrary to what was just treated in the previous section, the upcoing discussion will try to provide a TAG account that avoids assuming inconsistent dependency relations among the modifiers of U20 both in the derivation and in the derived trees. An alternate TAG approach to U20 that follows these lines can be presented by hinging on an adaptation of a theoretical tool first proposed by Cinque (2005), and that many following works (willy-nilly, as in Abels & Neeleman 2009, 2012) cannot do without, namely pied-piping. In rather abstract terms, one can present the essential ingredients of pied-piping as in (195).

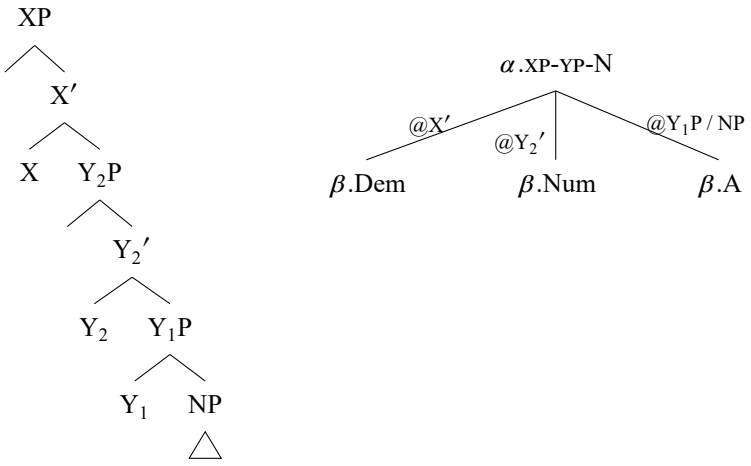
(195) Pied-piping applies under the following conditions:

a. A projection XP and a projection HP are first-merged so that the former dominates the latter.
b. The left specifier of X is free and can host HP.
c. The structural portion that intervenes between XP and HP must contain some overt material Y that c-commands the first-merged position of HP without being involved in the pied-piping of HP.

Both (195a) and (195b) are the rather standard conditions under which pied-piping occurs. The leftwardness of Spec,XP is the only additional stipulation, but can be grounded on any of the possible motivations that were covered in chapter 2. The condition in (195c) establishes instead the presence of some Y that can correspond to many projections, to a single YP or even to a non-projecting head (Bruening 2009: 31, Salzmann 2020: Sec. 2.1.3). The presence of such Y in the structure fulfils a twofold purpose. On the one hand, it (correctly) predicts that not everything that is dominated by XP should be pied-piped. On the other hand, the presence of some Y allows to prevent a Comp-to-Spec raising of the pied-piped material that would be problematic as per anti-locality (already discussed in section 2.4.1).

Based on (195), it becomes possible to formulate an L-TAG conversion of the analysis provided by Abels & Neeleman (2009, 2012) and related works, previously discussed in section 3.2.4. The necessary premises of such an account are provided in (196), and they consist of a nominal extended projection *α.XP-YP-N* whose structure allows to have all the modifiers of U20 adjoined at distinct nodes, as indicated in the contextually given derivation tree family. In (196), two distinct projections are both labeled YP because they correspond to what (195c) establishes, namely the presence of some structural portion intervening between XP and the raised HP. In (196), the latter is instantiated by NP, which can raise leftwards and in one fell swoop either to Spec,XP or to Spec,Y_2P.[148]

(196)



In the present account, the modifiers of U20 anchor (non-wrapping) elementary trees that can be indifferently predicative or modifier auxiliary trees. In fact, the internal structure of such trees is not relevant, as they are never directly combined among one another, as opposed to the analysis of U20 presented in section 5.2.2. Moreover, the hierarchy that Abels & Neeleman (2009, 2012) and related works assume for the derived tree follows from the assumed addresses shown on the edges of the derivation tree in (196).

Thus, two aspects of *α.XP-YP-N* seem to truly matter. On the one hand, the presence of some node (e.g. X′ and/or Y_2′) that dominates the head(s) that the Noun has to climb over (i.e. X and/or Y_2) while being c-commanded by the landing site of the raised Noun itself. On the other hand, the node targeted by *β.Num* must dominate a node that can be shoehorned by *β.A* while being dominated by a node at which *β.Dem* adjoins. This also implies that it is

[148] Alternatively, one could follow Georgi & Müller (2010) and argue that such specifiers are absent from (196) and that the Noun rather moves past the X head or the Y_2 head by reprojecting. However, such an alternative would raise some concern with respect to the exact node at which *β.A* must adjoin to derive an order like N-A-Dem-Num.

not strictly necessary that the nodes at which $\beta.Num$ and $\beta.Dem$ are adjoined be intermediate projections, just like it is not important whether XP, Y_2P and Y_1P are head-initial or head-final in (196).

The presence of Y_1P in (196) largely depends on one's take on two matters, namely Comp-to-Spec anti-locality and the copy theory of movement. If one argues that the former is not an issue while the latter holds (so that an unpronounced copy of the Noun is left behind whenever it raises leftwards), then Y_1P can be dispensed with. The latter maximal projection is in fact necessary to prevent the remerger of the Noun within the same projection in which it is first-merged, as well as to account for permutations like N-Dem-A-Num and N-Num-A-Dem, where N raises while A is left in a structural position which is lower than those of the other two modifiers of U20.

However, Abels & Neeleman (2009, 2012) and related works have to assume that A can be dragged along by N when the latter raises leftwards in order to account for permutations like A-N-Dem-Num or N-A-Dem-Num. In the present account, the latter assumption can be captured by assuming that $\beta.A$ can adjoin at NP, no matter if raised or not. Thus, if NP leaves an unpronounced copy in its first-merged position whenever it raises to some specifier (no matter how close), N-Dem-A-Num and N-Num-A-Dem can be accounted for by doing without Y_1P.

The crucial insights of the TAG analysis just presented would largely coincide with those of Abels & Neeleman (2009, 2012) and related works. The eight orders that they argue to be base-generated can be derived with no raising of NP in (196) while having each modifier of U20 anchor an auxiliary tree that is either of the left or of the right (non-wrapping) type. The six orders that are further attested according to Cinque (2005, 2021, 2023) can be derived (together with some reduntant alternatives to the base-generated orders) by raising NP to the left and by having $\beta.A$ adjoined to one of the various legitimate options discussed above. Similarly to what was previously discussed for the original proposal in Abels & Neeleman (2009, 2012), such a TAG adaptation of their work fails at providing a thorough account for the typological observations presented in chapter 4.

The major differences between Abels & Neeleman (2009, 2012) and the TAG adaptation just presented are due to what (L-)TAGs bring to the table more generally. For a start, the optional nature of the modifiers of U20 can be linked to their adjoining at the initial tree anchored by the Noun. The crosslinguistically consistent hierarchy among the various derived

trees following (196) descends instead from the conditions on pied-piping summarized in (195). Given that Abels & Neeleman (2009, 2012) are not committed to the Functional Head Hypothesis for the nominal domain discussed in sections 2.2.3 and 2.2.4 (actually, Georgi & Müller 2010 rather go in the opposite direction, i.e. towards the Lexical Head Hypothesis), the articulated structure of the initial tree in (196) represents another peculiarity of the TAG adaptation currently discussed.

Contrary to the proposal discussed in section 5.2.2, the present account seemingly allows to formulate a crosslinguistically consistent interpretation of U20. In fact, the present alternative assumes that the various derivation trees are constructional homonyms in the sense of Chomsky (2002–1957: 28, 86-87), namely they can be traced back to the same template given in (196), even though they vary among one another by altering the structural details within the elementary trees involved in the derivation itself.

However, the derivation tree family in (196) is problematic because the various auxiliary trees neither combine with one another nor target the same node (and for this very reason, the daugthers of α.*XP-YP-N* do not form an independent derivation and thus do not need to be ordered as per Schabes & Shieber 1992, 1994). This implies in turn that the often assumed scope relations among the modifiers of U20 are not due to established dependencies in the derivation trees, but rather follow from the structure of the corresponding derived trees. Although this view is consistent with what Chen-Main & Joshi (2014: Sec. 6) argue, it is in contrast with the idea (reported in section 5.2.1) that the crucial semantic dependencies are encoded by derivation trees.

Any attempt to account for the semantics on the derived tree is further weakened by another problem raised by the present account, namely the theoretical necessity of the projections that dominate NP in the initial tree in (196). Given that their presence in the structure is grounded on the properties of pied-piping presented in (195), it can be argued that they are unnecessary whenever the Noun does not raise to the left. By pushing the argument further, one could claim that only the six orders that are not base-generated in Abels & Neeleman (2009, 2012) do need an explanation based on the FHH, whereas the eight orders that they assume to be base-generated are rather the result of derivations that comply with the LHH.

The eight orders that an approach *à la* Abels & Neeleman (2009, 2012) base-generates can be obtained by revising the dependent derivation assumed on page 258 for the account discussed in section 5.2.2. Although being still constrained by the independent principle

given in (191), this time the spines of the auxiliary trees anchoring the nominal modifiers are not constrained with respect to directionality, and such trees can combine with one another as soon as it is further argued that they are predicative trees that ultimately target the maximal projection of NP. This refinement would further weaken the relevance of the derivation trees in the present account, as such dependent derivations determined by the LHH would counter the previously mentioned constructional homonymity among the various derivation trees, because in the counterparts following the LHH the modifiers of U20 do not target distinct nodes of the same initial tree (independently of one another).

5.3 Empowering the formalism

It is undeniable that the refinements proposed by Schabes & Shieber (1992, 1994) improve TAGs by stating "a range of interactions between dependent vs. independent derivation on the one hand, and syntactic constraints, linear ordering, and scopal vs. nonscopal semantic dependencies on the other hand" (Gardent & Narayan 2015: 64). However, they establish unappealing dependencies between main and subordinate clauses in the derivation trees. Moreover, they remain silent on the issues to be discussed from section 5.3.1 on, namely the syntactic phenomena that cannot be captured by means of the L-TAGs discussed so far.

To face these problems and make TAGs provide more adequate accounts, Frank (2002: Sec. 5.6, 5.7, 6.1) himself ends up appealing to a more powerful TAG alternative that section 5.3.1 will illustrate. More generally, three alternate strategies have been adopted in the literature for the same purpose. The first descends from a fact which was hinted at in (186) on page 252. As long as they are generated by different grammars, "two derived trees may exhibit quite divergent structure[s] while sharing a single derivation tree" (Nesson 2009: 23). This property can be exploited to develop a Synchronous TAG (or STAG; see a.o. Nesson 2009, Longenbaugh 2014: Ch. 2, Findlay 2023: 2088), so that the two derived trees are either the realization of the same semantic proposition in two different natural languages (Nesson 2009: Ch. 7) or the "pairing [of] a TAG grammar with a tree-based semantic representation" (Findlay 2023: 2088). In the latter case, a STAG must comply with the mechanisms that Findlay (2023: 2088) summarizes as follows:

> In STAG, elementary trees from one grammar are paired with those from another, and links are established between individual nodes in those trees. Then, when

> adjunction or substitution applies in one grammar, it must also take place in the other, at the linked node, and using the equivalent, paired tree. By pairing a TAG grammar with a tree-based semantic representation, we can therefore implement a "rule-to-rule" approach [...].

Given that the meaning of the sentence can be read off the derived semantic representation (Nesson 2009: 124), in STAGs the derivation tree is framed again as the full protocol of the derivation, just as it was claimed to be on page 221. What is more, the "extent of the isomorphism between the derived structures" becomes a matter of inquiry in STAGs (Nesson 2009: iii), as "[n]othing requires the paired trees to be isomorphic" (Findlay 2023: 2088). This is also worth mentioning because Longenbaugh (2014: Sec. 2.1) exploits this premise to propose a STAG where the Phonetic Form (PF) and the Logical Form (LF) assumed in Minimalism are generated by distinct grammars whose formal powers are only weakly equivalent (see again section 5.1.1 or Joshi 2003: 243).

As interesting and promising as they are, STAGs will no longer be a central concern from now on (note however that the approach proposed by Longenbaugh 2014 will be discussed further in section 5.4.3). In fact, Schabes & Shieber (1992, 1994) arguably have a twofold merit that is of little interest for STAGs. They firstly demonstrate that "there are several possible ways of defining TAG derivation trees" (Gardent & Narayan 2015: 41), and by doing so they consequently encourage research aimed at obtaining "a greater understanding of the role of the derivation structure" in TAGs (Chen-Main 2012: 106). It can thus be argued that it is worth following either the second or the third strategy adopted in the literature so as to provide more adequate TAG accounts, as they both descend from the insights just mentioned. Kallmeyer & Kuhlmann (2012: Sec. 6) summarize such strategies in the following way:[149]

> On the one hand, several authors have proposed to use a variant of TAG that yields different derivation trees than standard TAG. On the other hand, some approaches keep standard TAG while exploring ways to obtain the missing [dependency] links from the derivation tree.

In opposition to Kallmeyer & Kuhlmann (2012) and references therein, the next sections will be committed to the the first of the two possibilities that they sketch. This will lead

[149] A sub-branch of the second approach summarized by Kallmeyer & Kuhlmann (2012: Sec. 6) consists in having a standard L-TAG with the missing dependency link retrieved from the derived tree. See Chen-Main & Joshi (2014: 1013 ff.).

to section 5.4, where a solution of the "missing link problem" discussed by Kallmeyer & Kuhlmann (2012) and Chen-Main & Joshi (2014: Sec. 6.1) will be provided. The insight resulting from this improvement of TAGs will eventually turn out to be crucial for section 5.5, where Greenberg's U20 will be recast by means of a TAG analysis that manages to capture the insights emerging from chapter 4.

5.3.1 Multi-Component TAGs

Since the TAG literature has provided and discussed several TAG alternatives that are more powerful than an L-TAG (see a.o. Weir 1988, Chen-Main & Joshi 2007, Kallmeyer 2009, Nesson & Shieber 2009, Lichte 2015 for good overviews), the choice among one of those is a serious methodological concern. The prosecution of the already cited quotation from Bleam (2000: 193) on page 212 provides good guidelines for this purpose:

> [Since TAGs] restrict the range of possible analyses, and some of the burden of explanation is shifted to the formal system[, then t]he most profitable research strategy [...] is to use the most restrictive version of TAG available. Only when this strategy fails should we increase the power of the formal system and supplement it with a greater number of linguistic constraints to take over the burden of explanation.

With such premises in mind, the improvement of an inadequate TAG approach must also consider what Findlay (2023: 2071) formulates with respect to CFGs. "[T]here are two different kinds of response for the syntactic theorist", where the first one is to "replace the [TAG] with something more powerful", while the second one is to "beef up the [TAG] with something extra, so that the *combination* is more powerful" (italics in the original). Thus, the first alternative must be concerned with the weak generative power of the replacing formalism, so that its strong power is eventually increased as well. On the contrary, the term *combination* in the second alternative exclusively refers to the gain of strong generative power of the beefed-up TAG, whereas its weak generative power must be proven to remain intact. The discussion from now on will thus be concerned with such a second option, as section 5.1.2 has already shown that a well-known property of L-TAGs is that their weak generative capacity is highly likely to be the perfect fit for the strings that are empirically findable in natural language syntax (Frank 2004: 675-676).

Many works in the TAG literature are concerned with tree rewriting systems where two or more elementary trees are grouped together in a set because they are united by some established criteria (to be provided in due course). Such sets of trees are known as *multi-components* (because each elementary tree is understood as a *component*, namely a member, of a set) and TAGs allowing them are thus called MC-TAGs (Kallmeyer 2009). An important condition applying to these approaches is that "if an instance of an elementary tree set is used, then all trees from this set must occur in the derivation tree" (Kallmeyer 2009: 64).

There are two further properties that make lexicalized MC-TAGs different from L-TAGs. First, in MC-TAGs "lexicalization means that every tree set in the grammar contains at least one lexicalized tree. It does not require all trees in the grammar to be lexicalized" (Kallmeyer 2009: 80; see also Joshi 2014: 152). This means that the clearly distinct strategies that can be adopted to lexicalize an L-TAG (see Krivochen & Padovan 2021: Sec. 3 for references and discussion) become interestingly blurred in MC-TAGs, as sections 5.3.2 and 5.4.3 will show and discuss, respectively. Second, whereas any L-TAG establishes a single domain of locality, i.e. the one inside its elementary trees, MC-TAGs force the definition of two further domains of this kind. The first is concerned with the relation among members of the same multi-component (i.e. elementary trees belonging to the same set of trees), while the second constrains the degree of variation that is allowed in derivation trees (Nesson & Shieber 2009: Sec. 3).

Although these facts compel one to define three domains of locality for well-formed MC-TAGs, in the upcoming discussion of this chapter the characteristics that determine the single elementary trees will be fewer and less central than they are in L-TAGs. In fact, the elementary trees of any L-TAG can be seen as singleton sets of an MC-TAG. More formally put, if *h*-MC-TAG is used to say that the maximum arity of the components in an MC-TAG is established by the value of *h*, then any L-TAG is a 1-MC-TAG, with $h = 1$ (Boullier & Sagot 2011: Sec. 2.2).

Seeing L-TAGs as a specific subset of MC-TAGs is important for the discussion at hand because this implies that Frank's FTH in (152) and CETM in (153), both discussed from page 214 on, may actually define the domain of locality *in a set*, which happens to be defined with respect an elementary tree because any such set is a singleton in an L-TAG. In fact, this is also remarked by Bleam (2000: 204, 212-213), Frank (2002: 221) and Frank (2004: 693 fn. 6), so that one can safely apply the following adjustments to any *h*-MC-TAG:

(197) a. Revised FTH: Every syntactic dependency is expressed locally within an elementary tree *set*.

b. Revised CETM: The syntactic heads in a *set* of elementary trees and their projections must form an extended projection of a single lexical head.

The theta-criterion in (154) also needs to be restated. Although (198) gives a reworked definition, section 5.4.3 will discuss facts suggesting that some improvement might be needed (despite being a concern for further research).

(198) Revised θ-criterion (MC-TAG version):

a. If H is the lexical head of *a set of trees S*, H assigns all of its θ-roles in *S*.

b. If A is a frontier non-terminal node of *a component of S*, A must be assigned a θ-role in *S*.

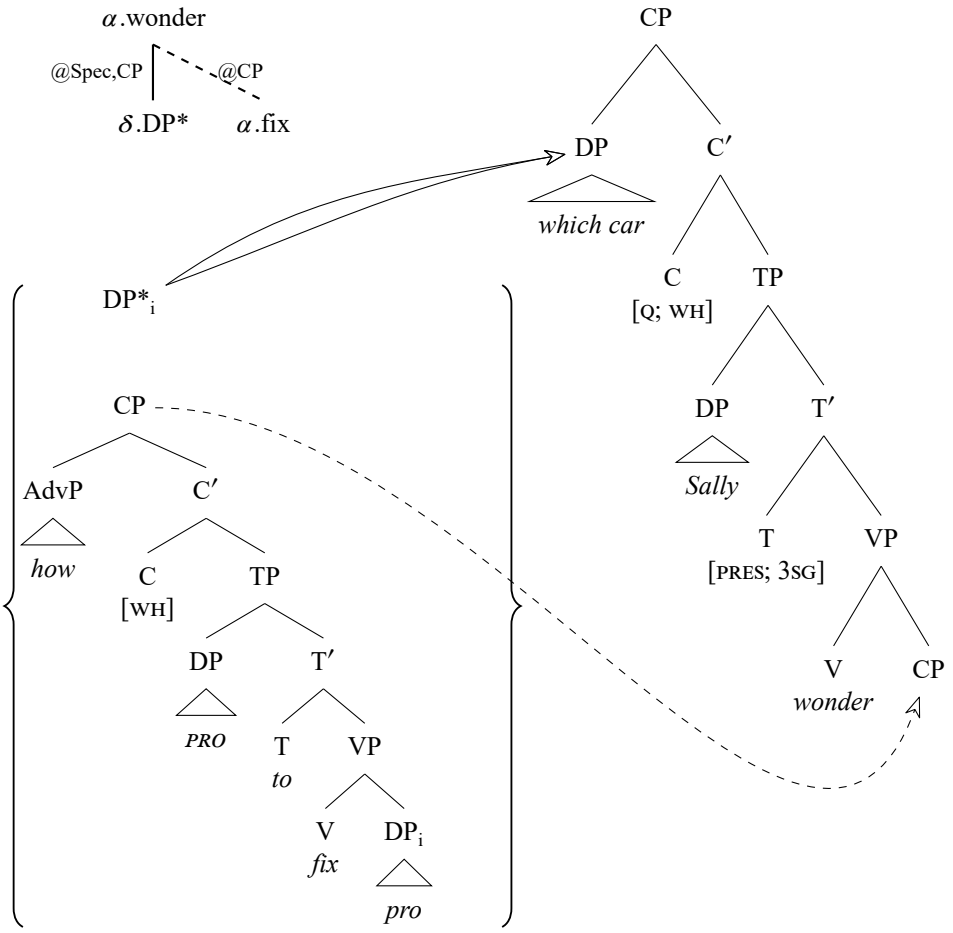
An MC-TAG that complies with the revised well-formedness conditions in (197) and (198) is provided by Frank (2002: Sec. 5.6, 5.7, 6.1), where a 2-MC-TAG (where 2 indicates the maximum arity *h* allowed for each multi-component) is adopted so as to account for some extraction phenomena that can be observed in different languages, including English. Frank essentially uses two lines of argument for all the cases he discusses, one being repeated throughout Frank (2002: Sec. 5.6, 5.7) and the other being discussed in Frank (2002: Sec. 6.1). Leaving the latter for section 5.4.4, the former will now be presented by means of a single representative example, i.e. the sentence *Which car does Sally wonder how to fix?* (Frank 2002: 219 ff.).

As Merlo (1996: 161-162) points out, an L-TAG approach *à la* Frank faces the same problem of previous successive-cyclic accounts, namely the difficulty of explaining how a single specifier of CP can host both *how* and *which car* raising there. In fact, an L-TAG does not differ from many other syntactic proposals with respect to two assumptions, namely that (i) *wh*-operators raise to Spec,CP and that (ii) each CP must have a single specifier in English (Kroch & Joshi 1985: Sec. 6.1). Thus, any L-TAG approach must resolve the problematic starting point that has just been mentioned (i.e. the "basic step" in Merlo 1996: 190) before explaining the subsequent *wh*-displacements (i.e. the "iterative step" in Merlo 1996: 190) in a manner that is not successive-cyclic.

As a solution to this problem, Frank (2002: 219 ff.) formulates a TAG adaptation of a previous analysis provided by Cinque (1990b), where the chain headed by *which car* "involves

the binding of a null pronominal variable by a *wh*-operator that is base-generated in the specifier of CP of the main clause" (Frank 2002: 219, italics in the original). The key tool that Frank uses to meet this desideratum with TAGs is a multi-component set of trees where the embedded clause is paired with a DP auxiliary tree that is coindexed with the object pronominal selected by the verb *fix*. Frank (2002: 222-223, 228, 288-290) calls any auxiliary tree of the latter kind *degenerate*, as it consists of a single point carrying some relevant features (such as the required index). Frank argues that, by virtue of their zero-dimensionality, degenerate trees are peculiar auxiliary trees whose spine (in the sense of Nesson 2009: 4, Kuhlmann & Satta 2012: 623) is made of a single node. In fact, root and foot nodes are collapsed together and no other node is allowed in the path supposed to connect them (see again section 5.2.1). Thus, the 2-MC-TAG account just summarized can be represented as in (199), where curly braces encompass the members of the same multi-component.[150]

(199)



[150] For the sake of precision, a pair of curly braces should also be used around any tree which is the unique member of a singleton set. However, such braces are left implied for reasons of practicality.

The derivation tree shown in (199) features an innovation that (to the best of my knowledge) is not present in previous works. In fact, the degenerate tree is indicated as a δ tree albeit it is used as an auxiliary (i.e. β) tree. Although the formal properties previously illustrated for degenerate trees (footnote 142 on page 243 included) already back up this choice, section 5.3.2 will provide further arguments in this respect. What matters at this point of the discussion is that the support of a degenerate tree which pairs with an extended projection within the same multi-component is sufficient to account for the problematic cases discussed by Frank (2002: Sec. 5.6, 5.7, 6.1). However, it is easy to see that this kind of solution is rather *ad hoc* and prone to overgeneration. This holds even if one follows Frank (2002: Sec. 5.6, 5.7, 6.1) in adopting a widely assumed constraint for MC-TAGs (to be discussed in due course), namely the "simultaneity requirement" (Kallmeyer 2009), which establishes (among other things) that any two trees belonging to the same multi-component cannot combine with one another. In fact, the degenerate DP can still adjoin to the subject DP *Sally* as easily as it does with the *wh*-operator *which car*, as such two DPs are both part of the same elementary tree which belongs to another (singleton) set of trees. As a consequence of this alternate adjoining, the ungrammatical coindexing in (200a) can be generated. One may try to fix this by stating that degenerate DPs can only adjoin to *wh*-operators, but the latter restriction would still give rise to (200b).

(200) a. * Which car does Sally$_{i}$ wonder how *PRO*$_{i}$ to fix *pro*$_{i}$?
b. * Who$_{i}$ wonders how *PRO*$_{i}$ to fix *pro*$_{i}$?

The examples above demonstrate that degenerate trees can be so powerful that one ultimately needs to stipulate the exact cases where it is possible to safely adjoin them, and (200b) further casts doubt on whether "listings of (permitted/required) auxiliary trees" that are "not particularly linguistically illuminating" (Findlay 2023: 2082) could actually be suitable for the purpose, even though TAGs originally constrained adjoining operations in this way (see Abeillé & Rambow 2000: 11-12, Frank 2013: 244, Lichte 2015: 130, Findlay 2023: Sec. 3.3). In a similar fashion, section 5.4.2 will show another case where degenerate trees allow coindexed pairs both when they are allowed and when they are ruled out under standard assumptions of the Binding Theory (see Sportiche 2017 and references therein).

From the discussion above one might be led to agree with Longenbaugh (2014) and claim that the MC-TAG approach proposed by Frank (2002) raises so many problems "that it be-

comes uncertain [...] what exactly has been gained by switching to a TAG formalism" (Longenbaugh 2014: 42). However, the discussion from now on will depart from Longenbaugh (2014: Ch. 3) with respect to the decisions to be taken as a reaction to this situation.[151] Instead of arguing that it is safer to move back to an L-TAG and shift all the problematic cases to the post-syntactic component, this chapter will claim that it is worth exploring the possibilities given by MC-TAGs so as to choose a formal system and linguistic constraints aimed at improving all the TAG analyses proposed so far in the literature.

With these premises set, it is now time to focus on a further domain of locality that MC-TAGs permit to establish, namely the one in the derivation tree (Nesson 2009: Ch. 3, Nesson & Shieber 2009: Sec. 3). Let us take the MC-TAG *à la* Frank (2002: Sec. 5.6, 5.7) shown in (199) as an example illustrating one first possibility, which consists in having all the trees from the same multi-component combined to "a single, non-constructed elementary tree" (Longenbaugh 2014: 24) contained in a different set of trees. This option is defined as a Tree-Local MC-TAG (henceforth TL-MC-TAG), which can be ruled by the following principle:

(201) Tree-Locality Requirement: An MC-TAG is Tree-Local if, in any derivation tree it generates, all the nodes (standing for elementary trees) that belong to a single multi-component set are sisters, i.e. daughters of the same (mother) node.

The statement in (201) is formulated in accordance with another restriction that has been anticipated above in this section, namely the fact that any two elementary trees from the same set cannot combine with one another. If they did, the targeting tree would either give rise to a problematic non-Markovian derivation (see section 5.1.4) or fail to be a daughter in the derivation tree to the otherwise common mother node, as imposed by (201). The ban of such problematic scenarios is due to (the first part of) the simultaneity requirement, which Kallmeyer (2009: 65) divides into the following two distinct criteria:

(202) Simultaneity requirement (SIM), as per Kallmeyer (2009: 65):

a. [O]ne tree from an instance of an elementary tree set cannot be substituted or adjoined into another tree from the same set instance.

b. [T]wo tree sets cannot be interleaved, i.e., they must be added one after the other. Consequently, derivation trees [...] are not allowed [... if] there is a cycle in the

[151] The arguments put forth by Longenbaugh (2014: Ch. 3) will be countered explicitly in section 5.4.3.

order of adding the tree sets.

On the one hand, SIM plays no relevant role in L-TAGs because its requirements are always trivially satisfied. Since L-TAGs are 1-MC-TAGs, all the sets of trees they use are singletons, so that no combination of two trees belonging to the same set is possible. On the other hand, SIM is important in *h*-MC-TAGs where $h \geq 2$, because it can be checked by considering the nodes of a derivation tree that are at the same depth as a single generation (Krivochen 2023c: 8). As a consequence, such nodes are understood to be combined with their mother(s) either all at once or in an uninterrupted sequence of steps.

SIM in (202) is a restriction that must hold true not only for TL-MC-TAGs, but also for a specific superset of MC-TAGs called Set-Local MC-TAGs (henceforth SL-MC-TAGs). In SL-MC-TAGs, all the trees belonging to the same multi-component must combine to "the same elementary tree set, not necessarily to a unique tree and not necessarily to all the trees in this elementary tree set" (Boullier & Sagot 2011: Sec. 2.2). This can be determined by means of an informal definition inspired by (201):

(203) Set-locality requirement: An MC-TAG is Set-Local if, in any derivation tree it generates, each node (standing for an elementary tree) that belongs to a single multi-component set Σ_2 is a daughter of one of the nodes (standing for distinct elementary trees) that belong to a different multi-component set Σ_1.

The requirement in (203) follows SIM by establishing that all the elementary trees which are part of the same multi-component must belong to the same generation in a derivation tree, i.e. they must share the same degree of kinship with respect to the "nearest common ancestor" (Nesson 2009: 85-86, Nesson & Shieber 2009: 95). The existence of a single common ancestor for any MC-TAG is often not given a lot of attention in the literature (as in the just mentioned works addressing it), and yet it is really important because it affirms that any derivation tree generated by any conceivable TAG must be single-rooted just like its derived and elementary trees must be (see again the discussion and references in section 5.1.1). The following definition elaborates on the statement that Kallmeyer (2009: 57) provides in this respect:

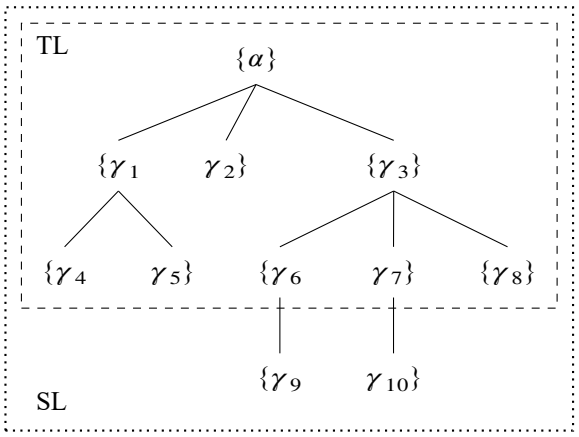
(204) Single Mother Condition for TAG derivation trees: Each TAG derivation must start "from an initial tree α being in a singleton set". Therefore, the number of singleton sets of trees involved in each TAG derivation must be ≥ 1.

It is now possible to clarify and summarize the requirements discussed so far by discussing the derivation tree in (205), where the following sets of trees are involved:

$\Sigma_0=\{\alpha\}$; $\Sigma_1=\{\gamma_1, \gamma_2\}$; $\Sigma_2=\{\gamma_3\}$; $\Sigma_3=\{\gamma_4, \gamma_5\}$; $\Sigma_4=\{\gamma_6, \gamma_7\}$; $\Sigma_5=\{\gamma_8\}$; $\Sigma_6=\{\gamma_9, \gamma_{10}\}$.

For the present concern, it is not important if the elementary trees just listed are initial or auxiliary trees, as long as they $\notin \Sigma_0$. Therefore, all the solid edges in (205) and in the alternate derivation trees that will follow indicate unspecified TAG operations, instead of specific occurrences of adjoining.

(205)



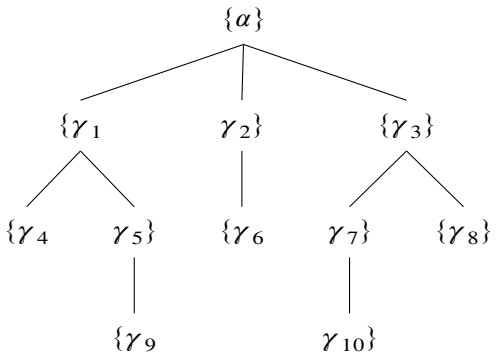
As the dotted box clarifies, the whole derivation in (205) complies with set-locality, while the dashed box highlights that the daughters and grandchildren of the root node α also follow tree-locality (which is a stricter requirement that is consistent with set-locality). These facts are confirmed by exploring the derivation tree by means of the breadth-first level-order traversal already described in section 5.1.4, where Gorn addresses were introduced. In the case of (205), it would lead to the following sequence: (α, γ_1, γ_2, γ_3, γ_4, γ_5, γ_6, γ_7, γ_8, γ_9, γ_{10}). As the curly brackets added at each node show, once a member of a set of trees Σ is added to the derivation tree in (205), all the other members of Σ are also computed before moving to the next set of trees.[152]

MC-TAGs whose derivation trees do not consistently follow either the Tree-Locality Requirement in (201) or the more relaxed set-locality restriction in (203) are said to be Non-

[152] The readers who want to be rigorous about it will observe that one can formulate isomorphic alternatives to (205) where at least one of γ_3 and γ_8 is neither the first nor the last node in its own generation. Hence, it could be that the SIM requirement in (202b) is slightly too strict, as it fails to capture the well-formedness of such isomorphic alternatives where Σ_1 is interleaved with Σ_2 or Σ_4 crosses Σ_5. This issue is not important here, and the interested readers are redirected to Krivochen (2023b: Ch. 3) and references therein for pertinent discussion.

Local MC-TAGs (henceforth NL-MC-TAGs). As Lichte (2015: 134) remarks, NL-MC-TAGs can be distinguished depending on their degree of compliance with SIM in (202). Let us see how by generating a new derivation tree where the same elementary trees used in (205) are combined anew. The result in (206) must be generated by a NL-MC-TAG because two of the sets of trees involved violate set-locality.

(206)

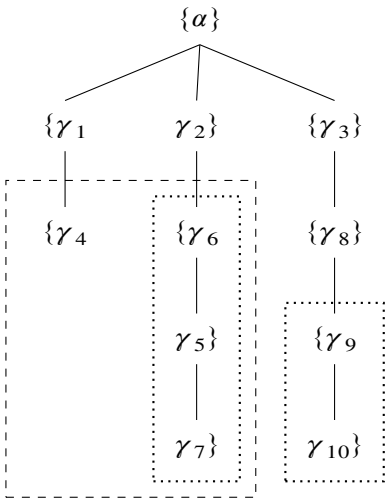


On the one hand, the two members of Σ_4, namely γ_6 and γ_7, break set-locality because they combine with a member of Σ_1 and with the only member of Σ_2, respectively. On the other hand, the two members of Σ_6, i.e. γ_9 and γ_{10}, do the same by being respectively combined with a member of Σ_3 and with a member of Σ_4. However, no SIM requirement in (202) is violated, as no elementary tree combines with another component belonging to the same set, and all the members of a set of trees are computed before moving to the next, just like in (205).

Things change in the equally non-local derivation shown below in (207), where the same elementary trees used to derive (205) and (206) are involved. On the one hand, the dotted boxes illustrate where (202a) is violated. In fact, in (207) a member of Σ_4 eventually combines to the other; ditto for the members of Σ_6. On the other hand, the dashed box indicates that both Σ_3 and Σ_4 violate (202b) because the two members of each of such sets belong to different generations. As a result of all these observations it follows that, in (207), the members of Σ_4 do not comply with SIM at all.

With respect to weak and strong generative powers, TL-MC-TAGs are well-known for being weakly equivalent to L-TAGs while having an increased strong power. In fact, it is for this very reason that Frank (2002: Sec. 5.6, 5.7, 6.1) uses a TL-MC-TAG to account for syntactic phenomena that L-TAGs fail to capture. As far as SL- and NL-MC-TAGs are concerned, Weir (1988) shows that both their weak and strong generative powers are increased

(207)



with respect to L-TAGs, so that the following weak equivalences hold:[153]

(208) L-TAGs = TL-MC-TAGs < SL-MC-TAGs < NL-MC-TAGs

One may therefore argue that the discussion done so far with respect to SL- and NL-MC-TAGs is not relevant for the concerns of natural language syntax, as section 5.1.2 has shown that the weak generative power of L-TAGs is the perfect fit for the purpose (Frank 2004: 675-676). However, section 5.3.3 will illustrate that both set-locality and non-locality have played a relevant role in the TAG literature concerned with natural language syntax, and section 5.4 will elaborate on a formalism (to be introduced in section 5.3.3) where a highly constrained form of non-locality is eventually allowed. Before introducing such aspects, section 5.3.2 will discuss a couple of further assumptions of the MC-TAG approach to be adopted and how they can be reconciled with GB/Minimalist inquiries.

5.3.2 On the interaction between Merge and TAG operations, pt. 2

As Longenbaugh (2014: 46) observes, the adoption of an h-MC-TAG where $h \geq 2$ brings to the table "a host of new questions concerning what sorts of limits to impose on, as a small example, the cardinality of these sets and the structure of the trees within them". Although the adjusted versions of FTH, CETM and the theta-criterion presented from page 270 on already

[153] Joshi, Kallmeyer & Romero (2003: 245) argue that a 2-SL-MC-TAG is weakly equivalent to a 2-TL-MC-TAG if each non-singleton set of elementary trees in the former includes a degenerate tree δ. However, both the previous and upcoming discussions will argue that degenerate trees are problematic and that TAG alternatives which dispense with them are more adequate.

address the latter concern, one can also benefit from the following three important generalizations, which are elaborated on quoted parts from Chen-Main & Joshi (2007: Sec. 6):[154]

(209) Common MC-TAG tenets:

a. In any *h*-"MC-TAG analysis" where the maximum arity *h* is assumed to be > 1, "a two-component set is typically used", i.e. $h = 2$.

b. "One of the trees [belonging to a single multi-component] is often a very small piece of structure that corresponds to the «base position», surface position, or scope position of a single [phonologically realized] element".[155]

c. In sets of trees which are not singletons, the "auxiliary tree components typically have elements with phonological content only on one side of the foot [node]", in case they are present.

The adoption of (209a) together with the lexicalization requirement for *h*-MC-TAGs provided by Kallmeyer (2009: 80) and Joshi (2014: 152) has the advantage of making the number of nodes in a derivation tree bounded and relatively small. In fact, it is subsequently "always linear in the input length. If [*h*] is the maximal number of trees per tree set in the grammar, then every derivation tree for an input of length *n* has no more than [*h*]*n* nodes" (Kallmeyer 2009: 80, italics in the original).[156] For these reasons, it will be assumed from now on that $h \geq 2$. Although the starting premises lead to assume that $h = 2$, the refinements in section 5.4 will lay the foundation for an MC-TAG approach where $h = 3$.

Another aspect to underline is that (209c) is explicitly expected to hold only if an auxiliary tree is actually present in a non-singleton set of trees. Chen-Main & Joshi (2007: Sec. 6) do not specify this for reasons that will be made clear in section 5.4.1 (see in particular footnote 172), but it is really important to highlight it here. In fact, (209c) will not be relevant for many multi-components to be discussed in this chapter, as they will exclusively include initial trees where the root node of one of them corresponds to a leaf node of the other. The latter situation can be illustrated by means of (210) on the facing page:

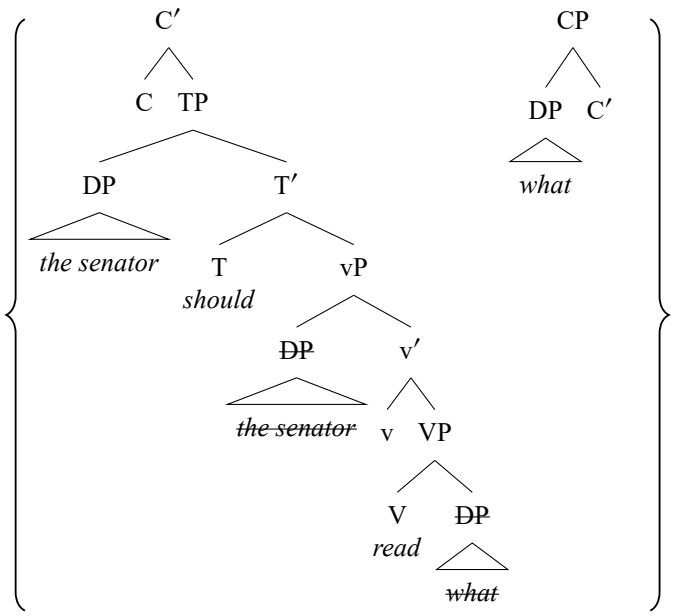
It is easy to see that, if the two components in (210) formed a single elementary tree,

[154] Chen-Main & Joshi (2007: Sec. 6) actually collapse (209a) and (209b) together, so that they end up claiming that the generalizations they formulate are only two.

[155] In (209)b, *the base position* arguably refers to what generative syntactician address as the *in-situ* position, i.e. where a syntactic object is assumed to be base-generated, whereas the *surface position* denotes the locus in which a displaced object is overtly realized.

[156] The original letter *k* has been substituted with *h* so as to make the quotation consistent with the rest of the present text.

(210)



they would provide a single structure that could be adopted to account for the sentence *What should the senator read?*. This is by no means a random fact, as the TAG approach to be proposed in section 5.4 will eventually adhere to the following assumption:

(211) Structural equivalence between a multi-component and a single legitimate elementary tree: In a 2-MC-TAG, any two distinct members γ_1 and $\gamma_2 \in \Sigma$ (where Σ is a multi-component set) are well-formed if a substitution of the one to the other would give rise to a derived tree that is structurally equivalent to a potentially legitimate elementary tree $\gamma \in \Sigma'$, where Σ' would be a singleton set.

The principle in (211) is "consistent with a characterisation of structure composition as graph union" (Krivochen 2023d: 339-340) and follows Chen-Main & Joshi (2008: 11) in explicitly establishing a «top» and a «bottom» tree for each multi-component, so that the foot node of the former always dominates the root of the latter in each derived tree. In the case of (210), the component rooted at C′ is the bottom tree, whereas the component rooted at CP is the top counterpart. However, (211) raises some concerns nonetheless. Why should it hold true? And how can multi-components following it, e.g. (210), be reconciled with common generative assumptions? And why should the node that the two trees in (210) share be an intermediate projection, which is hard to reconcile this with Bare Phrase Structure (Chomsky 1993) or with other Minimalist assumptions (see Krivochen 2023b: Sec. 4.1)?

On the one hand, a refinement that tackles the latter issue must wait until section 5.4.4. The discussion preceding that section will present a TAG approach where bar-nodes play a central role. This is to ensure consistency with previous TAG proposals adopting GB/Minimalist assumptions, e.g. Frank (2002, 2004, 2006, 2010, 2013), Storoshenko (2006), Frank & Storoshenko (2012), Longenbaugh (2014) and Georgi (2019). On the other hand, the answer to the first two questions of the previous paragraph can be discussed without further ado, as the principle in (211) *de facto* establishes that multi-components must be structurally equivalent to the splitting of single well-formed extended projections.

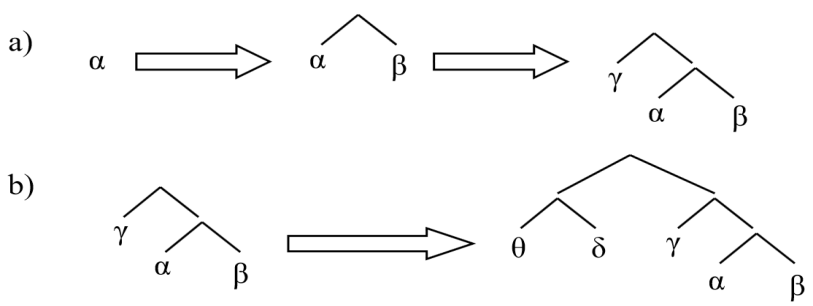
The principle in (211) can thus be seen as a novel application of the concept of *non-monotonic derivations* in the sense of Heck (2016, 2023). To illustrate the proposal, it is important to start by pointing out that Heck's idea of non-monotonicity does not coincide with the one proposed by Uriagereka (1999, 2012: Ch. 1, 4), but rather represents a superset of cases that also encompass what is discussed by the latter author. In fact, the comparison between the two types of non-monotonicity can start from the following quotation from Krivochen (2018b: Sec. 3.1), where the view *à la* Uriagereka is neatly summarized (italics in the original):

> In a derivational theory like the Minimalist Program, in which structure is built in a dedicated syntactic component by means of discrete combinatorics, we need to distinguish between operations that extend the phrase marker uniformly introducing a single terminal at a time (thus producing a series of structures head, non-head; see Chomsky 2013, 2015; Rizzi 2016) and operations which extend the phrase marker by introducing not a terminal, but a complex object (itself derived by the combinatoric operation). The first case [...] is referred to as *monotonic Merge* because the phrase marker grows uniformly (a single terminal at a time); the second [...] is referred to as *non-monotonic Merge*, because although the phrase marker does grow at every point, it does not grow uniformly due to the introduction of both terminals and nonterminals in the derivational space[.]

Krivochen (2018b: Sec. 3.1) further illustrates what has just been quoted by means of the representations on the next page, where (212a) and (212b) are examples of monotonic and non-monotonic Merge, respectively (see also Kosta & Krivochen 2012: 400-401).

Both alternatives depicted in (212) rely on a principle already introduced in section 3.2.4 and slightly reformulated here for convenience (see a.o. G. Müller 2023 for further discus-

(212)



sion):

(213) Strict Cycle Condition (SCC):
Only the head of the present root can have features that trigger Merge operations.[157]

The necessity of complex specifiers derived as in (212b), namely by means of non-monotonic applications of Merge that comply with the SCC in (213), lead Heck (2016: Sec. 2.4.2, 2023: 307-309) to follow Uriagereka (1999) in arguing in favor of the parallel construction of different syntactic objects during a single derivation. What keeps them separate is the further assumption of different spaces where they are built up and held ready, called *workspaces*. Although Heck (2016: 29) defines them informally as "storage devices", Krivochen (2018b, 2023d) observes that the notion of *workspace* is understood in several ways due to the lack of a shared and detailed definition. For the present purposes, the one put forth by G. Müller (2023: 23 fn. 20) will be taken as a starting point:

(214) The workspace of a derivation comprises all trees generated thus far in the derivation, outside of the current tree, and all lexical items in the numeration.

The definition in (214) must be regarded as provisional for three reasons. The first is that it departs from Chomsky et al. (2023) and related works, where both trees and numerations are dispensed with (for discussion with respect to the latter, see again section 5.1.5). It must be noted however that having the numeration as part of the workspace not only allows the former to be subintended while talking about the latter, but also permits to dispense with the concept of numeration altogether (if necessary. See Bruening 2021b,c for arguments favoring the

[157] Heck (2016: 11 ff.) further expands the breadth of the SCC in (213) by arguing that it must apply to any syntactic operation, be it Agree or some form of Merge, even though his discussion is limited to "what concerns the syntax of arguments" (Heck 2016: 23). This detail is not relevant here, and the interested readers are redirected to the cited works for further discussion.

necessity of some numeration). In fact, the lexical items that the latter is supposed to contain can be argued to be independently included in the workspace defined by G. Müller (2023: 23 fn. 20). Moreover, understanding the workspace as a repository for both lexical items and complex syntactic objects makes it possible to argue that they are equally accessible to Merge operations, even though the former are combined monotonically while the latter are not.[158]

A second reason to hold (214) as provisional is that it departs in a twofold manner from the way in which Heck understands the workspaces. On the one hand, Heck identifies the structure being derived (i.e. the *current tree* in 214) as a workspace (see e.g. the representations in Heck 2023: 308-311), while G. Müller (2023: 23 fn. 20) keeps it as a separate theoretical object. On the other hand, Heck assumes (when needed) two further separate workspaces, respectively posited for the derivations of a complex head and of a complex non-head. By following the quotation from Krivochen (2018b: Sec. 3.1) on page 280, the former is responsible for the creation of a sub-structure growing monotonically, whereas non-monotonic applications of Merge are allowed in the latter (see Heck 2023: 312, 314 for illustrations).

Leaving the technicalities for future research, one can however reconcile the two views by accepting the possibility that a workspace can contain within itself the current tree as well as other workspaces. In addition to (or instead of) this option, the two notions of *workspace* used by Heck and G. Müller (2023: 23 fn. 20) can be addressed with different names to avoid ambiguities. For the remainder of this chapter, *workspace* will indicate something that elaborates on (214). By comparing the latter to a metaphorical hospital, the workspaces discussed by Heck will be addressed as the *wards* of a workspace *à la* G. Müller (2023).[159]

The last issue that (214) raises is arguably the most serious. In fact, if Uriagereka (1999) and Heck (2016: Sec. 2.4.2, 2023: 307-309) are right in arguing (among others) for the parallel construction of different syntactic objects, how can one decide which tree must be held as the "current" one in (214)? Given that (214) tries to distinguish it so as to make it the reference point in a derivation driven by Merge, an MC-TAG perspective on the matter provides a straigthforward solution to this issue. In fact, the TAG analogue which assigns a similar

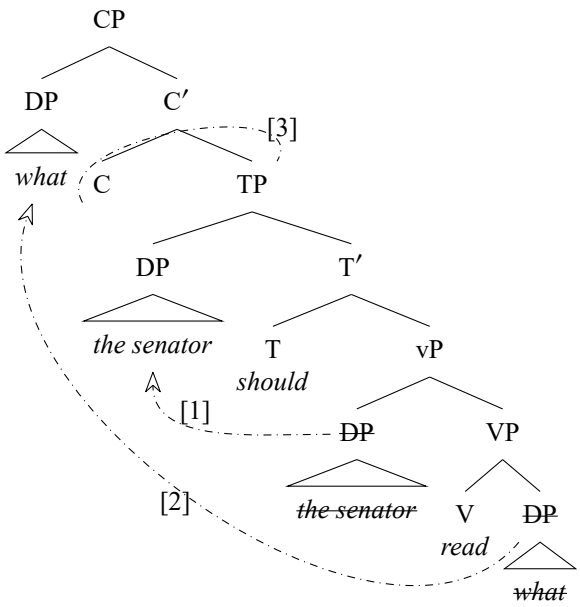
158 The claim made in the main text can be put on a par with a similar remark made by Krivochen (2023d: 327). However, the parallel cannot go further, because the latter provides a thorougly detailed definition of workspaces that is based on a partially different set of premises from those discussed in the present work.

159 The wards that have been defined based on Heck (2016, 2023) partly overlap with the *neighborhoods* discussed by Krivochen (2023d). However, the two notions do not coincide, as wards presuppose two or more distinct regions with different topological properties within the same workspace (Krivochen 2023d: 345 ff.). See section 5.4.3 for further discussion.

centrality to a single object of the derivation is the Single Mother Condition for derivation trees, introduced in (204) on page 274. Hence, the definition of workspace provided by G. Müller (2023: 23 fn. 20) can be adapted to a TAG derivation by arguing that the current tree must be nothing but the root of the derivation tree. By doing so, there is no longer a need to keep this initial tree separate from the workspace as in (214), since it can be intended as a ward of the workspace which is independently singled out by the Single Mother Condition for derivation trees.

With these theoretical premises in mind, it is now possible to illustrate the way in which Heck (2016: 7-8, 2023: Sec. 3.3) extends the already discussed concept of non-monotonicity, and how it can be adapted to the present needs. While Uriagereka (1999, 2012: Ch. 1, 4) limits it to the growth of a syntactic structure at a non-constant pace, Heck (2016, 2023) pushes the argument further by observing that a derivation is non-monotonic also if there exists at least one derivational step where it does not grow at all, but rather *shrinks*.[160] This can be illustrated with (215) on the current page, namely a non-monotonic derivation (where *v*P is not a phase, as discussed on page 230) that arguably leads to the eventual formation of the multi-component set of trees illustrated in (210).

(215)



[160] Thus, the shrunk structures proposed by Heck (2016, 2023) have nothing to do with the ones discussed in Abels & Neeleman (2009: Sec. 4.4.1), as the latter are called as such due to a formal comparison with the ones proposed by Cinque (2005, 2023). See again section 3.2.4, and in particular footnote 75.

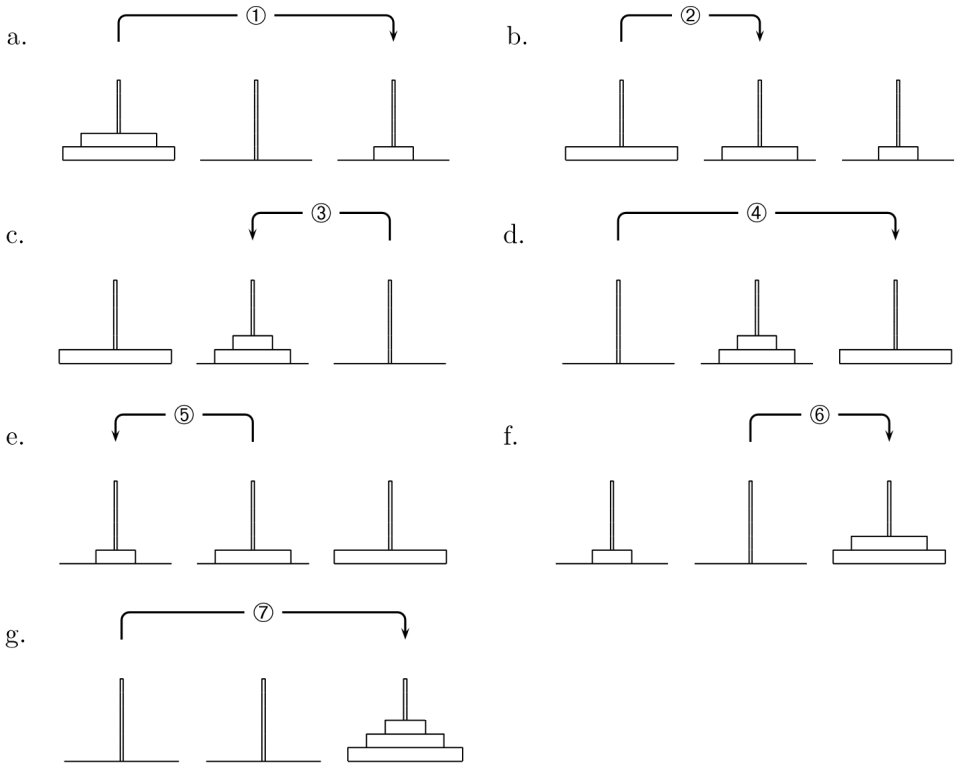
In (215), a single tree is derived by means of the usual Merge operations (applying non-monotonically in the sense of Uriagereka 1999, 2012). The attention must be directed towards the point of the derivation where the C head is externally merged to the structure. If one uses *n* to refer to such a merger, then *n* + 1 can indicate the subsequent derivational step where the *wh*-element undergoes A′-movement. As the arc crossing the C′ node in (215) aims to suggest, the crucial operation where the DP is removed from Spec,CP applies at *n* + 2. Although this shrinkage seems to restore the structure as it was back at *n*, there are still some important differences. In fact, all the features that have been checked at *n* + 1 remain as such at *n* + 2, so that the head C "«remembers» that there is the category" DP outside of the structure to which C is merged (Heck 2016: 32).

The last point represents the hallmark of the approach, as it allows Heck (2016, 2023) to reconcile the Strict Cycle Condition in (213) with head-movement and other counter-cyclic operations that are otherwise inexplicable in a derivational, narrow-syntactic account. To help the readers understand the rationale of his approach, Heck (2016: Sec. 2.4.1) exploits the Tower of Hanoi as an analogy. The latter is a puzzle that assumes three poles and a pile of disks that are ordered from the largest (at the bottom) to the smallest (on top). The goal is to find a way to rearrange the disks, which at first are stacked in the leftmost pole, into the rightmost "in an order preserving way, if necessary by using" the intermediate pole "as an auxiliary storage" (Heck 2016: 28). The illustration in (216) on the facing page follows Heck (2016: 28-29) in showing the shortest sequence of steps that the player must follow to correctly solve the puzzle when three disks are involved.

Both the Tower of Hanoi and syntactic derivations posit for the manipulable objects an established hierarchy which can never be violated. However, Heck (2016: 29) highlights that, in order to construct a well-formed derived object, "the derivation [be it (215) or (216)] sometimes has to destroy an intermediate representation that seemed to be a step towards the target configuration". In addition to that, Heck (2016: 29) observes that the following three rules further strengthen the parallel between the Tower of Hanoi (which he abbreviates as ToH) and the non-monotonic derivations that he assumes:

> [T]he first constraint of ToH (only one disk can be affected per move) reflects the general assumption that rule application in syntactic derivations is sequential and not parallel; the second constraint (only disks at the top are accessible) is an analog to the MLC [see footnote 143 on page 243]; and the third constraint

(216)



(disks may only be added to the top of a pole) corresponds to the SCC [in (213) on page 281].

The discussion made so far allows to point out another contrast. In a non-monotonic derivation *à la* Uriagereka (1999, 2012), two syntactic objects occupying two distinct wards of the same workspace must be unrelated, as they are still to be externally merged with each other. On the contrary, in a non-monotonic derivation *à la* Heck (2016, 2023) it does not have to be that way. The same two syntactic objects can yet be unrelated because of what has just been said, or they can already be connected because of their former participation to the same structure. If multple objects are evicted from the same structure, they can (but need not) be combined with one another if they are then sent to the same ward of the workspace, and in some subsequent derivational step(s) they must be remerged to the structure from which they were previously detatched.

After these observations, the concept of multi-component sets of trees sneaks in easily, as its members can be seen as the outcome of a shrunk structure. Hence, any multi-component ultimately indicates that two (or more) distinct wards in a workspace are related (as opposed to further syntactic objects contained in other wards) because they used to belong to the

same structure in some previous derivational step. This observation suggests that *h*, namely the maximum arity for each set of trees in the adopted MC-TAG (Boullier & Sagot 2011: Sec. 2.2), can be ruled by something like the following:

(217) In an *h*-MC-TAG (where $h \geq 1$) there can be at most $h - 1$ syntactic objects that do not remerge directly to the elementary tree from which they have been temporarily removed.

The principle in (217) is formulated by assuming a model inspired by Frank (2002, 2004, 2006, 2010, 2013), where structures can no longer be merged and shrunk once the derivation starts to deal with TAG operations (see again section 5.1.5). In fact, allowing a derived tree generated by a TAG to shrink would arguably complicate the matters in a way which goes beyond the scope of the present work.

It is now time to focus on the aspects that make the TAG approach to be adopted here diverge from the non-monotonic derivations proposed by Heck (2016, 2023). The discussion can start from the multi-component set in (210), which is formed by shrinking (215). The derivation presented here agrees with Heck (2016: 38) in arguing that a phrasal object "must not leave a trace/copy" behind in the structure from which it is removed. However, Heck (2016, 2023) would not allow the shrinking in (215), as will be discussed shortly. But even if he did, the CP node would not appear as the mother node of the severed DP *what*, and neither would C′, *contra* (210). The retainment of both CP and C′ in such a syntactic object can be defended by arguing that such nodes are not only a relic of the structural relations that the *wh*-element entertained when merged with the projection of C, but also stand for the features that were contextually checked. In any case, this issue will be taken up again in sections 5.4.3 and 5.4.4.

The present account and the one proposed by Heck (2016, 2023) can also be said to agree that, given a certain derivational step *n*, the current root *R* of the structure is evicted from it at the step $n + 1$ if a daughter of *R* is simultaneously removed from the structure. This fact cannot be surprising, because it is what makes the structure eventually shrink. However, Heck (2016, 2023) assumes that the latter principle applies only if the daughter of *R* is a head, whereas it is not necessarily so in a case like (215), where such a daughter is a non-head (see in particular Heck 2016: 37-38). In fact, in his approach a non-head can be detatched either if it is c-commanded by a head or if the structure has already shrunk in a previous step. It has to be so because the shrinkage process must start from a head immediately dominated by *R*,

for a twofold reason. On the one hand, only heads are assumed to bear the attracting features permitting the removal from the structure of both themselves and some other element they c-command. On the other hand, only a daughter of the current root *R* can discharge its features without violating the Strict Cycle Condition in (213).

A 2-MC-TAG whose elementary trees stem from non-monotonic derivations *à la* Heck (2016, 2023) must depart from what has just been discussed, for two related motivations. The first actually holds if one assumes the strong compartimentalization between Merge and TAG operations discussed in section 5.1.5, as it corresponds to the inherently counter-cyclic nature of the TAG operations pointed out by Frank & Hunter (2021: Sec. 1).

In fact, both substitution and adjoining can lead to the rewriting of nodes which do not coincide with the root of a structure (nor with the nodes that the root immediately dominates, insofar a tree larger than the previous one is created, Safir 2008: 337-338), *contra* the Extension Condition (see Chomsky 2014–1995: 227 ff.). Thus, the claim that the SCC in (213) is relevant in a TAG derivation only holds if one supports the weak compartimentalization presented in section 5.1.5, and an assessment of whether it is better to support it goes beyond the scope of the present work.

A second reason to conceive non-monotonic derivations differently from Heck (2016, 2023) is that it is in the interest of an adequate TAG approach to allow the operations leading from (215) to (210), as section 5.4 will show. In fact, the formation of a multi-component set of trees Σ cannot be done for the benefit of what Σ already contains, as (211) ensures the existence of an equivalent singleton counterpart Σ' containing a single legitimate elementary tree, which corresponds to the union of the two graphs in Σ. The formation of any multi-component rather follows from the Single Mother Condition for derivation trees, because the splitting of a structure is done in anticipation of (at least) one distinct tree that has been derived in parallel in a different ward of the workspace (or, differently put, in another set of trees).

The TAG account presented here agrees with Heck (2016, 2023) when it comes to the removal of a head or a projection which are not immediately dominated by the current root *R*. On the one hand, a projection HP must be maintained after its head *H* is evicted if "there is a specifier within HP" of "if HP is the complement to another head" (Heck 2023: 310), as in both cases the removal of HP would severe connections in the tree that one would want to keep. On the other hand, the removal of an embedded phrasal object "leaves behind an

impoverished category, bereft of any substantial morpho-syntactic features, which may just serve as a mark of reconstruction" (Heck 2016: 12), while leaving the rest of the structure intact. Heck (2016: 32-33) further argues that the latter assumption must extend to any type of phrasal movement. In fact, such an operation can be broken down into two distinct steps, where a phrase is first removed from the structure and sent to a different ward of the workspace, and then remerged when needed to the same (but further grown) structure.

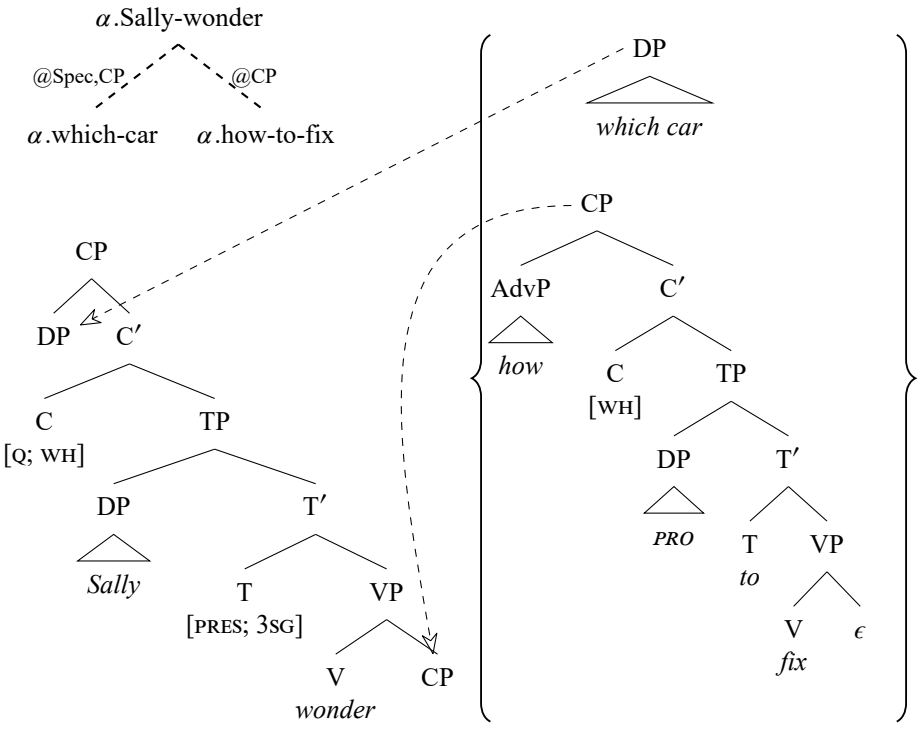
In principle, a TAG approach benefiting from the non-monotonic derivations *à la* Heck (2016, 2023) does not have to be committed to his view with respect to phrasal movement. In fact, one could argue that within a single elementary tree phrasal movement is driven by Internal Merge, while structure shrinkages are relegated to the cases where the formation of multi-component sets of trees is required for syntactic movements that span across distinct trees. However, the latter choice leads to the proliferation of the principles that guide syntax, which is in turn a theoretically unwelcome possibility. In this respect, Heck (2016, 2023) shows that non-monotonic derivations not only allow to handle different syntactic phenomena by means of a single principle, but also effectively circumvent some of the problems raised by the copy theory of movement (see a.o. Witkoś 2021, Gärtner 2022, Krivochen 2023d).

Although a thorough discussion of the reasons favoring one theoretical choice over the other go beyond the scope of the present work, it is worth underlining nonetheless that non-monotonic derivations represent a good theoretical premise for an adequate treatment of long distance displacements. Given that this aspect will turn out to be crucial for the ultimate analysis of U20 that will be presented in section 5.5.1, it is better to linger a little longer on this point. Therefore, the removal of an embedded phrasal object that leads to the formation of a multi-component can be illustrated by giving a novel 2-MC-TAG analysis of the sentence *Which car does Sally wonder how to fix?*, whose account provided by Frank (2002) has already been discussed and questioned in section 5.3.1. The alternative in (218) elaborates the structures suggested by Frank (2002) (whence the absence of *v*P from the representation) without resorting to degenerate trees, as they are hard to reconcile with non-monotonic derivations *à la* Heck (2016, 2023). In fact, the structures proposed by the latter author shrink only if (at least) one non-terminal leaf has been removed, so that the present MC-TAG adaptation of these premises must assume the formation of multi-components whose members cannot be degenerate trees.

The derivation in (218) not only follows these desiderata, but is also arguably better than

the previous account, as it avoids the aforementioned problems of overgeneration while still being compatible with the addition of further intermediate *wh*-islands, as in *Which car does Sally wonder* who to ask *how to fix?*, a sentence which Frank (2002: 226-227) discusses following Cinque (1990a). The multi-component in (218) departs from those in Frank (2002: 226-227) because it has two other forerunners. The first is Kroch (1989: 79 ff.), who takes cases like (218) as further evidence in favor of the multi-component sets first assumed in Kroch (1987). The second is Scheffler (2003: Sec. 3.2), who observes that one does not need to follow Kroch (1989: 79 ff.) in assuming that the extracted *wh*-phrase anchors an auxiliary tree (though see footnote 162 on the next page), as the latter can be re-analyzed as an initial tree substituting at a specifier in the left periphery of the main clause.[161] The latter dichotomy follows from an issue that will be discussed from section 5.4.1 on, namely the problematic way in which many TAG approaches deal with the structure that dominates the clausal core.

(218)



In (218), *which car* can be argued to be severed from its *in-situ* position in *α.how-to-fix*

[161] It must be admitted that Scheffler (2003: Sec. 3) ultimately dismisses an MC-TAG approach to long extractions *à la* Kroch (1987, 1989) by favoring instead an L-TAG analysis that is further elaborated in Kallmeyer & Scheffler (2004). The main reason that makes Scheffler (2003: Sec. 3.2) and Kallmeyer & Scheffler (2004: 34) reject Kroch's multi-component sets of trees is that they violate CETM as defined on page 214, based on Frank (2002). However, both the non-monotonic derivations *à la* Heck (2016, 2023) and the premises supporting the MC-TAG adaptation of CETM on page 270 show that such an objection is not as compelling as it first appears.

due to the failure of the latter structure to check its [WH] feature.[162] In fact, (218) agrees with Frank (2002: Sec. 5.6) in arguing that a CP cannot have multiple specifiers in the elementary trees assumed for English. It is also important to observe that if *α.how-to-fix* did not shrink, it would fail to be a well-formed derived tree of a TAG derivation as per (211) because of the unchecked [WH] feature of *which car*. Thus, the splitting that leads to the formation of the multi-component in (218) is a last resort, allowed by the maximum arity of the MC-TAG as established in (217), and performed in an attempt to comply with the Maximal Checking Principle presented in (169) on page 233. Although this move is not decisive per se, it is made in the hope of a convenient complementary structure which allows to perform the desired repair. And in fact, *α.Sally-wonder* suits the purpose by having *which car* substituted at its Spec,CP. However, a problem arises, as there is no a priori reason to posit a DP label for the Spec,CP in *α.Sally-wonder*. This is a look-ahead issue that allows the latter initial tree and *α.which-car* to combine with each other, but which must eventually be fixed. Section 5.4.4 will try to do this.

The representation in (218) also shows that no new copy is entailed in the non-monotonic operation that separates *which car* from *α.how-to-fix* and makes them members of the same multi-component in (257). No PP is left behind within the elementary tree in which the *wh*-element is first-merged, and the rest of *α.how-to-fix* is not affected by its removal. Thus, the previous presence of the object in such an initial tree can only be inferred from the selectional requirements that were satisfied and the features that were checked during the derivation of the same tree. Similarly, the PP would be different from other hypothetical objects of the workspace that have not been externally merged, because the former is θ-marked while the latter would not be.

For these reasons, the representation of *α.how-to-fix* in (218) bears a symbol which is commonly used in formal language theory to indicate an empty terminal node, i.e. ϵ. TAG works like Abeillé & Rambow (2000), Bleam (2000), Chen-Main & Joshi (2008), Chiang & Scheffler (2008), Kallmeyer (2009), Chen-Main, Bleam & Joshi (2012), and Lichte (2015)

[162] Of course, the problem stated in the main text does not arise in languages that allow multiple *wh*-fronting (see a.o. Stepanov & Tsai 2008: Sec. 6.2 and references therein). This type of crosslinguistic variation can be captured by stating different properties for the sets of trees in each language, and the ways in which they can be combined. In particular, a tentative TAG analysis would consist in arguing that in certain languages *wh*-phrases are allowed to anchor modifier auxiliary trees (as per Kroch 1989: 79 ff.), so that they can multiply adjoin at the matrix CP node. The presence vs. absence of Superiority effects (see again footnote 143 page 243) could be captured by having such auxiliary trees respectively involved in dependent vs. independent derivations *à la* Schabes & Shieber (1992, 1994), as discussed in section 5.2.1.

resort to it so as to fill the structural positions that generative syntacticians would argue to be occupied with a trace or a copy related to a displaced element (see a.o. Abels 2015, Gärtner 2022 and references therein). The representation in (218) diverges from this difference between the two syntactic strands (which are bridged in Krivochen & Padovan 2021) by arguing that ϵ must be intended as a truly empty terminal node regardless of one's theoretical background.

5.3.3 Delayed tree-locality

Chen-Main & Joshi (2007: Sec. 6) provide a summary of various syntactic facts that can be captured only by means of h-MC-TAGs where $h \geq 2$. Among those, the clitic climbing phenomena attested in Spanish and several other Romance languages represent a bone of contention in the TAG literature, because the range of (optional, Bleam 2000: 196) syntactic variation among structures with identical semantic interpretations defies the explanatory power of TAGs. This fact might appear unsurprising to the readers who are already aware of the several analyses provided in the generative tradition for the very same phenomenon (see a.o. Wurmbrand 2001: Ch. 3, 4 and the various sources mentioned in Villa-García 2019: Sec. 2).

The discussion that follows bears unavoidable commonalities with the latter proposals, but the TAG approaches to be discussed are not committed to a certain line of analysis in the generative tradition (say, a restructuring view *à la* Rizzi 1982) to the detriment of another (e.g., Kayne 1989 or Sportiche 1998: Ch. 4-7). Therefore, possible similarities among the TAG and the generative analyses will not be discussed in the main text, but will only be mentioned in footnotes 168 and 181 on pages 299 and 347, respectively.

As Bleam (2000: 195) summarizes, "clitic climbing [...] obtains when a clitic appears in a higher clause than [...] the one in which it is a thematic dependent". Although clitics "can climb out of non-tensed clauses [...] but not out of tensed" ones (Bleam 2000: 196), "[o]nly a subset of the verbs which take infinitival complements allow a clitic from the complement to climb [...]. The trigger/non-trigger verb distinction cuts across the categories of subject raising, subject control, object control, and exceptional case marking (ECM) or raising-to-object verbs" (Bleam 2000: 195). Finally, clitic climbing "appears to be unbounded since, as long as only triggering verbs intervene, the clitic can move over unboundedly-many clauses" (Bleam 2000: 196).

The relevant data are now reported, with clitics boldfaced. See Chen-Main, Bleam & Joshi (2012: 2) for (219), Bleam (2000: 197) for (220) and (221), and Chen-Main, Bleam & Joshi (2012: 8) for (222) and (223). The interested readers are redirected to Bleam (2000), Kulick (2000: Sec. 4), Chen-Main, Bleam & Joshi (2012) and references therein for further examples and discussion.

(219) a. *Mari quiere ver-**lo***
Mari wants to.see-3SG.M

b. *Mari **lo** quiere ver*
Mari 3SG.M wants to.see
'Mari wants to see it.'

(220) a. *Mari quiere permitir-**te** ver-**lo***
Mari wants to.permit-2SG to.see-3SG.M

b. **Mari **lo** quiere permitir-**te** ver*
Mari 3SG.M wants to.permit-2SG to.see

c. **Mari **te** quiere permitir-**lo** ver*
Mari 2SG wants to.permit-3SG.M to.see

d. *Mari **te** quiere permitir ver-**lo***
Mari 2SG wants to.permit to.see-3SG.M

e. *Mari **te** **lo** quiere permitir ver*
Mari 2SG 3SG.M wants to.permit to.see

f. *Mari quiere permitir-**te-lo** ver*
Mari wants to.permit-2SG-3SG.M to.see
'Mari wants to permit you to see it.'

(221) a. *Quiero mostrár-**te-los***
want.1SG to.show-2SG-3PL.M

b. ***Te** *quiero mostrar-**los***
2SG want.1SG to.show-3PL.M

c. ***Los** *quiero mostrar-**te***
3PL.M want.1SG to.show-2SG

d. ***Te los*** *quiero mostrar*
2SG 3PL.M want.1SG to.show
'I want to show them to you.'

(222) a. *Que quiere mostrar-**te** Juan?*
What wants to.show-2SG Juan

b. *Que **te** quiere mostrar Juan?*
What 2SG wants to.show Juan
'What does Juan want to show you?'

(223) a. *Luis suele comer-**las***
Luis tends to.eat-3PL.F

b. *Luis **las** suele comer*
Luis 3PL.F tends to.eat
'Luis tends to eat them.'

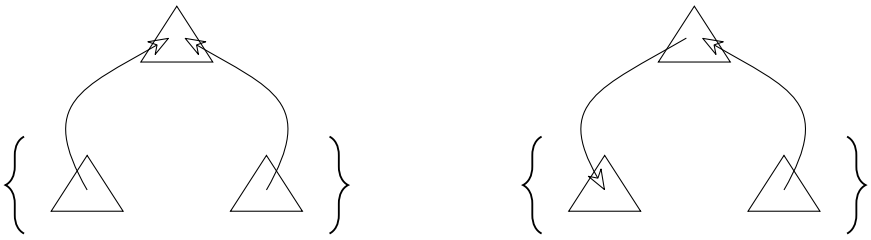
Three lines of TAG analyses have been pursued to address the issue. The first one is endorsed by Longenbaugh (2014: 46-47), where the burden of explanation is shifted to a post-syntactic component. This choice will not be discussed here, for a twofold reason. On the one hand, it is ultimately grounded on a general rejection of MC-TAGs that section 5.4.3 will eventually question. On the other hand, it also relies on legitimate considerations that are equally brought to the table by the second line of thought, represented by Kulick, Frank & Vijay-Shanker (2000), where it is argued that L-TAGs can capture the clitic climbing facts by revisiting how c-command rules the combinations among distinct elementary trees so as to generate the attested *derived* trees (see Kulick, Frank & Vijay-Shanker 2000: Sec. 5 for detailed discussion).

Kulick, Frank & Vijay-Shanker (2000) try to provide a satisfactory analysis of clitic climbing in L-TAG terms because they argue that Bleam (2000), i.e. the alternate 2-MC-TAG account they are aware of, fails to do so. The latter analysis starts from a twofold assumption. On the one hand, Bleam (2000: 203) holds that a clitic is encoded in the elementary tree anchored by the verb assigning a theta-role to it if the clitic itself is realized *in situ* in the derived tree. If on the other hand the clitic is eventually displaced in the derived tree, it is due to the fact that, at the start of the derivation, the verb and the clitic constitute distinct components included in the same set of trees (Bleam 2000: 206). Although the latter premise requires the

adoption of a 2-TL-MC-TAG (see again section 5.3.1), Bleam (2000: Sec. 8.5) shows that the locality restrictions imposed by TL-MC-TAGs are too severe to effectively account for the facts. Thus, Bleam (2000: Sec. 8.6) eventually decides to relax such constraints at the cost of adopting a SL-MC-TAG. However, such a move is arguably "not desirable" (Longenbaugh 2014: 47) because the weak generative power of SL-MC-TAGs is greater than both L-TAGs and TL-MC-TAGs (Weir 1988).

Kulick (2000: Ch. 4) points out another issue with respect to the premises made by Bleam (2000: 203, 206). Since a singleton set and a multi-component anchored by the same lexical verb are assumed to be independent objects of the formalism, the analysis cannot capture the similarities between the two types of structures, nor does it manage to explain why they should both be considered legitimate primitives. Although non-monotic derivations *à la* Heck (2016, 2023) provide the basis for a response to this objection (see again section 5.3.2), Kulick, Frank & Vijay-Shanker (2000: Sec. 4) follow Kulick (2000: Ch. 4) in emphasizing an even more serious issue. In fact, one needs the TAG derivation reported in (224b) on this page so as to make clitic climbing occurr together with long distance wh-movement as in (222b); ditto for clitic climbing co-occurring with raising, as in (223b). The alternatives in (224) adapt the representations in Kulick, Frank & Vijay-Shanker (2000: 41), so that triangles stand for the elementary trees involved in a derivation. All the arrows indicate unspecified TAG operations (instead of some instance of adjoining) and always point to the targeted trees.

(224) a. b.



The fact that (224b) must be a viable option is something that questions the adequacy of the whole analysis provided by Bleam (2000), because (224b) breaks the part of the simultaneity requirement preventing any two members included in the same set of trees from combining with one another, as previously stated in (202a) on page 273 by following Kallmeyer (2009: 65). To put it more succintly, (224b) is illegal to both TL- and SL-MC-TAGs, so that even the latter are seemingly incapable of accounting for the facts.

Chen-Main, Bleam & Joshi (2012: Appendix A) counterargue that a 2-SL-MC-TAG can

also account for the problematic (222b) and (223b) once the analysis in Bleam (2000) is slightly refined. They add a new assumption which states that, if an elementary tree γ anchored by a certain lexical verb features a subject in Spec,TP and/or a *wh*-element moved to Spec,CP, then the clitic selected by the same verb can not only be included in γ in its *in situ* position, but also raise past the verb itself so as to c-command the node to which the raising or trigger verbs are then adjoined.

However, Chen-Main, Bleam & Joshi (2012) end up favoring a third approach, which is the one that will be endorsed from now on. This alternative consists in adopting a type of 2-MC-TAG that Chiang & Scheffler (2008: Sec. 4) prove to be weakly equivalent to both L-TAGs and 2-TL-MC-TAGs. The fundamental property of this approach is that it completely dispenses with the simultaneity requirement in (202) on page 273, by hinging on what is usually addressed as *delayed tree-locality*, following Chiang & Scheffler (2008: Sec. 3).[163]

One should immediately point out that the notion of *tree-locality* intended by Chiang & Scheffler (2008), Han, Potter & Storoshenko (2010), Chen-Main (2012), Chen-Main, Bleam & Joshi (2012), and Chen-Main & Joshi (2014) is rather misleading, because it does not assume the Tree-Locality Requirement formulated in (201) on page 273, but rather the Single Mother Condition for TAG derivation trees presented in (204) on page 274. In fact, what matters in MC-TAGs of this kind is that all the trees included in the same multi-component be *eventually* tree-local (see Chiang & Scheffler 2008: 19), which is just another way of saying that they must share a single "common ancestor" in the sense of Nesson (2009: 85-86), Nesson & Shieber (2009: 95), and Kallmeyer (2009: 74).[164]

This apparent sublety leads to an important observation, namely that the Tree-Locality Requirement in (201) imposes irreflexive relations among the nodes it checks in a derivation tree, while the Single Mother Condition in (204) does not. As already discussed from page 273 on, the components belonging to the same unique set of trees Σ can comply with the Tree-Locality Requirement only if they all directly combine to a unique elementary tree

---

163 The type of MC-TAG discussed in the main text is not the only one which drops the simultaneity requirement. See Kallmeyer (2009), Nesson (2009: Ch. 3), and Nesson & Shieber (2009: Sec. 3) and references therein for surveys where such alternatives are discussed.

164 The present work will depart from Chiang & Scheffler (2008: 19) and subsequent works by addressing the formalism under discussion as k-*delayed* h-*MC-TAGs* (where *k* is the maximum number of delays allowed at any node of the derivation tree, while *h* is the maximum arity allowed in each set of trees), instead of k-*delayed TL-MC-TAGs*. Although the reference to tree-locality can also be due to the proven equivalence between *k*-delayed (TL-)MC-TAGs and TL-MC-TAGs in the strict sense (see again Chiang & Scheffler 2008: Sec. 4), the omission of *TL* makes the name of the formalism shorter and avoids the potentially misleading interpretation discussed in the main text.

which is different from any of them (and thus necessarily included in a different set of trees), as depicted in (224a) by adapting the representation in Kulick, Frank & Vijay-Shanker (2000: 41). On the contrary, the same components belonging to Σ can obey the Single Mother Condition for derivation trees as soon as one of the components of Σ is the common ancestor of all the other members of Σ. This is exactly what happens for example in (224b), where an elementary tree included in the multi-component eventually combines to the other by using a third elementary tree (belonging to a different set of trees) as a proxy.

The tree within the multi-component in (224b) which features only an incoming directed edge can thus be said to be local with respect to itself. In fact, it represents the ultimate domain of locality for the whole TAG derivation in (224b). However, this contravenes an assumption (204) on page 274 that is formulated as per Kallmeyer (2009: 57), namely that the root of a derivation tree must be in a singleton set. The explicit assumptions discussed by Kallmeyer (2009: 67, 74-75) show however that one can implicitly drop this detail as soon as the simultaneity requirement is dispensed with as well. Therefore, the Single Mother Condition in (204) can be reworked as follows:[165]

(225) Single Mother Condition for non-simultaneous *h*-MC-TAG derivation trees:
Each derivation must start from an initial tree α. The number of singleton sets of trees involved in each of such derivations is ≥ 0.

At this point, another problem seems to emerge. Given that this alternate MC-TAG approach dispenses with the whole simultaneity requirement stated by Kallmeyer (2009: 65), it must also allow derivations that are typical of NL-MC-TAGs. However, it has been reported in (208) on page 277 that the weak and strong generative powers of the latter are way greater than those of L-TAGs and TL-MC-TAGs (Weir 1988), so that one may wonder how the approach that Chen-Main, Bleam & Joshi (2012) apply to clitic climbing can eventually be weakly equivalent to that of an L-TAG or of a 2-TL-MC-TAG. This is made possible by the introduction of *delays*, which can be clarified by following Chiang & Scheffler (2008: 19) verbatim (italics in the original):

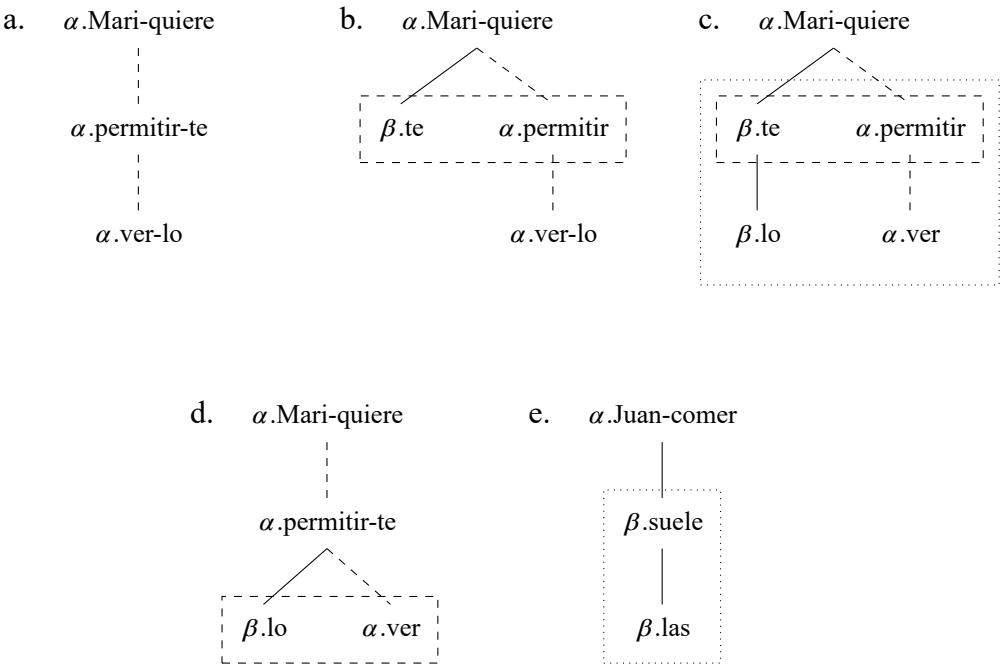
> A *k-delayed tree-local multicomponent TAG* is a multicomponent TAG whose derivations have the following property. Let the *destination* of an elementary tree set instance *S* be the lowest derivation node that dominates all the members

[165] As discussed from page 273 on, any L-TAG (i.e. any *h*-MC-TAG where $h = 1$) cannot be non-simultaneous. Hence, the value of *h* in (225) is necessarily ≥ 2.

> of *S*. Let the *delay* of *S* be the union of the paths from the destination down to each member of *S*, minus the destination itself. Then no derivation node can be a member of more than *k* delays. [...]
>
> Intuitively, this means that the members of an elementary tree set can adjoin into different trees, arriving at the same elementary tree (the destination) after some delay; and there can be at most *k* delays at any point in the derivation. (Note that this definition also allows one member of an elementary tree set to adjoin into another.)

In the hope that readers will forgive the wordplay, *k*-delayed MC-TAGs essentially make their derivation trees allow local portions of non-locality, so that all the derivation trees in (226) are legitimate under this approach. In particular, (226) on the current page accounts for the grammatical clitic climbing examples shown before in (220) and (223), so that the correspondence is between (220a) and (226a), (220d) and (226b), (220e) and (226c), (220f) and (226d), (223b) and (226e). All the trees in (226) adjust those shown by Chen-Main, Bleam & Joshi (2012: 5, 8) to the representational conventions adopted in this chapter.

(226)



Three clarifications are needed in order to fully appreciate (226). The first is that Chen-Main, Bleam & Joshi (2012) follow Bleam (2000) in assuming that at the start of the TAG derivation each to-be-displaced clitic is contained in an auxiliary tree whose foot node is to

the left of the clitic itself. This implies that all such trees are non-wrapping right auxiliaries (see again section 5.2.1). Thus, if a clitic *c* is stacked onto another clitic *c′* by adjoining the former at the root of the latter, *c* is eventually linearized to the right of *c′* (see the cited works for further illustration and discussion). What has just been said can be put in more concrete terms by referring to (226c) and equating *c* with $\beta.lo$ and *c′* with $\beta.te$. In the derivation tree in (226c), the former is a daughter of the latter, and the just mentioned properties of such trees dictate that the content of $\beta.lo$ must follow that of $\beta.te$ in the corresponding string.

The second point to highlight with respect to (226) is that the number of multi-components involved in a single MC-TAG derivation coincides with the number of delays in the corresponding derivation tree. The readers are invited to note that the present formalism does not limit the overall number of delays in the derivation tree, nor their lengths, but rather the ways in which they are allowed to interact with one another (Nesson 2009: 94). From (226) on, delays will be represented (whenever present) on derivation trees by means of boxes, as it is a rather standard convention among works using *k*-delayed *h*-MC-TAGs.

The present chapter will elaborate on this tradition by introducing a further distinction. On the one hand, a delay will be represented with a dashed box if it contains only and exclusively all the nodes standing for the components of a single set of trees Σ. On the other hand, delays in derivation trees will be indicated with dotted boxes if they also include a least one node $\notin \Sigma$.[166] The derivations in (226b)-(226d) show that the delays complying with the former property can satisfy the Tree-Locality Requirement, whereas the latter type of delays cannot, as they determine local portions on the derivation tree that would only be allowed in SL- or NL-MC-TAGs. In fact, (226c) bears a set-local delay, while the unique delay in (226e) is non-local.[167]

The last aspect that must be pointed out with respect to (226) is that any two members of a single dashed box necessarily target distinct nodes of the tree in which they are tree-local. This is due to the fact that in (226b)-(226d) all the multi-component sets of trees are made of an auxiliary tree, which thus must adjoin at an inner node of the target, and an initial

[166] It is worth noting that in *k*-delayed *h*-MC-TAGs the root of a derivation tree can never be part of a delay, as it must satisfy all the locality requirements that no other node in the same derivation tree has taken care of.

[167] However, the relation between dashed delays and tree-locality is bijective only in *h*-MC-TAGs where $h = 2$. Taking a 3-MC-TAG as an example of an *h*-MC-TAGs with $h > 2$, one can conceive a derivation where γ_1 and γ_2 are daugthers of the root α, whereas γ_3 is a daughter of γ_1 (or γ_2). It is easy to see that, if γ_1, γ_2 and γ_3 all belong to the same multi-component Σ, they would all be be eventually local in α. Thus, in the corresponding derivation tree the members of Σ would be included in a dashed box although the derivation does not comply with tree-locality.

tree that rewrites one of its leaf nodes by substituting there. The assumption that the bridge verbs and the main lexical verbs combine with one another by means of substitution goes against Candito & Kahane (1998: Sec. 2), where it is argued that the trees for bridge verbs are "[t]ypical predicative auxiliary trees" (see again section 5.2.1). According to Candito & Kahane, the tree anchored by *permitir* should hence feature a foot node whose label coincides with the root of the same tree, while Chen-Main, Bleam & Joshi (2012: 3) posit that XP is the label of the node in α*.permitir* where α*.ver* is substituted.[168]

This contrast between the two works just mentioned highlights a vulnerability in the account provided by Chen-Main, Bleam & Joshi (2012: 3). In fact, assuming that α*.permitir* bears a non-terminal leaf labeled XP implies that α*.ver* substitutes there even if the labels do not match. Section 5.4 will lay the foundation for a seamless reconciliation of the two perspectives, without the need of (implicitly) tweaking the conditions under which substitutions can apply. Moreover, section 5.4.3 will suggest a possible account for an aspect which is not explicitly touched upon in the derivation trees in (226), i.e. the control relations holding in each derivation among the verbs that anchor initial trees.

It is now time to establish the value of k, i.e. the maximum number of delays of which any node in the derivation tree can be a member of. For a start, it must be observed that k must be greater than 1, for a twofold reason. The first is empirical, as (226c) is an instance of a grammatical derivation that exploits the power of a 2-delayed MC-TAG by having β*.te* and α*.permitir* be part of two delays. The second is grounded on the theoretical proof of weak equivalence between 2-delayed 2-MC-TAGs and 2-TL-MC-TAGs provided by Chiang & Scheffler (2008). On these bases, the literature has not only agreed that k cannot be lower than 2, but has also assumed that $k = 2$ due to the lack of evidence in favor of a greater value of k (see Chen-Main, Bleam & Joshi 2012: Sec. 5 for discussion on this point). By following the guidelines given by Bleam (2000: 193) in this respect (see page 268 once more), a 2-delayed

[168] As will be discussed in section 5.4.1, the root of α*.permitir* is assumed to be VP because Chen-Main, Bleam & Joshi (2012) do not resort to a projection of T to account for the non-finite nature of the bridge verbs involved in clitic climbing. Still, the absence of an S projection in the trees they propose suggests that they would agree with Frank (2002: Sec. 5.2), who argues in turn that the extended projection of bridge verbs fails to span up to CP, similarly to what has been seen for raising verbs since section 5.1.5. Although the absence of a CP in the elementary trees of bridge verbs is in line with the restructuring proposed by Rizzi (1982), the very presence of distinct elementary trees anchored by the various verbs is in line a multi-clausal view *à la* Kayne (1989). Moreover, the non-monotonic formation of the multi-component set of trees might also foster a reconciliation with the data discussed in Sportiche (1998: Ch. 4-7), where it is argued that clitics are merged only where they are overtly realized. Although these matters need of course further research, they will be touched upon once more in footnote 181 on page 347.

2-MC-TAG will thus be adopted here as well, provided that it never undergenerates.[169]

To conclude this subsection, the readers are invited to compare the radically different analyses that the compliance with the simultaneity requirements (see again Kallmeyer 2009: 65 and section 5.3.1) and the conformity with the counting of the maximum delays per node (see Kallmeyer 2009: 74-75, 90 for a formal implementation) provide for the same derivation tree(s). This can be done by recollecting the three examples previously shown from page 275 on (and to be reproposed shortly) which were generated by combining the following sets of trees:

$\Sigma_0=\{\alpha\}$; $\Sigma_1=\{\gamma_1,\gamma_2\}$; $\Sigma_2=\{\gamma_3\}$; $\Sigma_3=\{\gamma_4,\gamma_5\}$; $\Sigma_4=\{\gamma_6,\gamma_7\}$; $\Sigma_5=\{\gamma_8\}$; $\Sigma_6=\{\gamma_9,\gamma_{10}\}$.

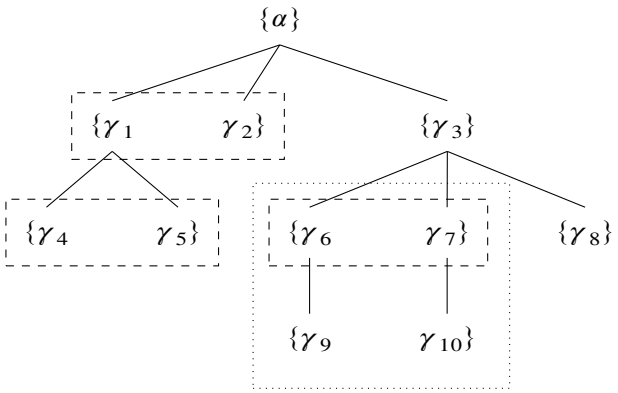
By using the simultaneity requirements as a criterion of analysis, the first tree in (205) was said to be entirely compliant with set-locality, while the last in (207) was a non-local derivation also breaking the simultaneity requirements in all the conceivable manners. The derivation tree in (206) was an in-between case, as it was a non-local derivation still managing to follow all the simultaneity requirements.

The same structures are shown again in (227)-(229) in the same order, but this time they are to be analyzed on the basis of the delays they feature.[170] As the readers can see, this modification leads to completely different considerations, as all the three derivation trees fully exploit the power provided by a 2-delayed 2-MC-TAG in a legal way. In fact, in (227) the nodes that are members of two delays are γ_6 and γ_7. In (228), the same happens with respect to γ_1, γ_2, γ_3, γ_5 and γ_7. In (229), this phenomenon affects γ_1, γ_2 and γ_5.Compared to the others, the derivation tree in (228) further stands out for featuring five boxes. This is due to the fact that, if one were to capture the delay of Σ_6 by means of a single box in (228), γ_2 and γ_6 would be erroneously included as well. Thus, the two paths leading "from the destination down to each member of" Σ_6 (Chiang & Scheffler 2008: 19) must be represented separately, though connected by means of a double-pointed arrow (labeled $\cup$) which indicates their union into a single delay.

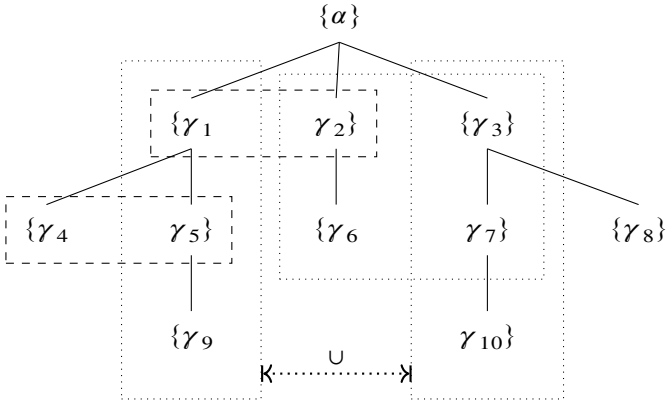
[169] See again sections 5.3.1 and 5.3.2 for the arguments that led to assume $h = 2$, where h is the maximum arity allowed per set of trees in an MC-TAG.

[170] Just as in section 5.3.1, it is not important if the elementary trees just listed are initial or auxiliary trees. Therefore, all the solid edges in (227)-(229) indicate unspecified TAG operations, instead of specific occurrences of adjoining.

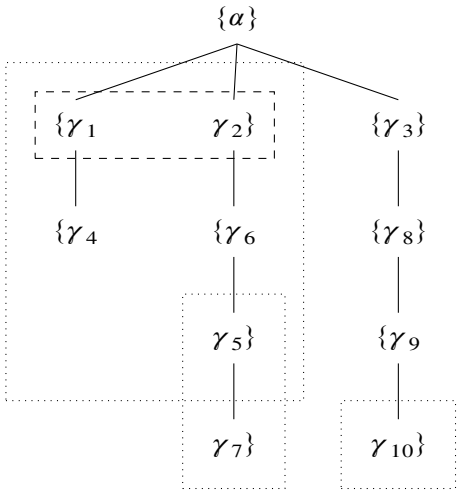
(227)



(228)



(229)



5.4 Fixing the problems of L-TAGs

The premises are now almost set in order to finally propose a solution to the main issues that emerged throughout the previous sections of the present chapter. First and foremost, it will be shown that a 2-delayed 2-MC-TAG approach allows to propose an analysis where a

wh-question is the targeted element, instead of the targeting one, if it combines with a subordinate clause it is supposed to select. The solution to this problem, to be explained in section 5.4.1, fosters the twofold content in section 5.4.2. On the one hand, novel formal definitions that may interest both syntacticians and computational linguists will be introduced. On the other hand, 2-delayed 2-MC-TAG will be proven to provide a solid counterargument to Georgi (2019), where TAGs are said to be inadequate because of some problems that an L-TAG approach *à la* Frank (2002, 2004, 2006, 2010, 2013) eventually faces. Although the latter argumentation will in turn raise another apparent problem, in section 5.4.3 it will be claimed that such an apparent issue is actually a welcome result that points to an empirical obervation already formulated by Longenbaugh (2014: Sec. 3.4).

The fact that different TAG approaches lead to the same conclusions is an interesting aspect that is not new to TAGs in general, as Chen-Main & Joshi (2007: Sec. 6) and Chen-Main, Bleam & Joshi (2012) discuss, and arguably demonstrates that L-TAGs and the approaches considered by such papers are all close to an adequate account of the formal power of natural language. However, section 5.4.3 will argue that the novel TAG approach proposed here is superior to L-TAGs (especially if a Feature-Based account like the one proposed in section 5.4.4 is adopted) because it solves many of the potential issues raised by L-TAGs while paving the way for the formulation of a single syntactic component that, depending on theoretical and empirical needs, can encompass Merge and TAG operations as well as head movement.

5.4.1 Reversing the perspective on adjoining

The usual TAG derivation for the sentence *What do you think the senator should read?*, discussed in sections 5.1.2 and 5.1.3, proceeds in a way which goes against the intuitive expectations. In fact, the main clause must adjoin to the embedded clause it should be selecting, so that the former is dominated by the latter in the derivation tree. Thus, the sentence can be bracketed as in (230a) below, where the subscripted β indicates the auxiliary tree and C′* stands for its recursive node. However, similar counterparts with no *wh*-movement to the left periphery of the main clause can be analyzed either as a main clause adjoining at the root of an embedded clause, or vice versa as an embedded clause substituting at the CP node in the frontier of the main clause (Kulick 2000: 54, Chen-Main & Joshi 2014: 1011). The former strategy is adopted by Frank (2013: 232) and the latter by Frank (2006: 155) for the sentences

in (230b) and (230c), respectively.

(230) a. What C′[$_\beta$ do you think C′*] the senator should read?

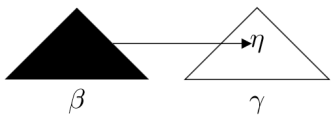
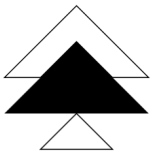
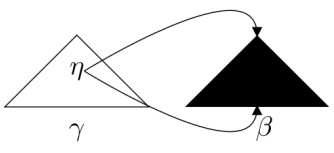
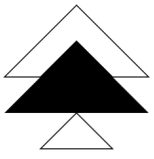
b. [$_\alpha$ CP [$_\beta$ Every student hopes CP*] the professor finishes the lecture quickly]

c. [$_\alpha$ John wondered CP [$_\alpha$ what the senator should read]]

The starting point towards an alternate structural description for the sentences in (230) is given by following Joshi (2003: 252-254) and reversing the perspective on the adjoining operation. This idea, first addressed as *wrapping* (see also Kallmeyer & Joshi 2003: 13-14), has been named in several ways in the TAG literature, i.e. *flexible composition* (Joshi, Kallmeyer & Romero 2003: Sec. 3, Chen-Main & Joshi 2014: 1009), *reverse adjunction* (Chiang & Scheffler 2008), *wrapping substitution* (Kallmeyer, Osswald & Van Valin 2013: 178) and *reverse adjoining* (Joshi 2014: 152, 156). Despite the different names, this operation always shows that an auxiliary tree β shoehorning an internal node labeled *N* of an elementary tree γ can be reanalyzed by splitting the latter tree at *N*, so that one obtains two subparts γ_1 and γ_2, where *N* is both a leaf of γ_1 and the root of γ_2. The reversal of the adjoining operation is complete once γ_1 and γ_2 target the root and the foot node of β, respectively.

Chiang & Scheffler (2008: 18) illustrate this procedure with what is reported in (231) on this page. In (231), *adjunction* stands for the adjoining operation (see discussion in section 5.2.1), the node *N* is labeled as η, and the labels on the edges of derivation trees indicate the type of operation involved (see section 5.1.3).

(231)

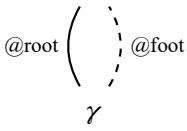
	Operation	Derivation	Action	Result
(a)	adjunction	γ $\mid_{\text{adj@}\eta}$ β		
(b)	reverse-adjunction	β $\mid_{\text{rev@}\eta}$ γ		

The comparison between the two derivation trees in (231) shows that the one in (231b) manages to obtain the desired outcome, namely the reversed kinship between β and γ. How-

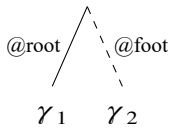
ever, such a derivation tree is not an adequate account of the facts, as it endorses the wrong interpretation of the "action" depicted in (231b). The latter is ambiguous as to whether γ is performing one or two TAG operations targeting β. The former interpretation stems from the conceptual reversal of the common adjoining, but the two arrows pointing at different nodes should suggest the latter. In fact, the two white triangles in the "result" of (231b) correctly imply that the derived tree is made of three elementary trees, namely β together with the two subparts γ_1 and γ_2, each perfoming one of the two operations shown by the "action" in (231b). On the one hand, the arrow pointing at β's root must be understood as γ_1 adjoining there (as the root of a tree is the non-terminal node par excellence), while the other arrow pointing at β's foot node depicts γ_2 substituting there.

The issue pointed out for (231b) shows that the reversal of an adjoining operation must go hand in hand with a reversal of the Markovian property of TAG derivations discussed in section 5.1.4. Just like an auxiliary tree must be "born and not made" before the adjoining (Frank 2002: 68, 2006: 160, 2013: 245), γ_1 and γ_2 must be born-and-not-made elementary trees before they target β. As a result of the present discussion, (232a) shows the actual derivation graph for the action in (231b), which is not a tree because the two nodes are connected by two parallel edges (see Krivochen 2023a: 129-130 and references therein for related terminology and observations), while (232b) shows the actual well-formed derivation tree for (231b). The latter can also be observed to be an instance of the tree-local scenario in (224a) on page 294, which was adapted from Kulick, Frank & Vijay-Shanker (2000: 41).

(232) a.



b.



If a single elementary tree β adjoining to γ can alternatively be understood as γ_1 and γ_2 combining with β, then γ_1 and γ_2 should arguably form a single multi-component that complies with FTH and CETM in (197). As consequence, one can claim that the reversal of the adjoining operation requires an *h*-MC-TAG whose maximum arity be > 1 (Joshi 2014: 142-143 fn. 7). Thus, the sentence in (230b) on page 303 can be reanalyzed as in (233a), where γ_2 is (as expected) an initial tree α, while γ_1 is realized by a degenerate CP tree δ. The latter structural description bears strong and interesting similarities with (230c), which is repeated in (233b) to make the comparison easier.

(233) a. [$_\beta$ [$_\delta$ CP] Every student hopes CP* [$_\alpha$ the professor finishes the lecture quickly]]

b. [$_\alpha$ John wondered CP [$_\alpha$ what the senator should read]]

It can now be seen that the reversed perspective on the adjoining operation allows to put the subordinate clauses in (233a) and (233b) on a par, as they are both initial trees substituting at a CP node in the frontier of the respective main clauses. This has the advantage of sketching a consistent treatment of similar sentences, in contrast to analyses that oscillate indecisively between substitution of initial trees and adjoining of auxiliary trees (a problem first raised by Vijay-Shanker 1992: Sec. 5). The parallel between (233a) and (233b) can be enhanced further by arguing that initial trees having a leaf labeled CP can be seen as auxiliary trees in disguise, since CP is also the label of their root. With this in mind, it is possible to see the α tree anchored by *wonder* in (233b) as a β tree, with a degenerate tree δ (locally related to the subordinate by coming from the same multi-component) targeting its root. This further revision is provided below:

(234) a. [$_\beta$ [$_\delta$ CP] Every student hopes CP* [$_\alpha$ the professor finishes the lecture quickly]]

b. [$_\beta$ [$_\delta$ CP] John wondered CP* [$_\alpha$ what the senator should read]]

The bracketings in (234) manage to provide a common structural description for sentences with arguably similar structures, thus improving the original analyses given by Frank (2006: 155, 2013: 232). However, what is proposed in (234) suffers from a few problems. For a start, in section 5.3.1 it has been argued that degenerate trees lead to *ad hoc* overgenerating analyses (and section 5.4.2 will further strengthen this claim). In this case it can however be seen that degenerate CPs undergenerate, as they are of little use in order to account for the sentence that was our major concern from the beginning, i.e. *What do you think that the senator should read?*. In the analysis provided in (230a) on page 303 (as well as in section 5.1.2), the main clause adjoins at the C′ node of the subordinate clause. In order to comply with the derivation tree in (232b), the reversed perspective on the same operation requires a multi-component where γ_2 is an initial tree made of C′ and all the nodes it exhaustively dominates, while γ_1 is a non-degenerate auxiliary tree recursive at C′. The latter requirement is however problematic, because the *wh*-operator should be dominated by a CP node, rather than C′.

As a consequence of these premises, the reversal of (230a), reported in (235a) below for convenience, leads to the problematic structural description in (235b), where both γ_1 and the subsequently derived tree are rooted at C′. Adding a degenerate tree to the multi-component

formed by γ_1 and γ_2 would not fix this conondrum. In fact, if δ were recursive at CP, it could not adjoin anywhere. If instead it were recursive at C′, it would not manage to improve the structural description.

(235) a. What C′ [$_\beta$ do you think C′*] the senator should read?

b. [$_\beta$ C′ [$_{\gamma_1}$ What C′*] do you think C′* [$_{\gamma_2}$ the senator should read]]?

Another issue that is arguably even more serious is that the derivation tree in (232b) and those corresponding to the derived structures in (234) and (235b) violate a detail of the Single Mother Condition for (non-simultaneous *h*-MC-)TAG derivation trees stated in (225) on page 296, i.e. they all have a β node (which stands for an auxiliary tree) as a root, instead of a α node (corresponding to an initial tree). This aspect is not an exclusive peculiarity of the formalizations that Kallmeyer (2009: 57) proposes for *h*-MC-TAGs with $h \geq 2$, but is a rather common assumption among TAG approaches in general. In fact, it is also assumed for L-TAGs in Kallmeyer (2010: 53) as well as in the *derivational completness* introduced by Frank (2013: 244) and previously presented in (159a) on page 221.

From the discussion so far, it seems that adjoining cannot be effectively reversed in a TAG approach that tries to be consistent with many generative assumptions. However, a solution to the present issue can be found by starting from a less recent work, where an ever more stringent version of the requirement hindering the reversal of the adjoining operation is proposed: "The root of a derivation tree for TAGs *is labeled by an S-type initial tree*" (Joshi & Schabes 1997: 75, italics added). The *S* label mentioned there identifies the root of a sentence in works predating, or not adopting, the C-T-v-V system discussed in sections 2.2.3 and 2.2.4. Therefore, the assumption in Joshi & Schabes (1997: 75) can be adapted to common generative parlance by stating that, in a TAG, the root of a derivation tree is always an initial tree whose root is CP.

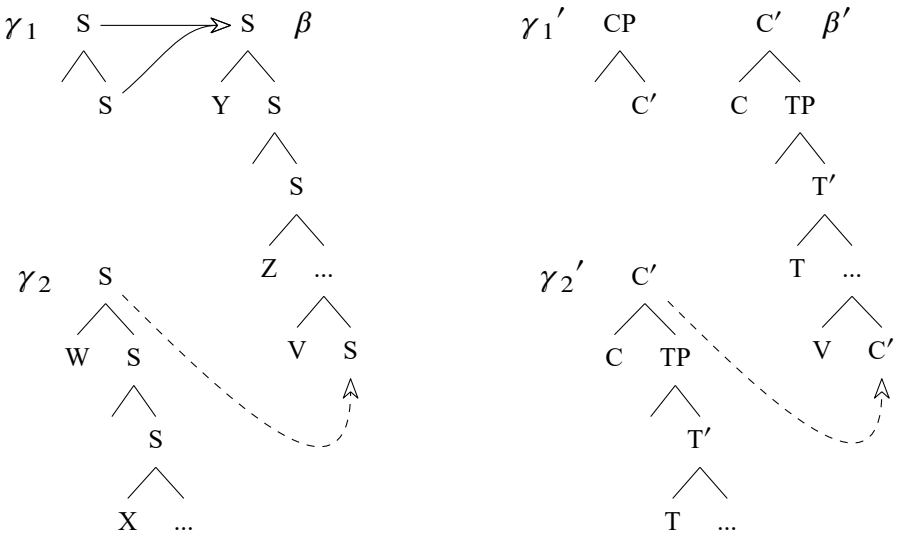
The point just made (which section 5.2.1 needed to anticipate) is highly relevant, because all the TAG works discussing the reversal of an adjoining operation also adopt stacked S projections to account for the verbal extended projection (and stacked Ns for the nominal ones).[171] Because of this, any confrontation among such analyses and the ones provided by Frank (2002, 2004, 2006, 2010, 2013), Storoshenko (2006), Frank & Storoshenko (2012),

[171] An arguably non-exhaustive list of works that explicitly state that adjoining can be reversed is the following: Joshi (2003, 2014), Kallmeyer & Joshi (2003), Joshi, Kallmeyer & Romero (2003), Chiang & Scheffler (2008), Chen-Main & Joshi (2008, 2014), Nesson (2009), Nesson & Shieber (2009), and Kallmeyer, Osswald & Van Valin (2013).

Longenbaugh (2014), Georgi (2019) and Krivochen & Padovan (2021) not only requires additional scrutiny and careful reflection, but can also raise some unexpected contrast which eventually highlights problems that would not be so obvious otherwise.

This fact can be exemplified by means of (236), where two equivalent representations of the same reversal of an adjoining operation are supposed to be given. Instead of having β and β' respectively adjoined at an inner S node of γ and the inner C′ node of γ', γ and γ' are assumed to be split at S and C′, respectively. Hence, the former is decomposed into γ_1 and γ_2 so that they eventually combine with β, while the latter is divided into γ_1' and γ_2', which are supposed form a derived tree together with β' just as easily.

(236)



As the readers can see, (236) not only raises the same problem already discussed with respect to (235b), but also shows why such an issue has not been pointed out by previous works. As long as inner nodes are all labeled with *S*, the concern with γ_1 is not whether it can be adjoined to β, but rather how to ensure that it adjoins at its root and not elsewhere (see a.o. Vijay-Shanker 1987, Abeillé & Rambow 2000: Sec. 1.2.4, Frank 2013: 244 ff., Lichte 2015: Sec. 5.2.1, Findlay 2023: Sec. 3.3 for discussion on possible ways to rule this). The tables are turned if the supposedly equivalent γ_1', γ_2' and β' are considered, where C's replace the previous Y and W and T's do so with X and Z. As long as C and T heads project their labels in a way that complies with the X-bar schema discussed in section 2.2.3, the S labels that pervaded γ_1, γ_2 and β are dispensed with. As a consequence, γ_1' either fails to adjoin as shown in (236) or does so by raising the issues already discussed for the derived

structure in (235b) and, more generally, for any derivation tree following (232b).[172]

Joshi (2014: 142-143 fn. 7) hints at the key to solve this relevant issue, but does not elaborate it further (see also Vijay-Shanker 1992: Sec. 5 for related discussion):

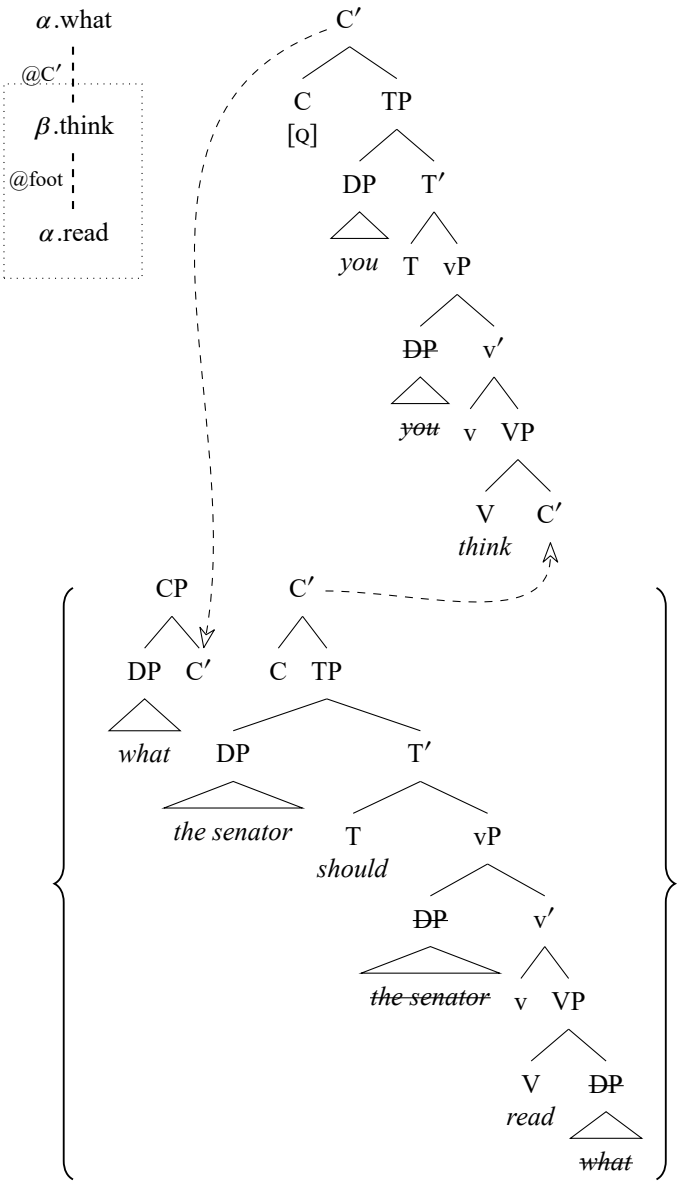
> If a node X in a tree t is targeted for adjoining by an auxiliary tree, say t', then the adjoining operation can be thought of as a pair of substitutions such that the tree t is pulled apart at the node X, creating two X nodes (top and bottom), and then t' is spliced into the tree t by making a pair of substitutions.

By following the suggestion from the quotation above, in (236) it should not be γ_1' which adjoins at the root of β', but rather the latter which substitutes at the matching leaf node of the former. Therefore, a more effective reversal of the adjoining operation must not follow the schema in (224a) on page 294, but rather the alternative in (224b), which is legal only in *k*-delayed MC-TAGs (see again section 5.3.3). As soon as the latter approach is adopted to ensure the desired corrections, the derivation tree for the combination of γ_1', β and γ_2' in (236) improves as well, because it becomes rooted at γ_1', which is an initial tree with a CP root.

At long last, it is now possible to provide a rather promising analysis for the sentence that was introduced in section 5.1.2 and that guided the discussion of the present section, namely *What do you think the senator should read?*, whose derivation is reworked as in (237) on the next page. By featuring the same multi-component already presented in section 5.3.2 on page 279, (237) neatly illustrates the just mentioned adjustments while also representing a concrete instance of the operation that is illegal for TL- and SL-MC-TAGs, i.e. the one presented in (224b) on page 294. As a consequence, (237) cannot be obtained by doing without a *k*-delayed *h*-MC-TAG. With respect to the questions that were left open at the end of 5.1.2, (237) manages to fix the wrong relation between the main clause and the embedded one, but still bears troublesome bar-projections and an unsatisfactory explanation of the asymmetric T-to-C raising in English. Such issues must wait until section 5.4.4.

[172] The problem pointed out in the main text also clarifies why in the TAG literature non-singleton sets of trees are usually assumed to include at least an auxiliary tree (see again the common MC-TAG tenets stated in section 5.3.2 following Chen-Main & Joshi 2007: Sec. 6), whereas in the present chapter that need not be the case.

(237)



5.4.2 A rejoinder to Georgi (2019)

The structure just shown in (237) illustrates a really important difference between the previous TAG accounts and the one proposed here, namely that an auxiliary tree does not necessarily have to adjoin to another elementary tree (though see Vijay-Shanker 1992: Sec. 5 for a work hinting at this possibility). This eventually justifies the seemingly redundant decision of following Joshi & Schabes (1997: 76) and Joshi (2014: 151), where the distinction made through Greek letters on the nodes of derivation trees goes along with dashed edges for substitutions and solid lines for instances of adjoining (see again section 5.1.3). In fact, (237) represents the simplest scenario of the alternative operation in which an auxiliary tree can be

involved. Such an operation will be called *chain of substitions* and defined as follows:

(238) A well-formed chain of substitutions $\mathcal{S}$ is made of:

a. A *tail* $\mathcal{T}$, namely an initial tree α whose root is N, where N can be a legitimate label for both root and leaf non-terminals in any elementary tree.

b. A *body* $\mathcal{B}$, where $\mathcal{B}$ is a list of auxiliary trees β. All the trees $\in \mathcal{B}$ are recursive at N and $j \geq 1$, where j is the integer number of auxiliary trees $\in \mathcal{B}$.
Given any value of j, $\alpha \in \mathcal{T}$ substitutes at the foot node $N \in \beta_j$.
If $j \neq 1$, then β_j substitutes at the foot node $N \in \beta_{j-1}$, β_{j-1} substitutes at the foot node $N \in \beta_{j-2}$..., β_2 substitutes at the foot node $N \in \beta_1$.

c. A *head* $\mathcal{H}$, namely an elementary tree γ with a leaf node N at which β_1 substitutes. If a foot node $F \in \gamma$ is also present, then $F \neq N$.

In (237), the head and the tail of the chain of substitutions (i.e. α*.what* and α*.read*, respectively) constitute the two members of the same multi-component (first discussed in section 5.3.2). As can be seen in the related derivation tree, this implies the use of a delay in which both the tail and the body of the chain are included, whereas the head represents the destination of the delay itself (see again the definitions reported in section 5.3.3 following Chiang & Scheffler 2008: 19). As the readers will observe, all the chains of substitutions to be discussed in the present chapter will share these properties with (237), even though (238) does not impose that the head and the tail of the chain be necessarily members of the same set of trees. It is a matter of future research to assess whether some syntactic phenomenon can be captured insightfully only by means of a chain of substitutions whose head and tail belong to different sets of trees.

There is another observation that must be made with respect to (237). In the substitution chain $\mathcal{S}$, the precise label of the node that each tree targets only needs to be specified on the edge connecting the head $\mathcal{H}$, i.e. α*.what*, with its daugther in the derivation tree that also belongs to the body of $\mathcal{S}$, i.e. β*.think*. In fact, (238) rules that, once the label N of the leaf node in the head tree $\mathcal{H}$ is established, the root of the tail tree $\mathcal{T}$ and both root and foot nodes of the auxiliary trees in the body $\mathcal{B}$ must be N as well. To put it differently, the information that β*.think* substitutes at C′ on α*.what* percolates on the rest of the chain $\mathcal{S}$. This can be indicated on the other edges of the derivation tree by means of what ensures it, namely the fact that each member of the chain (but the head) substitutes to another by targeting the foot

node of the latter. As a result, telling that $\alpha.read$ substitutes at the foot node of $\beta.think$ is more informative than telling the targeted label, as the latter information can be inferred from the opposite end of the chain.

Now that (237) has been clarified, the focus will shift to an in-depth discussion of (238). The first point to make explicit is that the formulation in (238) is guided by the following assumption:

(239) Derivational completeness for auxiliary trees: Each auxiliary tree must have its recursive nodes (i.e. root and foot) both involved in a TAG operation.

Readers coming from section 5.2.1 may have noted a couple of things. First, (238b) distinguishes auxiliary trees by means of indexes, just as it has been done for the strong co-occurrence constraints discussed on page 254 and on page 255. This observation has far-reaching consequences that must be postponed to section 5.5.

Second, both (238) and (239) are in contrast with definitions which posit that foot nodes are not substitution nodes, e.g. Kallmeyer (2010: 54) and Joshi & Schabes (1997: 71). Such an assumption must be dispensed with in the present approach, so that the only property that singles out the foot node among the other leaves of an auxiliary tree is the coincident label of the foot with the root belonging to its same spine (in the sense of Nesson 2009: 4, Kuhlmann & Satta 2012: 623).

The principle in (239) is what determines that the tail $\mathscr{T}$ in (238a) must be an initial tree. In fact, a unique auxiliary tree forming an alternate tail $\mathscr{T}'$ would fail in two possible ways. On the one hand, $\mathscr{T}'$ could be an effective tail at the cost of not having its own foot node rewritten, thus violating (239). If on the other hand such an auxiliary tree does have its own foot node rewritten, it would fail to effectively close the chain of substitutions $\mathscr{S}$, as $\mathscr{T}'$ would subsequently be indistinguishable from any auxiliary tree in the body of $\mathscr{S}$.

On the contrary, the nature of the elementary tree forming the head $\mathscr{H}$ is not relevant to (238c). The most pressing aspect concerning the latter is that the node *N* recurring across the whole substitution chain be different from the foot node in $\mathscr{H}$, whenever $\mathscr{H}$ is not instantiated by an initial tree α but rather by an auxiliary tree β. Otherwise, the root of such $\mathscr{H}$ would also be an instance of *N*, thus making $\mathscr{H}$ indistinguishable from any other auxiliary tree in the body $\mathscr{B}$. In fact, the only way to make $\mathscr{H}$ different under such circumstances would be to have it not substituted to another elementary tree, thus violating both (239) and the Single Mother Condition for non-simultaneous *h*-MC-TAG derivation trees in (225) on page 296.

The way in which the body of a chain of substitutions $\mathscr{B}$ is defined in (238b) shows that the reversal of the adjoining operation discussed in section 5.4.1 and exemplified by (237) corresponds to the specific case in (238) where there is a unique auxiliary tree in $\mathscr{B}$, i.e. $j = 1$. This suggests that a chain $\mathscr{S}'$ in which $j \geq 2$ can be weakly equivalent to an alternate derivation where the auxiliary trees in $\mathscr{B}$ are all adjoined. In fact, the counterargument to Georgi (2019) to be illustrated shortly will demonstrate that there exists some syntactic phenomenon where this formal prediction is borne out. However, it is first necessary to define this conceptual alternative to (238), as follows:

(240) A well-formed chain of adjoining operations $\mathscr{A}$ is made of:

a. A *body* $\mathscr{B}$, where $\mathscr{B}$ is a non-empty list of auxiliary trees β that are all recursive at N, where N can be a legitimate label at least for the set of internal nodes in any elementary tree.
   If $j \geq 2$, where j is the number of auxiliary trees $\in \mathscr{B}$, then β_1 adjoins at the root node $N \in \beta_2$, β_2 adjoins at the root node $N \in \beta_3$..., β_{j-1} adjoins at the root node $N \in \beta_j$.

b. A *head* $\mathscr{H}$, namely an elementary tree γ with an internal node N to which β_j adjoins. If γ is an auxiliary tree, then N cannot be its root R, so that $R \neq N$.

One can immediately observe that any $\mathscr{A}$ defined by (240) differs from any $\mathscr{S}$ established by (238) because the latter features a tail $\mathscr{T}$ while the former does not. An example of chain of adoining operations has already been provided in section 5.2.1, where the derivation of a clause with two raising verbs was discussed in (182) on page 248. It will be reproposed shortly with a slight modification, as it will eventually support the comparison between (238) and (240). In order to get to that point of the discussion, it is important to first mention three cases that Georgi (2019) shows to be problematic for an L-TAG *à la* Frank (2002, 2004, 2006, 2010, 2013), i.e. stylistic inversion in French and cases of reconstruction for Condition A of the Binding Theory occurring with both long-distance A- and long-distance A′-movements. All of them have a common matrix, which can be highlighted by adapting the following quotation from Georgi (2019: 87) (note that *[IT+AT2]* is part of the original text):

> The only way to circumvent this [type of] problem[s] would be to first adjoin AT2 [i.e., the auxiliary tree that is the most embedded in the derived tree] to IT [i.e., the only initial tree], and then to adjoin AT1 [i.e., the auxiliary tree that

> is the least embedded in the derived tree] to [IT+AT2]. [...] However, such a derivation is explicitly excluded in TAG.

At this point it is important to remark that Georgi (2019) focuses on L-TAGs without mentioning MC-TAGs as viable alternatives. The upcoming discussion will compensate this by eventually showing that the seemingly illegal operation which Georgi (2019: 87) mentions can be handled with the 2-delayed 2-MC-TAGs (introduced in section 5.3.3) just as easily as the derivation in (224b) on page 294. Therefore, what Georgi (2019: 87) argues to be an insurmountable problem for TAGs actually suggests the formalism to adopt in order to formulate an adequate account.

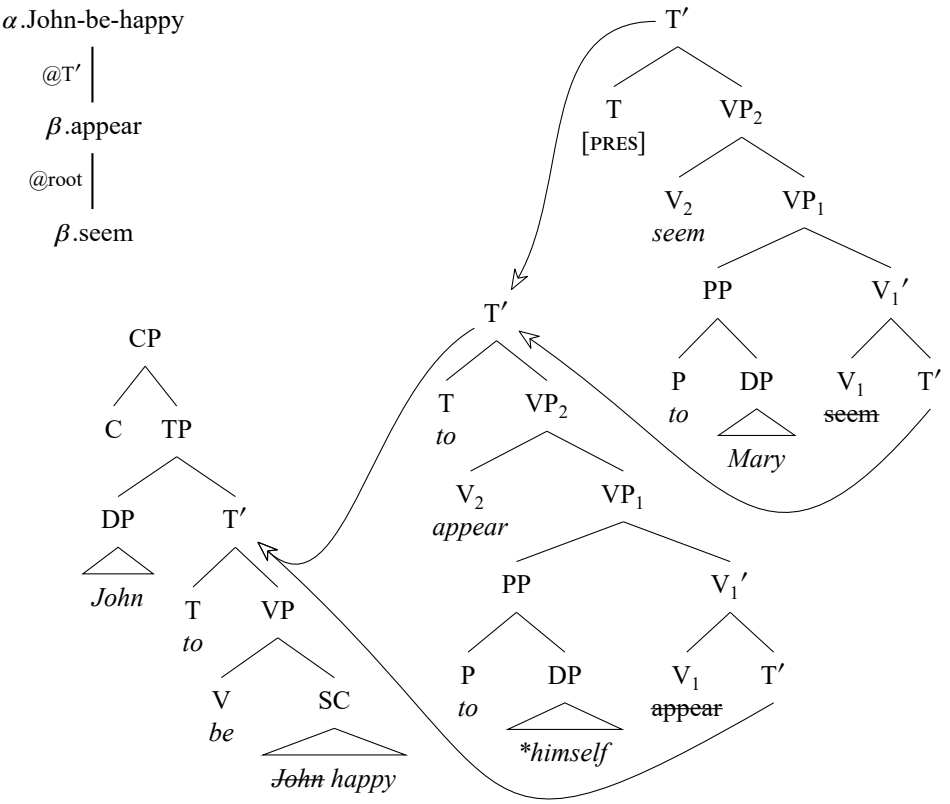
Given that the desired TAG operation that has just been quoted would be the key to generate all the three problematic derivations discussed by Georgi (2019), the argument will be made here only with respect to one of them, namely the scope reconstruction for A-movement with multiple raising verbs. In particular, Georgi (2019: 88) shows that an L-TAG *à la* Frank (2002, 2004, 2006, 2010, 2013) makes the wrong prediction with respect to the grammatical sentence that she discusses following Castillo, Drury & Grohmann (2009–1999: 93):

(241) John$_i$ seems to Mary to appear to himself$_i$ to be happy.

In fact, the anaphor in (241) is expected to be ungrammatical in the chain of adjoining operations applied to derive the analogous structure in (182) on page 248 (where the R-expression *Luke* was merged instead of the reflexive). The problem with the L-TAG derivation in (242) on the following page is that at no point the DP *John* reaches the Spec,TP in the extended projection of the raising verb *appear*. Thus, the anaphor in β*.appear* is never "in the required structural condition to fulfill Principle A" (Georgi 2019: 88). The readers may also note that the derivation tree for (242) is misleading in this respect, as the parent-child relation between α*.John-be-happy* and β*.appear* seemingly suggests the desired locality between the two at the expense of β*.seem*, contrary to fact.

The problem that arises with respect to the anaphor can be set aside temporarily, so as to look at (242) as the previously promised example of chain of adjoining operations $\mathscr{A}$. It can thus be observed what was already witnessed for the chain of substitutions instantiated by (237), namely that the precise label of the node that is targeted by each auxiliary tree in $\mathscr{A}$ only needs to be specified on one end of the chain in the derivation tree, namely at the edge connecting the head $\mathscr{H}$ α*.John-be-happy* with its daugther that also belongs to the body of

(242)

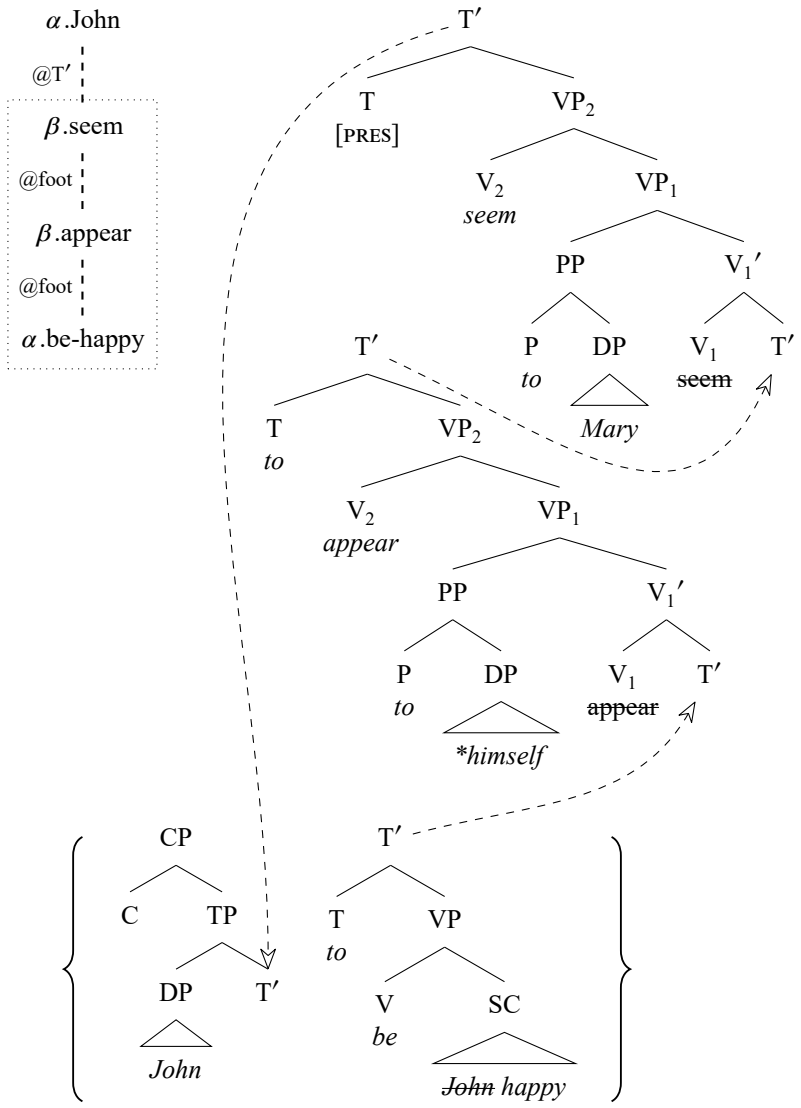


$\mathscr{A}$, i.e. *β.appear*. In a similar vein to what (238) does for the body and tail of any chain of substitutions $\mathscr{S}$, (240) warrants that the derivation in (237) be recursive at T′, so that this information percolates along the body of $\mathscr{A}$. However, $\mathscr{A}$ ensures such a percolation differently from $\mathscr{S}$, namely by making each auxiliary tree adjoin at the root of another elementary tree.

One first option that a 2-delayed 2-MC-TAG can generate as an alternative to (242) is a chain of substitutions whose body contains all the auxiliary trees in the body of $\mathscr{A}$ in (242), i.e. *β.appear* and *β.seem*. The latter are however combined in the reversed order in such $\mathscr{S}$, so that their parent-child relation in the derivation tree is swapped as well, as (243) shows on the next page. Moreover, the head and tail of $\mathscr{S}$, if composed together, are equivalent to the head of $\mathscr{A}$ *α.John-be-happy* in (242). In fact, $\mathscr{H}$ and $\mathscr{T} \in \mathscr{S}$ form a multi-component that is generated by means of a non-monotonic derivation where *α.John-be-happy* splits at T′, thus resulting in the trees *α.John* and *α.be-happy* in (243).

However, (243) illustrates that this alternative does not solve the locality problem between *John* and *himself* that (242) also raises. Therefore, it can be argued that (242) and (243) are weakly equivalent, so that they provide the anticipated empirical example in favor of the

(243)



following statement:

(244) Hypothesis of weak equivalence among chains: A chain of substitutions $\mathcal{S}$ and a chain of adjoining operations $\mathcal{A}$ are weakly equivalent if the unmediated substitution of $\mathcal{T}(\mathcal{S})$ to $\mathcal{H}(\mathcal{S})$ would generate a structure equivalent to $\mathcal{H}(\mathcal{A})$, and the auxiliary trees in $\mathcal{B}(\mathcal{S})$ form a sequence of trees whose order is reversed in $\mathcal{B}(\mathcal{A})$.

Although borne out by (242) and (243), the hypothesis in (244) needs further mathematical proof and empirical support, all tasks that are left for future research. In particular, it is a matter of verification whether the auxiliary trees involved in such weakly equivalent chains need to be of the same type, e.g. raising predicates as in the case above. For now, two things are important. The first is that (243) still fails at providing an account for the sentence in (241),

while the second is that the weak equivalence in (244) seems to hold only if the auxiliary trees involved in $\mathcal{B}(\mathcal{A})$ are in a reversed order opposed to $\mathcal{H}(\mathcal{S})$. This last point will be further discussed in section 5.4.3.

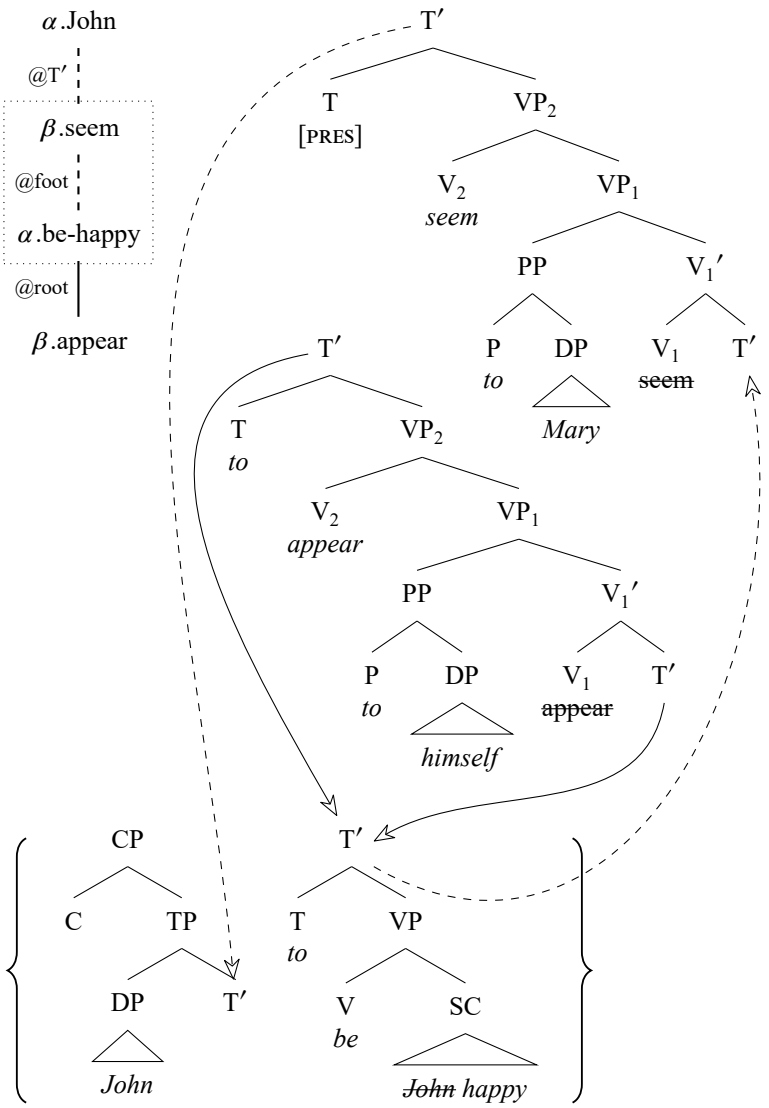
Before going any further, it is worth pausing to note that Georgi (2019) implicitly suggests that the binding conditions can be checked on some derived tree obtained in the course of the derivation, rather than in the one which is derived at the end. This is a reasonable option, as it has predecessors in works that Safir (2008: Sec. 2) dubs a "Movement-As-Coconstrual (MAC)" approaches. TAGs arguably allow to circumvent the problems that such analyses suffer from (see Safir 2008 for discussion), but detailed considerations in this respect must be reserved for future work.

Still, it would be interesting to complement this implicit premise in Georgi (2019) with what was discussed in section 5.1.5, namely the two alternate proposals that Frank (2002: Sec. 4.6) and Frank (2006: 195 ff.) formulate so as to account for the fact that, even though they originate in distinct trees, the subject DP and the T head of the tensed raising verb check each other's EPP- and φ-features. Frank (2006: 195 ff.) considers this process "as analogous to cases of cross-clausal agreement", and thus judges his previous proposal an avoidable complication. However, Frank (2002: Sec. 4.6) holds the specifier-head configuration formed by DP and T as a crucial factor for establishing a local relation between them. Thus, such an assumption could be reinterpreted as the key process that makes the DP enter in a local relation with the material contained in the elementary tree that combines at the sister node of the DP itself. In (242), this corresponds to β*.seem* adjoining at the T′ node of α*.John-be-happy*, while in (243) the same situation holds for β*.seem* substituting at the T′ node of α*.John*. If on the one hand the proposal in Frank (2002: Sec. 4.6) can be regarded as an interesting addition to what Georgi (2019) suggests, the same cannot be said with respect to the alternate option in Frank (2006: 195 ff.), as the ability of agreement to span across clauses (still under precise and restricted conditions) has arguably no parallel with necessarily local phenomena such as Condition A of the Binding Theory (see Sportiche 2017 and references therein).

Even though (243) does not manage to account for the sentence in (241), it is possible to propose two alternate delayed derivations where the same multi-component described for (243) is used anew, namely (245) and (246). In particular, one could question whether the former effectively ensures the desired locality requirement between *John* and *himself*, as nothing in (245) provides a straightforward reason to claim that this desideratum is met. Is it

enough for an auxiliary tree to adjoin to a member of a multi-component Σ so as to be also in a local relation with the other member of Σ? Answering this question in the affirmative would mean to state this claim as an axiom (just like CETM or the Markovian property previously discussed in this chapter) that must be further proven not to overgenerate. Alternatively, one could posit that *β.appear* adjoins to the main clause before its structure splits into two different trees forming a multi-component, but this option works only if one further argues that TAG operations can be interspersed in the course of (non-monotonic) derivations driven by Merge (see again sections 5.1.5 and 5.3.2).

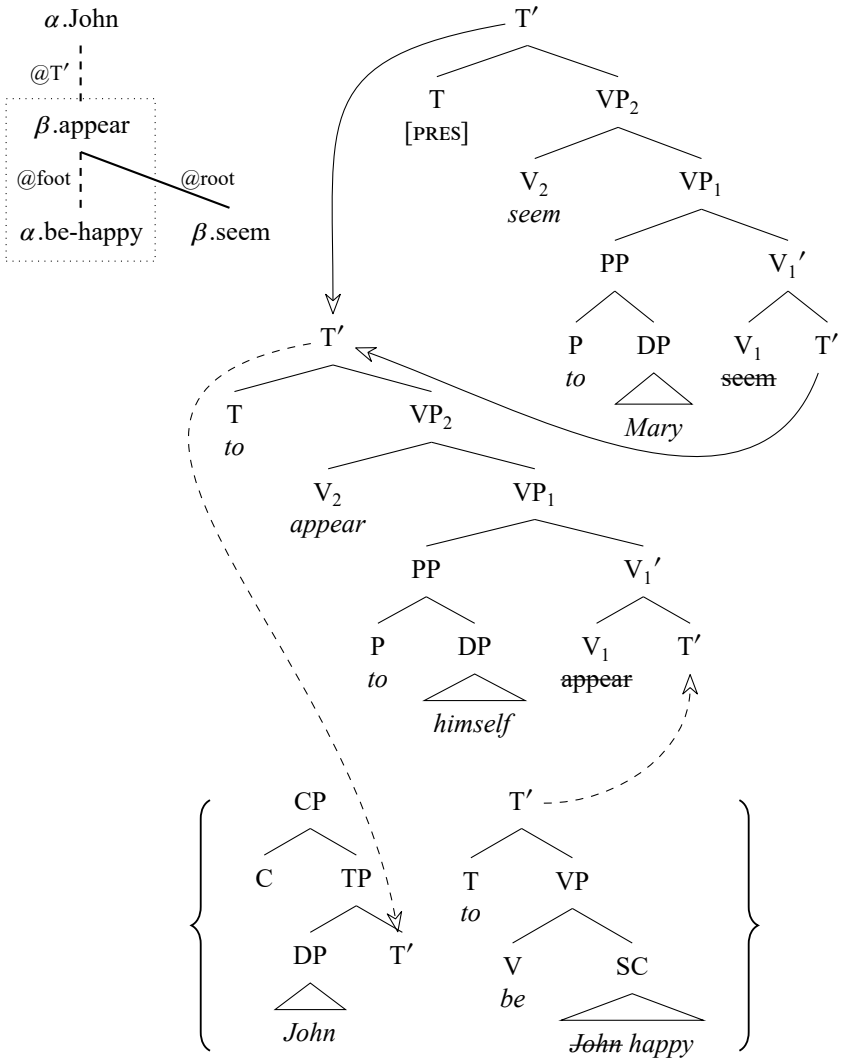
(245)



The delayed derivation in (246) satisfies the locality requirement that (241) demands in a way that is arguably less stipulative than (245). In fact, *β.appear* enters a local relation with the DP *John* by means of an unmediated substitution. Moreover, this locality is confirmed by

the derivation tree in (246), where *β.appear* is immediately dominated by *α.John* whereas *α.be-happy* and *β.seem* are not.

(246)



Although the problems raised by Georgi (2019) can arguably be solved by generalizing the strategy applied in (246), the latter representation brings at least two new concerns to the table. The first is that its derivation tree manages to effectively percolate the address T′ also to *β.seem*, even though the latter is not part of the chain of substitutions as it is defined in (238). The potential relevance of this observation must be investigated by future theoretical research. In fact, it may eventually be the case that both definitions in (238) and (240) need some complementation that manages to capture scenarios like (245) and (246), where an auxiliary tree adjoins to a tree involved in a chain of substitutions.[173]

[173] One must also observe that, in opposition to (242) and (243), (245) not only does not seem to be weakly

The second problem raised by (246) is that it exploits two different TAG operations so as to have the same type of elementary trees (i.e. auxiliary trees anchored by raising verbs) ultimately combined to the main clause. On the one hand, $\beta.appear$ is involved in a chain of substitutions. On the other hand, $\beta.seem$ adjoins to $\beta.appear$. This fact opens up to three different options. The first is that the derivation in (246) needs a slight modification so as to have the two auxiliary trees combined through the same operation. The second aims to establish the desired local requirements for the sentence in (241) in a different manner. Lastly, the third boils down to claim that the derivation in (246) is not problematic at all, but rather descends from linguistic facts that a 2-delayed 2-MC-TAG manages to account for with no further stipulation.

These three possibilities will now be explored in turn, starting from the first, which can be pursued by replicating a strategy already adopted in section 5.4.1 for the sentences discussed on page 305. It consists in (the arguably *ad hoc*) addition to the derivation of a degenerate δ tree that is recursive at the root node of one of the auxiliary trees, thus forming a multi-component with the latter (which, as a result, can no longer be included in a singleton set). In the case of (246), this can be done with respect to $\beta.appear$, which is thus paired with a δ recursive at T′, as (247a) illustrates on the following page.

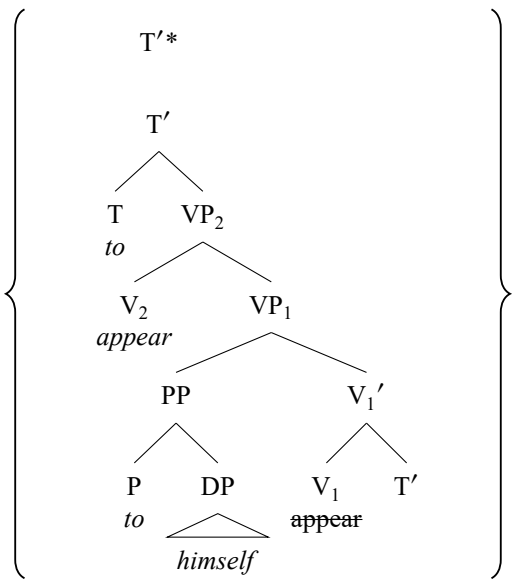
Therefore, a new alternate derivation tree can be proposed in (247b), which is well-formed as long as one follows Frank (2002: Sec. 5.6-5.7, 6.1) in assuming that degenerate trees are peculiar auxiliary trees. Given that all the elementary trees involved in (247b) are taking part to a chain of substitutions, (247b) can be seen as a revised alternative of (243) where $\delta.T'*$ acts as an intermediary between $\beta.seem$ and $\alpha.John$. By replicating the assumption already discussed for (245), $\alpha.John$ can be argued to be in a local relation with $\beta.appear$ because the former is combined with a member of the same set of trees which includes the latter, i.e. $\delta.T'*$.

The alternative that has just been proposed relies on two stipulations, namely the mechanism which ensures the locality relation between *John* and *himself* and the presence of a degenerate tree which makes all raising verbs be part of a chain of substitutions.[174] One could

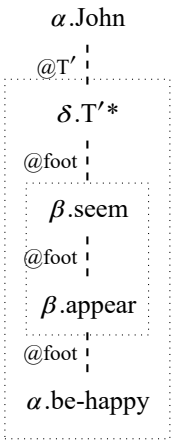
equivalent to (246), but also takes on all the issues that the main text points out for (246). The contrast between (245) and (246) also helps to clarify that two strongly equivalent chains of substitutions are no longer so if an auxiliary tree of a certain type (e.g. a rasing predicate) adjoins at different points, so that the hypothesis in (244) does not hold among such modified chains.

[174] One could further object that at this point it is not clear why $\beta.seem$ is still the only member of a singleton set, if this does not hold for $\beta.appear$. And in fact, one could assume the presence of another $\delta.T'*$ that forms a multi-component together with $\beta.seem$, and that the latter substitutes to the former without the

(247) a.



b.



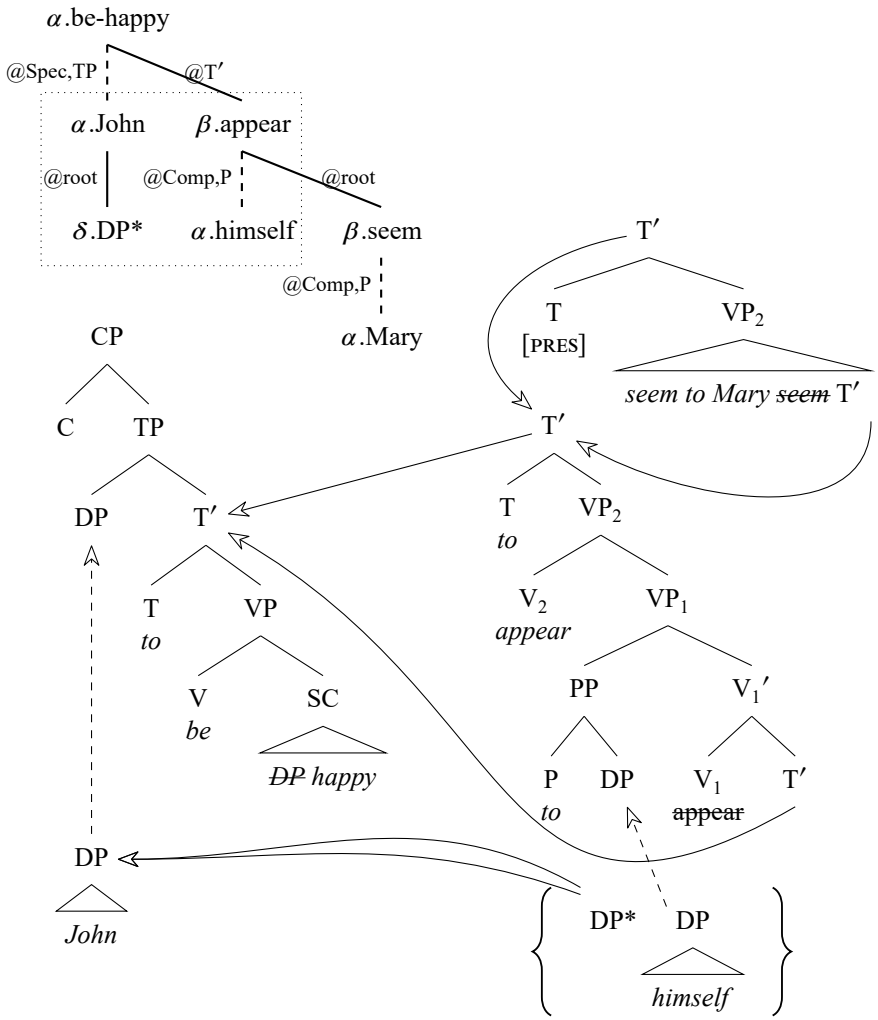
thus be tempted to explore another way to make TAG derivations account for the locality requirements for Condition A of the Binding Theory. This can be done by resorting to Chiang & Scheffler (2008: 22-23), who use a 2-delayed 2-MC-TAG to account for the same relation in the following case of raising to object/exceptional case marking (see a.o. Krivochen 2023a and references therein for related discussion on this syntactic phenomenon):

(248) John$_i$ believes himself$_i$ to be a decent guy.

Chiang & Scheffler (2008) generally assume that, in each tree anchored by a lexical verb, the arguments selected by the latter substitute to the structural positions in which they are overtly realized (see again sections 5.1.1 and 5.1.5). Furthermore, raising verbs must combine to the clause they eventually target as described in section 5.2.1 (namely by means of a chain of adjoining operations), because chains of substitutions have never been proposed before the present work. The machinery that can be transposed to the sentence in (241) hinges on these bases by positing a multicomponent set which includes the two trees *α.himself* and *δ.DP**. The former substitutes at *β.appear*, while the latter adjoins at the root of *α.John*, which in turn substitutes to *α.be-happy*. The whole derivation is thus shown in (249) on the next page, to the exception of the substitution of *α.Mary* to *β.seem*, which is indicated only in the derivation tree.

need of an intermediary tree. This would improve the consistency of the derivation, at the cost of adding a third stipulation.

(249)



The adequacy of the derivation in (249) follows by assuming once again that *John* and *himself* are in a local relation because the former combines to a component of the same set which includes the latter. The difference between such an alternative to (246) and the one presented in (247) ultimately lies in the trees involved in the mechanism ensuring that *John* and *himself* be locally bound in (241). This fact makes (249) as stipulative as (247), with the aggravation of being easily prone to overgeneration. In fact, establishing the locality directly between an anaphor and an R-expression is done at the cost of remaining guiltly silent on the structural distance between them. Hence, an approach *à la* Chiang & Scheffler (2008: 22-23) predicts the following sentence to be grammatical, contrary to fact:

(250) John$_i$ seems to claim that himself$_i$ is happy.

Given the issues highlighted for the two alternatives to (246) that have just been described,

it is time to explore the hypothesis that (246) is not only the best account for the sentence in (241), but that the major problem that it raises actually provides some relevant linguistic insight. In this respect, it is worth turning the readers' attention for a moment to the way in which Longenbaugh (2014: Sec. 3.4) analyzes the "extraction of dative experiencers passed raised subjects", which is an arguably related matter. Instead of reporting the illustratory example that Longenbaugh (2014: 54) provides, (251) shows an adaptation of a sentence from Nakajima (1992: 316), as the latter features the same embedded clause used across the various structures discussed so far:

(251) To whom does John seem ~~to whom~~ to be happy?

Given that neither L-TAGs nor TL-MC-TAGs can account for grammatical sentences like (251), Longenbaugh (2014: 55) argues that "the[ir] failure [...] to derive a particular example should be viewed not as a shortcoming [of TAGs] but as a prediction" indicating that things may not be what they first seem, so that (251) can eventually be accounted for by resorting to an L-TAG *à la* Frank. To motivate this claim, Longenbaugh (2014: 54-55) starts from the well-known truth that raising verbs do not assign external θ-roles, which means that "they don't place animacy or agency requirements on their subjects: expletives, adjuncts, meteorological it, and inanimate DPs may freely appear" in Spec,TP and "subject idiom chunks maintain [their] idiomatic force" (see also Diercks, Santorini & Kroch 2025: Ch. 11 for further discussion of these facts). It is then unexpected to witness the highly degraded grammaticality of the alternate sentences where the dative experiencer is A′-extracted, to the exception of the non-idiomatic reading of (256b). The supporting data are taken from Longenbaugh (2014: 55).

(252) Expletive there:

a. ?There seems to Sophie to be a riot.

b. ?*To whom does there seem to be a riot?

(253) Non-DP subject:

a. ?Under the desk seems to Sophie to be an unusual place to smoke.

b. ?*To whom does under the desk seem to be an unusual place to smoke?

(254) Meteorological it:

a. ?It seems to Sophie to be snowing.

b. ??To whom does it seem to be snowing?

(255) Inanimate subject:

a. ?The door seems to Sophie to be closed.

b. ?*To whom does the door seem to be closed?

(256) Subject idiom chunk:

a. The cat seems to Sophie to be out of the bag.

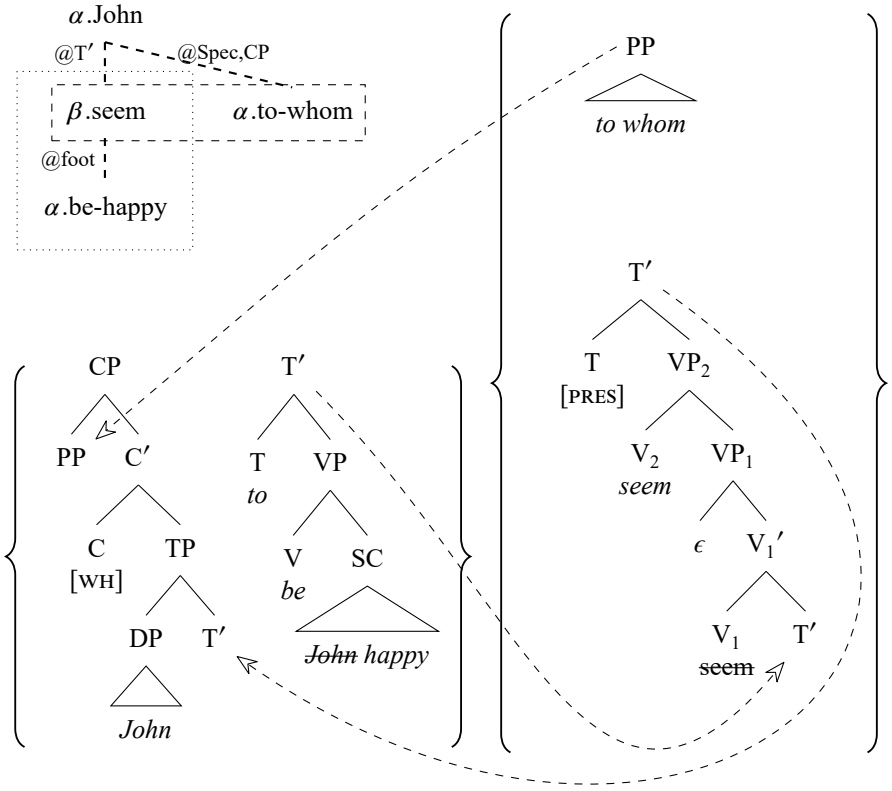
b. %To whom does the cat seem to be out of the bag? [No idiomatic force]

The loss of idiomatic force of *the cat* in (256b) and the worsened grammaticality of the *wh*-fronted dative experiencers in (252)-(255) lead Longenbaugh (2014: 55) to claim that "*seems* assigns an external θ-role and thus is unambiguously behaving as a control verb" (italics in the original). This conclusion consequently expands the number of contexts where it can be argued, following Potsdam & Runner (2001: Sec. 2.2) and references therein, "that predicates like *seem*, *appear*, *look*, and *sound* [...] are ambiguous between a thematic and non-thematic use" (italics in the original).

Let us now try to extend the argument just presented to the derivation in (246) for the sentence in (241), i.e. *John$_i$ seems to Mary to appear to himself$_i$ to be happy*. The fact that β*.seem* undergoes an adjoining and β*.appear* is involved in a chain of substitution can be seen as evidence that only the former is actually behaving as a raising verb, while the latter enters in a control relation with the subject of the main lexical verb. This interpretation would make the derivation on the following page in (257) consistent with what Longenbaugh (2014: Sec. 3.4) argues. In fact, (257) resorts to a chain of substitutions to account for the A′-movement of the dative experiencer in the sentence in (251) taken from Nakajima (1992: 316). Given that β*.seem* is included in two delays in the derivation tree in (257), a k-delayed h-MC-TAG with $k \geq 2$ and $h \geq 2$ is necessary to provide the desired structural description.

As already pointed out in section 5.3.2, no copy of the PP is left behind in the elementary tree in which it has been first-merged, and the rest of β*.seem* is not affected by its removal. Thus, the previous presence of the dative experiencer in the auxiliary tree can only be inferred from the selectional requirements that were satisfied and the features that were checked during the derivation of the same tree. However, (257) bears some limitations. Above all, it

(257)



remains unclear why the Spec,CP in *α.John* should be labeled in the same manner as the projection of *to whom*, if no prior relation between the two initial trees has been established. This is a recurring problem, as it had already emerged for the derivation (218) on page 289. However, a solution must be postponed to section 5.4.4, as the focus of the next section will be directed towards the implications brought by the derivations in (246) and (257).

5.4.3 On the interaction between Merge and TAG operations, pt. 3

The analysis provided in the previous section for multiple raising predicates strongly relies on the chain of substitutions formulated in (238) on page 310. Vijay-Shanker (1992: Sec. 5) and Joshi (2014: 142-143 fn. 7) provide observations that can be taken as forerunners of such a theoretical tool. A similar claim can be made with respect to Kuhlmann & Satta (2012: 623) and Frank (2013: 259 fn. 4). If on the one hand the former work will be discussed later on

in this section, on the other hand Frank (2013: 259 fn. 4) can be quoted on the spot:

> Because the other TAG operation, adjoining, can only introduce a single instance of recursion into a structural domain, substitution is necessarily implicated in recursive structures which are multiply recursive [...].

The affirmation above is compatible with chains of substitutions, although it originally refers to a case where a single elementary tree anchored by a lexical verb is targeted by two different clauses, so that a sentence that combines the properties of (230b) and (230c) on page 303 is eventually derived.

The eventual support of the structural description in (246) on page 318 as the best account of multiple raising in English came after scrutinizing four potential alternatives, three of which admitted under a 2-delayed 2-MC-TAG. Because of these many alternatives, the latter formalism *prima facie* seems to extend the strong generative power of L-TAGs in ways that give rise to unwelcome underdetermination. However, further reflection on the adequacy of the thus obtained structural descriptions has allowed to overcome this issue by excluding some of the candidates on several grounds and come to a single structural description for a certain target sentence.

Furthermore, the availability of several alternative accounts for a certain construction should be seen not as a hindrance, but as an advantage, because it adds some interesting potential to the formalism. In addition to an Optimality-Theoretic implementation (see A. S. Prince & Smolensky 2002–1993, A. Prince & Smolensky 2004 and related works), 2-delayed 2-MC-TAGs offer a new perspective on three distinct linguistic aspects, namely (i) the crosslinguistic syntactic variation, (ii) the seemingly puzzling behaviors of structures belonging to a single language, and (iii) the fact that native speakers of one single language may give conflicting grammaticality jugdgements for the construction being investigated. Although 2-delayed 2-MC-TAGs share the first two points with other weakly equivalent formalisms (by revisiting their structural descriptions for the better), the third stems by observing that competing structural descriptions like the ones discussed in the previous section predict different grammaticality judgements with respect to the same sentence. If speaker variability does emerge with respect to such a construction, then it may not (only) be the result of non-grammatical constraints on acceptability, as argued by Bruening (2018b), but rather (or also) follow from different closely-related grammars.[175]

[175] In this respect, Julie Doner (p.c.) observes that she is not sure whether she would accept with no qualms

The crucial turn in favor of (246) on page 318 was done by following Longenbaugh (2014: Sec. 3.4), who argues that, contrary to expectations, raising verbs selecting a dative experiencer must be control verbs when the argument is *wh*-extracted. Since his analysis is based on an L-TAG *à la* Frank (2002), the embedded lexical verb is then supposed to bear a PRO subject. Both the *wh*-element and the subject of the whole clause must be instead theta-marked by the control counterpart of what at first glance appears to be a raising predicate, and the elementary tree anchored by the latter must be a full-fledged CP (Longenbaugh 2014: 56). Arguably, other raising verbs can be adjoined further to such a structure in the manner assumed in Frank (2002). This means that, on the one hand, a raising verb can target the T′ node in the extended projection of the control verb if it also scopes over it. On the other hand, a raising predicate must shoehorn the equivalent counterpart projected by the embedded lexical verb if its scope is instead lower.

If the discussion presented towards the end of section 5.4.2 is on the right track, then the conclusions that Longenbaugh (2014: Sec. 3.4) comes to are accounted for in a surprisingly different manner. No PRO and no copy of the controlling element were necessary so as to ensure that some raising verb has the control properties that an L-TAG would fail to account otherwise. A 2-delayed 2-MC-TAG manages to hint at the same type of analysis by resorting to the twofold options in which auxiliary trees can be combined with the target structure. A raising predicate can either be adjoined if it bears no control property, or become (part of) the body of a chain of substitutions if it does.

Based on what has just been reaffirmed, one could wonder if chains of substitutions can give a «control flavor» to a raising predicate, or if they rather involve control verbs in the strict sense to begin with. If the latter turned out to be the case, it would be easy to reach an explanation for the control relations holding in the derivations presented on page 297 in section 5.3.3 with respect to the clitic climbing phenomenon. However, the broader theoretical consequences would be far-reaching, because this scenario would imply a thorough revision of the θ-criterion and lead to interesting questions with respect to the exact role that substitutions play in 2-delayed 2-MC-TAGs. Although this must necessarily be another aspect left for future research, one can nonetheless acknowledge that the language-internal variation con-

the sentence discussed in section 5.4.2 following Castillo, Drury & Grohmann (2009–1999: 93), i.e. *John$_i$ seems to Mary to appear to himself$_i$ to be happy*. She further argues that the *wh*-extraction of any of the two dative experiencers from such a sentence would be ungrammatical to her. However, she also points out that these "are tricky grammaticality judgments in terms of processing load. [... T]hese are the kinds of sentences that you should probably ask more than one person about, and set up a context". Hence, an inquiry made along lines of Bruening (2018b) and references therein awaits.

cerning what are usally classified as raising verbs finds interestingly comparable outcomes with respect to another phenomenon, i.e. the crosslinguistic variation of auxiliary chains.

The title of Wurmbrand (1999) summarizes her arguments supporting the fact that, in the Germanic languages, "modal verbs must be raising verbs". As long as Frank (2002, 2004, 2006, 2010, 2013) assumes that raising predicates can anchor their own elementary trees by way of derogation from CETM in (197) on page 270 (as discussed in section 5.2.1, raising verbs do not project a complete extended projection), Wurmbrand's statement implies that modals have the right to do the same in TAGs. However, Krivochen & Padovan (2021: Sec. 3.1) show that the observations made by Wurmbrand (1999) do not hold for Spanish, where modals have a «control flavor» which is comparable to what has been discussed until recently following Longenbaugh (2014: Sec. 3.4). In this respect, it is profitable to quote Krivochen & Padovan (2021: 10) verbatim:

> Summarising a lot of previous work, [...] (i) Spanish modals, unlike English modals, are morphologically unremarkable verbs, thus (ii) they can be selected by other auxiliaries (when they are, they «absorb» the information conveyed by auxiliaries to their left [...]), and (iii) it has been argued that they can be thematic assigners, just like lexical verbs. We know they are auxiliaries because they can combine with meteorological verbs and because of their semantic contribution to the compositional interpretation of the structures in which they appear (epistemic/deontic/dynamic in the most widely accepted classifications).

In order to capture the differences between English and Spanish modals, Krivochen & Padovan (2021: Sec. 3) assume that the former do not anchor their own elementary trees, whereas the latter do. By following Krivochen & Padovan (2021: 6), each head that anchors an elementary tree "must also be a predicate: it must, in GB terms, have a thematic grid and a subcategorisation frame". Moreover, the related semantic property that allows to identify a lexical predicate is that the latter "can both modify another object and be modified itself" (Krivochen & Padovan 2021: 7). To illustrate this, Krivochen & Padovan (2021: Sec. 3.1) take as an example the Spanish sentence in (258), where no contradiction arises because the lexical verb *leer* is not affected semantically by the perfective auxiliary *ha* since another lexical predicate, namely the deontic modal *debido*, interferes between the two.

(258) *el* *senador* *ha* *debido* *leer* *un* *discurso,*
DET.M.SG senator.M.SG AUX.PFV.3SG must.PTCP read.INF INDF.M.SG speech.M.SG
pero *no* *lo* *ha* *echo.*
but NEG 3SG.M AUX.PFV.3SG do.PTCP
'The senator has had to read a speech, but he hasn't done so.' [Spanish]

The 2-delayed 2-MC-TAG approach proposed in this work does not deny the insightfulness of what has just been reported. However, the variation in the behavior of modals need not be captured by dividing them into those who can anchor their own tree (by being lexical predicates) and those who cannot. The crucial difference is rather captured by having each auxiliary tree combined either by means of a chain of substitutions or by means of a chain of adjoining operations. In other words, lexical predicates are not the only possible anchors for elementary trees, but it can be argued that they always anchor trees that are eventually combined by means of (chains of) substitutions.

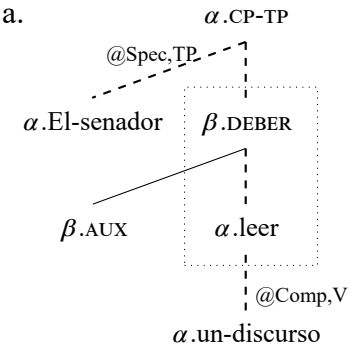
What has just been argued can be illustrated by pointing out that the behavior of a modal used with a certain reading (e.g. deontic, as in 258) does not vary only crosslinguistically, but also language-internally. In fact, in the Italian sentence (259a) the verb *deve*, i.e. the properly conjugated form of the deontic modal *dovere* in the present tense, interacts with another lexical verb as in the Spanish sentence in (258), hence leading to no contradiction. However (at least according to the intuitions of the writer and of other consultants), in a construction which is seemingly closer to (258) by having the auxiliary *ha* added (and the deontic modal in the participle form), a contradiction emerges due to the inability of the auxiliary itself to affect the modal to the exclusion of the verb *leggere*. The latter case is reported (259b).

(259) a. *Vera* *deve* *leggere* *un* *discorso,* *ma* *non* *lo* *fa.* [Italian]
Vera must.PRS.3SG read.INF INDF.M.SG speech.M.SG but NEG 3SG.M do.PRS.3SG
'Vera has to read a speech, but she does not.'

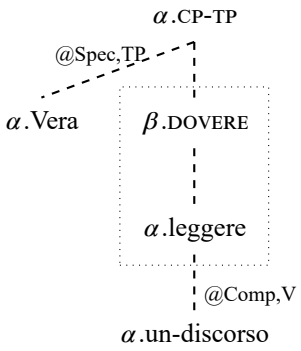
b. *Vera* *ha* *dovuto* *leggere* *un* *discorso,* # *ma* *non*
Vera AUX.PFV.3SG must.PTCP read.INF INDF.M.SG speech.M.SG but NEG
lo *ha* *fatto.*
3SG.M AUX.PFV.3SG do.PTCP
'Vera has had to read a speech, # but she hasn't done so.'

The representation in (260) shows how a 1-delayed 2-MC-TAG allows to depart from Krivochen & Padovan (2021) and assume either a raising-flavored or a control-flavored interpretation of modals while still having them always anchored by their own elementary trees. In particular, (260a) and (260b) capture the control-flavored cases in (258) and (259a), whereas (260c) accounts for the raising-flavored interpretation in (259b).

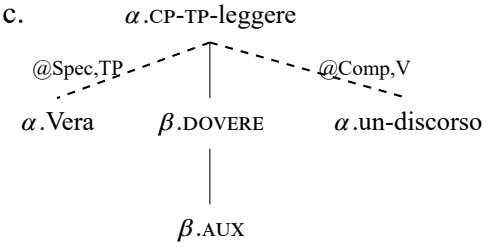
(260) a.



b.



c.



The representations in (260) depart from those discussed in section 5.4.2 by showing all the substitutions and the adjoining operations taking place for each derivation (tree). This allows to tackle a point that has not been addressed before, namely that the root of a derivation tree can be a non-lexicalized part of the extended projection of the clause with which it forms a multi-component. As already mentioned in section 5.3.1, this is not a problem for MC-TAGs because in such formalisms "lexicalization means that every tree set in the grammar contains at least one lexicalized tree. It does not require all trees in the grammar to be lexicalized" (Kallmeyer 2009: 80).

Therefore, *leer* and *leggere* not only account for the lexicalization of *α.leer* and *α.leggere* in (260a) and (260b), respectively, but also for the initial trees rooting the two structures (both indicated as *α.CP-TP*), which for their part end up being the destination of the delay that they help introduce. The equivalent singleton of these non-monotonically derived elementary trees (see again section 5.3.2) is the root of (260c), namely *α.CP-TP-leggere*. The last point being

made also clarifies that the subjects in (260a) and (260b) are theta-marked and subcategorized by the verbs *leer* and *leggere*, respectively. This process runs just as easily as it does for α.*Vera* with respect to α.CP-TP-*leggere* in (260c), although the exact details in this regard depend on whether one supports either the strong or the weak compartimentalization between the Merge and TAG operations discussed in section 5.1.5.

The derivation trees in (260) also depart from Krivochen & Padovan (2021: Sec. 3) in other two respects that will be discussed further in section 5.4.4. First, in (260) the perfective auxiliaries and the deontic modals are represented in smallcaps to suggest that they may not bear their own inflection before combining to the targeted tree (a point which will be explored further in section 5.4.4). Second, since the bijective relation between lexical predicates and anchors of elementary trees is no longer needed, the question of whether each perfective auxiliary can anchor its own tree as suggested in (260) becomes an orthogonal matter.[176]

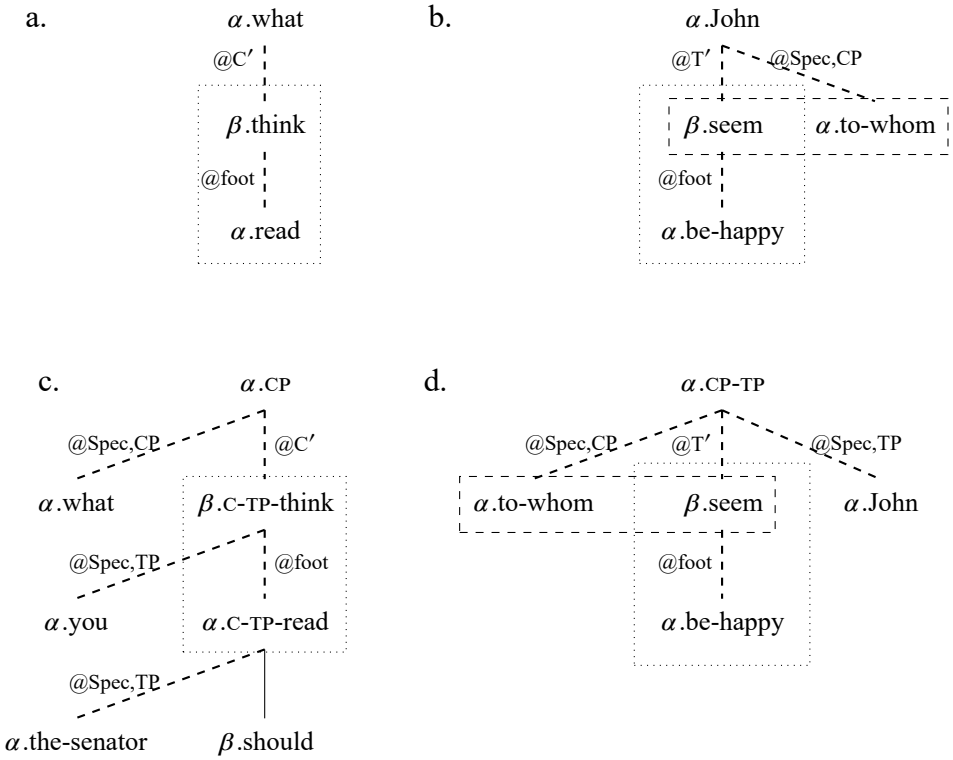
The observation made with respect to (perfective) auxiliaries also extends to the arguments that are subcategorized by *leer* and *leggere* in (260). In fact, Krivochen & Padovan (2021: 27 en. 2) explicitly point out that nominalizations are lexical predicates whereas "proper N, non-nominalised common N" are not (being rather defined as "lexical non-predicates"). Since Krivochen & Padovan (2021) argue that only lexical predicates anchor elementary trees, neither *el senador*, nor *Vera*, nor *un discurso* would be separate elementary trees in their account. The insights from Krivochen & Padovan (2021) and the assumptions posited in section 5.3.1 following Frank (2002, 2004, 2006, 2010, 2013) can thus be reconciled by means of *k*-delayed derivation trees like those in (260), where any kind of nominal extended projection represents an elementary tree of its own. What is more, these premises will be pivotal in the analysis of U20 that will eventually be presented in section 5.5.1.

The availability in the derivation trees of roots with no lexical materials is what allows to solve an apparent inconsistency in the *k*-delayed derivations of *wh*-questions previously discussed. To illustrate the problem, (261a) reports below the derivation tree proposed at the end of section 5.4.1 for *What do you think the senator should read?*, whereas (261b) does the same for the analysis of *To whom does John seem to be happy?*, which was presented at the end of section 5.4.2. It can be seen that in (261a) the *wh*-phrase seemingly lexicalizes the

[176] Auxiliaries can thus be accounted in three different ways. The first is the one where each of them anchors its own tree as in (260), the second consists in having (at least some of) them already included in the extended projection of the verb anchoring the main clause, whereas the third is the strategy supported by Krivochen & Padovan (2021: 11) for the sentence in (258), where the inflected auxiliary is included in the tree anchored by the deontic modal.

root of the derivation tree, whereas this does not happen for the *wh*-counterpart in (261b). In this respect, the latter turns out to be on the right track, although both structures need to be reworked. The derivation in (261c) corresponds to the revised version of (261a), ditto for (261d) compared with (261b). Section 5.4.4 will however discuss that (261c) is still not the ultimate derivation for *What do you think the senator should read?*, due to some issues in the derived tree that emerged at the end of section 5.1.2 and that the present chapter has not addressed yet.

(261)



Given the importance that Longenbaugh (2014: Sec. 3.4) plays in the discussion done so far with respect to control-flavored syntactic objects, it is now worth evaluating whether his approach has the potential to defy the 2-delayed 2-MC-TAG account proposed in section 5.4.2. As anticipated at the end of section 5.2.1, Longenbaugh (2014: Sec. 2.1) elaborates on an L-TAG *à la* Frank (2002) by proposing a Synchronous TAG where the LF and the PF objects are derived in parallel.[177] Although the LF component is assumed to exploit

[177] The discussion at the beginning of section 5.3 highlighted that the parallel derivation of two distinct structures by means of a STAG is possible only if a common derivation tree is assumed for them. Longenbaugh (2014: 23) argues that such a "reliance is implicit rather than overt" in STAGs. By leveraging this implicit, Longenbaugh (2014: 24) claims that, contrary to what the T-model and its subsequent variants state (see S. Müller 2023b: 88, 128-129, 150 and references mentioned there), it becomes possible to avoid "bifurcating a single structure that yields both the PF and LF objects". This latter move is arguably fallacious, as an

the power of a 2-TL-MC-TAG, Longenbaugh (2014: 40) strongly opposes the need of an MC-TAG for the derivation of the PF object and adopts two main strategies to account for phenomena that defy the supposed adequacy of an L-TAG *à la* Frank (2002). The first is to relegate certain problems as post-syntactic phenomena, with the idea that the MC-TAG accounts provided by Frank (2002: Sec. 5.7, 5.8, 6.1) can be dispensed with as soon as such issues no longer fall under the umbrella of the explananda. The second is to claim that certain phenomena are not what they *prima facie* seem, as the *wh*-fronting of an experiencer theta-marked by a raising verb, discussed at the end of the previous section. Hence, an alternate analysis that complies with L-TAG is not only possible, but even suggested by the fact that such a formalism would not manage to capture the facts otherwise.

Even though the points just reported following Longenbaugh (2014) are rather intriguing, it can be seen that the L-TAG approach he adopts is weaker than the formalism proposed here with respect to both generative and explanatory power. One first reason motivating this claim is that the L-TAG used by Longenbaugh (2014) would suffer from the same issues pointed out by Georgi (2019) and discussed in the previous section. Another empirical issue where 2-MC-TAGs fare better than L-TAGs is long distance *wh*-movement, as in *Which car does Sally wonder how to fix?*, which was analyzed in (218) on page 289 after revising the analysis provided by Frank (2002). In this respect, Longenbaugh (2014: 54 fn. 16) is forced to admit that he has "nothing to say about these data".

A second type of objection to the approach adopted by Longenbaugh (2014) is argumentative in nature. Longenbaugh (2014: 32-33) claims that (as far as the PF object is concerned) MC-TAGs must be discarded in favor of L-TAGs by relying on a proof where a multi-component derivation eventually generates illegitimate island-violating derived trees. Crucially, this scenario is claimed to be possible because multi-component sets of trees do not need to comply with the well-formedness conditions that an L-TAG *à la* Frank (2002) imposes on the single elementary trees.

The claim just presented is flawed on a twofold ground. First, the versions of FTH, CETM and the theta-criterion that have been adjusted for *h*-MC-TAGs since page 270 effectively rule out the multi-component set that Longenbaugh (2014: 32-33) illustrates in favor of his

approach *à la* Longenbaugh (2014) ought to claim instead that the common derivation tree that regulates both the PF and LF objects is what narrow syntax truly is. In this respect, it is also interesting to observe that this revised architecture of Longenbaugh's grammar would be conceptually similar to what Wiltschko (2014: 324) discusses and would provide a defined collocation for the independent principle presented in section 5.2.2 (and to be discussed again in the remainder of the chapter).

argument. Moreover, these restrictions are strengthened by the principle (211) proposed in section 5.3.2, which establishes a structural equivalence between a single legitimate elementary tree and a multi-component by seing the latter as a result of a non-monotonic derivation *à la* Heck (2016, 2023).

The second flaw in the argumentation formulated by Longenbaugh (2014: 32-33) is that, although one can design an MC-TAG which generates the multi-component set that he puts forth (as well as many more derived trees that break common syntactic assumptions), the same could be made with some eventually inadequate L-TAG, as discussed for instance with respect to the Markovian property in section 5.1.4. By following the same line of reasoning pursued by Longenbaugh (2014: 32-33), one should therefore conclude that L-TAGs (or any other formalism, for that matter) are not a desirable theoretical tool either. To get out of this paradox, it is necessary to recognize that the blame for unsatisfactory structural descriptions is not necessarily a consequence of the generative power of the chosen formalism, but can also be due to the bad well-formedness conditions assumed in the first place, as stated by Bleam (2000: 193) in the quotation reported on page 212. The issue raised by Longenbaugh (2014: 32-33) thus tends to be a matter of the constraints that the linguists posit for formal systems they choose. And in fact, it has been shown from section 5.3.1 on that the careful formulation of formal constraints is a pivotal concern when one needs to tailor the most adequate *h*-MC-TAG.

The third and last counterargument to the reasoning that makes Longenbaugh (2014: 32-33) reject *h*-MC-TAGs with $h \geq 2$ in favor of L-TAGs is theoretical. Svenonius (2024: Sec. 3.4) takes up the analysis that Haegeman & Riemsdijk (1986) originally put forth for West Germanic verb clusters and argues that their peculiar syntactic behavior could be a consequence of "processes that affect sequences of verbs at the postsyntactic level of phonological form (PF)" (Svenonius 2024: 16). This theoretical move leads to a crucial point, which will now be quoted verbatim from Svenonius (2024: 17):

> The possibility that verb clusters are formed at spell-out, rather than in the syntax, has an interesting consequence for the structure of grammar. The demonstration by Shieber (1985) and others that natural language is not context-free [...] was based on cross-serial dependencies [...]. A context-free grammar cannot keep track of the cross-serial dependencies, where each phrasal dependent to the left must be matched in order with a verb in the sequence to the right. If clusters

> are formed at PF, and cross-serial dependencies do not emerge elsewhere, then the structure that corresponds to the order prior to cluster formation might be described by a context-free grammar.

In other words, shifting the burden of the explanation of certain phenomena to the post-syntactic component does not allow Longenbaugh (2014: 32-33) a safe retreat to an L-TAG *à la* Frank (2002). This is due to the fact that the latter formalism is more powerful than a CFG (as discussed from section 5.1.2 on) precisely because it aims to capture the mildly context-sensitive phenomena that can be accounted for by a CFG which is empowered by means of the post-syntactic processes discussed in Svenonius (2024: Sec. 3.4).

Given that in section 5.1.1 it was observed that CFGs and TSGs are weakly equivalent, one could argue that the best fit for an approach *à la* Longenbaugh (2014) is an empowered TSG, which can be called TSG^{+} for the sake of brevity. As in normal TSGs, TSG^{+} can be conceived as a tree-rewriting system where substitution is the only possible operation. In order to account for the phenomena that Longenbaugh (2014) captures in the ways he does, TSG^{+} must however be empowered by means of two further stipulations. First, the rearragment of heads is seen as a uniform phenomenon (*contra* what will be discussed in section 5.4.4) that is consistently handled post-syntactically, just like in the CFG^{+} sketched by Svenonius (2024: Sec. 3.4). Second, auxiliary trees are allowed, but they can only participate in chains of substitutions.

The latter assumption neatly captures at least two facts that section 5.4.2 made emerge. On the one hand, it explains why both Longenbaugh (2014: Sec. 3.4) and the 2-delayed 2-MC-TAG adopted in the present work manage to come to the conclusion that, under certain circumstances, raising predicates acquire nuances that are typical of control verbs. On the other hand, TSG^{+} would account for two or more raising predicates embedding another lexical verb with the chain of substitution shown in (243) on page 315. This would be in turn consistent with Longenbaugh's belief that he is adopting an L-TAG *à la* Frank (2002), as the latter generates the same structure (which suffers from the shortcomings discussed in section 5.4.2 following Georgi 2019) in a weakly equivalent manner by means of the chain of adjoining operations shown in (242) on page 314.

To summarize, the short detour on the argumentative problems that arise from Longenbaugh (2014) not only allows to review many of the topics covered so far, but also leads to the formulation of a new list of alternate formalisms which are argued to be weakly equivalent in

the literature. By paraphrasing Abeillé & Rambow (2000: 6), the choice among one of such formalisms (262) is up to the linguist's preferences and needs. They are reported in (262), in order of strong generative capacity from the inferior to the superior and by assuming $k \geq 2$ and $h \geq 2$ where k and h are the maximum number of delays (5.3.3) and the maximum arity of a multi-component (as per Boullier & Sagot 2011: Sec. 2.2), respectively:

(262) $\text{CFG}^{+} = \text{TSG}^{+} = h\text{-TL-MC-TAGs} = k\text{-delayed } h\text{-MC-TAGs}$

Before concluding this section, two futher arguments can however be mentioned in support for k-delayed h-MC-TAGs with $k \geq 2$ and $h \geq 2$. The first additional argument that makes CFG^{+}, TSG^{+} and L-TAGs less appealing than h-MC-TAGs (where $h \geq 2$) is computational in nature. Besides the strong and weak generative capacity discussed so far following Joshi (2003: 243), Maletti & Engelfriet (2012: Sec. 1) and references therein, in the computational linguistics community there is in fact another criterion that makes certain formalisms preferable than others. It is called *lexicalization* and consists in "the process that transforms a grammar into an equivalent lexicalized one" (Engelfriet & Maletti 2017: 218).

Similar to what happens with the generative capacity, lexicalization can be either weak or strong. As neatly discussed by Kuhlmann & Satta (2012: 617), *weak* lexicalization corresponds to each procedure that casts the source grammar into another one which preserves the string language while "substantially alter[ing] the structure of the parse trees of the source grammar". One obtains instead a *strong* lexicalization whenever "the sets of parse trees that the two grammars assign to the common string language" are the same (Kuhlmann & Satta 2012: 617). In natural language processing, the latter type of lexicalization is to be preferred because it allows to assume a grammar whose parse tree can be used by other linguistic modules with no risk of formal impairment (see again Kuhlmann & Satta 2012: 617).

What matters for our present concerns is *closure under strong lexicalization*, which consists in the case where a certain subclass of grammars $\mathcal{G}$ strongly lexicalizes itself, so that any lexicalized grammar $\in \mathcal{G}$ can be recast into another one that is still both lexicalized and $\in \mathcal{G}$ (see Kuhlmann & Satta 2012: Sec. 2.2). Crucially, Joshi & Schabes (1991, 1997) assume that L-TAGs are closed under strong lexicalization by relying on "an informal argument" where adjoining is seen as "more powerful than substitution with respect to weak generative capacity" (Kuhlmann & Satta 2012: 618). However, Kuhlmann & Satta (2012: Sec. 4) blatantly disavow this assumption by providing a formal proof that leads to the conclusion that "[L-]TAGs are not closed under strong lexicalization", and neither are TSGs (Kuhlmann &

Satta 2012: Sec. 3). On the contrary, *h*-MC-TAGs with $h \geq 2$ bear the potential to allow strong lexicalization, as demonstrated by Engelfriet & Maletti (2017).

As anticipated at the beginning of the present section, Kuhlmann & Satta (2012: 623) provide a passage that is particularly worth quoting (italics and boldface in the original):

> There is a perspective under which adjunction quite closely resembles substitution [...]. Let us first look at substitution as an operation on the *yield* of the derived tree. Under this view, substitution is essentially context-free rewriting: It replaces a nonterminal symbol in the yield of a derived tree with a new string consisting of terminals and nonterminals, representing the yield of the tree that is substituted. Under the same perspective, [...] just as substitution can be seen as context-free rewriting on tree yields, adjunction can be seen as context-free rewriting on the *paths* of trees: It replaces a nonterminal symbol in some path of a derived tree with a string representing the **spine** of the tree that is adjoined–the unique path from the root node of the tree to the foot node.

The passage just reported contains two aspects discussed in section 5.2.1, namely the addressing of *adjoining* as *adjunction* and a definition of *spine* (which is not to be confused with the one discussed for FOFC in sections 2.3.2 and 2.3.3). More importantly though, the quoted observation is not only pivotal to the formal proof that Kuhlmann & Satta (2012: Sec. 4) give against the strong lexicalization of L-TAGs, but also fosters the chains of substitutions proposed in section 5.4.2.

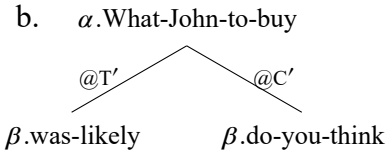
The second important observation that favors the adoption of a 2-delayed 2-MC-TAG is that this formalism "renews the viability of interpreting the derivation tree as a dependency analysis", as quite rightfully remarked by Chen-Main (2012: 106). In fact, it becomes possible to recast the issues pointed out by Frank (2002: 291 en. 3):

> Candito & Kahane (1998) propose a very simple process for computing one aspect of interpretation from the TAG derivation, namely, that of recovering predicate argument relations. They suggest that a Substitution link in a derivation tree be interpreted as a predicate-argument dependency between the elementary tree that is the locus of substitution and the substituted elementary tree, respectively. They take Adjoining to trigger a similar interpretation, but with the dependency reversed, so that the adjoined element is the predicate and the locus of adjoining is the argument.

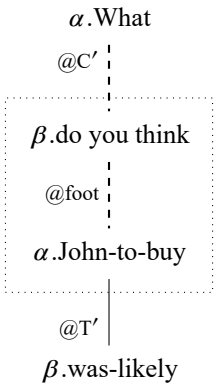
Section 5.4.2 provided an empirical observation in favor of what Candito & Kahane (1998) propose. In fact, the hypothesis in (244) on page 315 argues that a crucial ingredient that makes a chain of substitutions weakly equivalent to a chain of adjoining operations is the reversed order of the auxiliary trees in their respective bodies $\mathscr{B}$. In line with what was reported at the beginning of section 5.2.1, the prediction just quoted runs however into problems in a case like (263a), which Frank (2002: 291 en. 3) proposes as an adaptation of the problematic derivations discussed by Candito & Kahane (1998: Sec. 4) themselves. In fact, Frank accounts for (263a) by means of the L-TAG derivation tree in (263b), but based on Candito & Kahane (1998: Sec. 3) such a derivation tree should lead to an "incorrect interpretation" where "the *buy* predication is an argument of both *likely* and *think*" (Frank 2002: 291 en. 3, italics in the original).

(263) a. What do you think John was likely to buy?

b.



c.

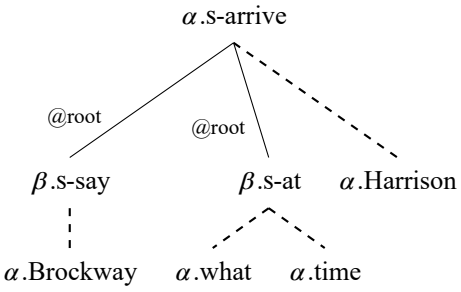


As shown since the end of 5.4.1, the exploitation of a delay and the subsequent introduction of a chain of substitutions allow instead to establish the desired dependencies among the various clauses, so that (263c) accounts for (263a) in a more adequate manner. In fact, (263c) represents a solution to the unexpected absence of a dependency relation between *think* and the raising predicate, which Kallmeyer & Kuhlmann (2012: Sec. 6) and Chen-Main & Joshi (2014: Sec. 6.1) address as the "missing link problem" on the derivation tree. Still, (263c) must be understood again as a simplification, as a more detailed derivation tree should trace what has been shown in (261c) on page 331.

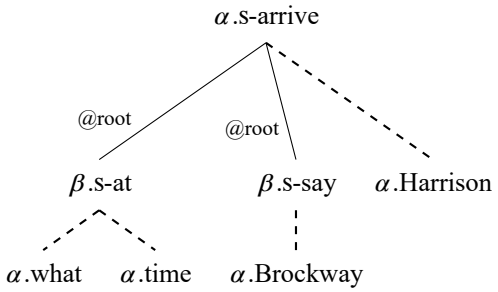
A similar point can be made with respect to the derivation trees discussed by Gardent & Narayan (2015: 54). Since their attention is directed towards the extended notion of derivation first given by Schabes & Shieber (1992, 1994) and previously discussed in section 5.2.1, Gardent & Narayan (2015: 54) propose the derivation trees in (264) as a counterexample

to the outermost predication rule posited by Schabes & Shieber (1994: 102), according to which a predicative auxiliary tree must be the rightmost sister in the derivation tree whenever a mixed derivation is generated (for the details, see page 253 and the referenced works). Although the sentence in (264a) can receive a twofold interpretation, the added brackets highlight that the two derivation trees in (264) should account for the same dependency, on the basis of which "it is the time of arriving (rather than the time of saying) which is questioned" (Gardent & Narayan 2015: 54).

(264) a. At what time [did Brockway say] Harrison arrived?



b. [Did Brockway say] at what time Harrison arrived?



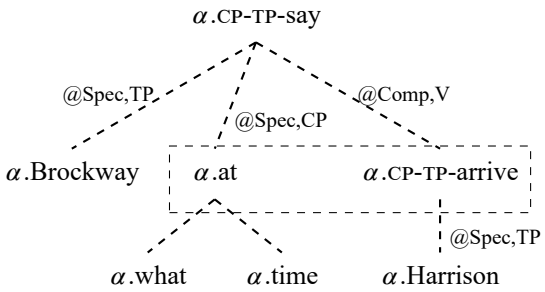
As highlighted by the additional labels, *α.s-arrive*, *β.s-at* and *β.s-say* all bear a S root in (264), so that the latter two can both shoehorn *α.arrive* at its root. For such an analysis to work, two conditions must be met. First, the C-T-v-V system must not be adopted, as discussed in section 5.4.1. Second, the extended notion of derivation introduced in section 5.2.1 (following Schabes & Shieber 1992, 1994) establishes that the two auxiliary trees in (264) must form either an independent or a mixed derivation. The latter alternative must be the correct one, for a twofold reason. On the one hand, *β.s-say* bears the typical traits of a predicative auxiliary tree established by Candito & Kahane (1998: Sec. 2) and Gardent & Narayan (2015: 45), i.e. it is anchored by a lexical verb which selects another clause (namely the extended projection of *arrive*) as one of its arguments. On the other hand, *β.s-at* must be a modifier (auxiliary) tree. Otherwise, it would not be allowed to adjoin at the same node that is targeted by *β.s-say*.

Crucially, Gardent & Narayan (2015: 54) propose the pair of derivation trees in (264) to show that the outermost predication rule assumed by Schabes & Shieber (1994: 102) must be dispensed with. Given that both *β.s-at* and *β.s-say* must be left auxiliary trees, the latter principle predicts in fact (264b) while ruling out (264a), because in (264a) the rightmost sister (which thus must be the highest in the derived tree) is not a predicative tree.

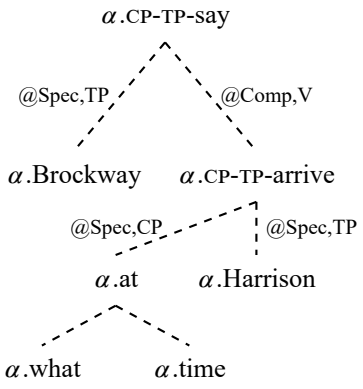
The 2-delayed 2-MC-TAG approach adopted here allows to argue that the argument just presented against the outermost predication rule is actually ill put. Both (264a) and (264b) manage to capture the desired dependency between the *wh*-phrase and *α.s-arrive*. However, this is done at the cost of wrongly making the two structures almost identical in their entirety (i.e. as if they were an example of *constructional hominimity* in the sense of Chomsky 2002–1957: 28, 86-87), even though they should be quite different by virtue of the fact that the sentence in (264b) is a polar question, while the counterpart in (264a) is not.

A counterproposal to the representations in (264) must capture the aforementioned desired dependency while also accounting for the completely different nature of the two sentences. This can be done by accepting uncritically the structure that Gardent & Narayan (2015) assume for the *wh*-phrase *at what time*, as the latter is not a central concern. Rather, the first important step consists in assuming that both auxiliary trees in (264), i.e. *β.s-say* and *β.s-at*, are initial trees instead. On the one hand, *α.at* is thus made capable of substituting to some Spec,CP. On the other hand, the initial tree anchored by the verb *arrive* substitutes at a leaf node which is subcategorized by the verb *say*. As a result, the provisional representations in (265) improve those in (264) by having the tree anchored by *say* correctly dominate the one lexicalized by *arrive*. The distant relation between the latter tree and the *wh*-phrase is further captured by having them included in the same multi-component and tree-locally combined by means of a delay.

(265) a. At what time [did Brockway say] Harrison arrived?



b. [Did Brockway say] at what time Harrison arrived?



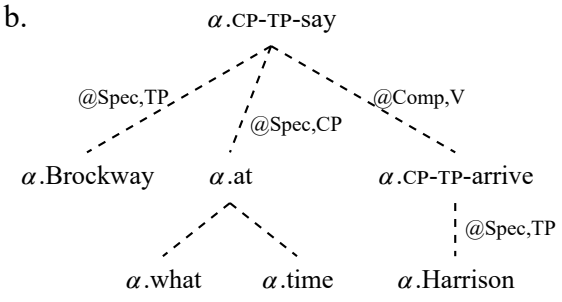
The derivation tree in (265a) suffers however from a major problem. If the elementary tree anchored by *arrived* projects up to CP, why does the *wh*-phrase fail to merge locally at Spec,CP, thus disobeying the Maximal Checking Principle stated by Frank (2002: 126)

and discussed on page 233 (in section 5.1.5)? An answer in this respect will be provided on page 350 (in section 5.4.4), where the major insights emerging from the present counterproposal to Gardent & Narayan (2015) will be kept while reworking the problematic details of (265a).

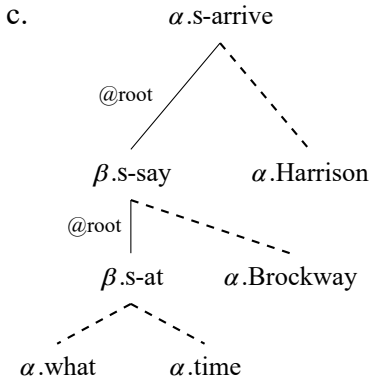
For the time being, it is worth mentioning another fact supporting the idea that the derivations in (265) improve the analysis presented in (264) following Gardent & Narayan (2015: 54). It consists in seeing how the two competing proposals capture the other possible reading of the sentence in (264a) and (265a), namely the alternative in (266a) where it is the time of saying which is questioned, rather than the time of arriving. Under the *k*-delayed approach adopted here, the structure in (266b) is identical to the one in (265a), except for the crucial absence of a delay, which goes to highlight that in (266b) there is no dependency between α.*CP-TP-arrive* and α.*at*. As a result, the latter ends up being theta-marked by α.*CP-TP-say* in (266b), and the ambiguity of a the string in (265a)-(266a) is both mirrored and captured in the corresponding derivation trees in (265a) and (266b). The approach of Gardent & Narayan (2015: 54) would instead account for (266a) by positing the chain of adjoining operations (i.e. a particular form of dependent derivation) shown in (266c). Similarly to what has been discussed for (263b) following Frank (2002: 291 en. 3), (266c) would however lead to incorrect semantic dependencies as soon as a raising predicate is adjoined to α.*s-arrive*.

(266) a. [At what time did Brockway say] Harrison arrived?

b.



c.



5.4.4 An alternative based on features and syncategorematicity

Since section 5.1.2, the derivation of the sentence *What do you think the senator should read?* has been the subject of repeated discussion. The structural descriptions proposed throughout

have constantly assumed that the verb *think* anchors an auxiliary tree, while the surrounding assumptions made for the representations in (155) and (156), discussed from page 218 on, have been reworked so as to come up with an arguably more adequate account, namely the *k*-delayed *h*-MC-TAG derivation in (237) proposed on page 309 at the end of section 5.4.1.

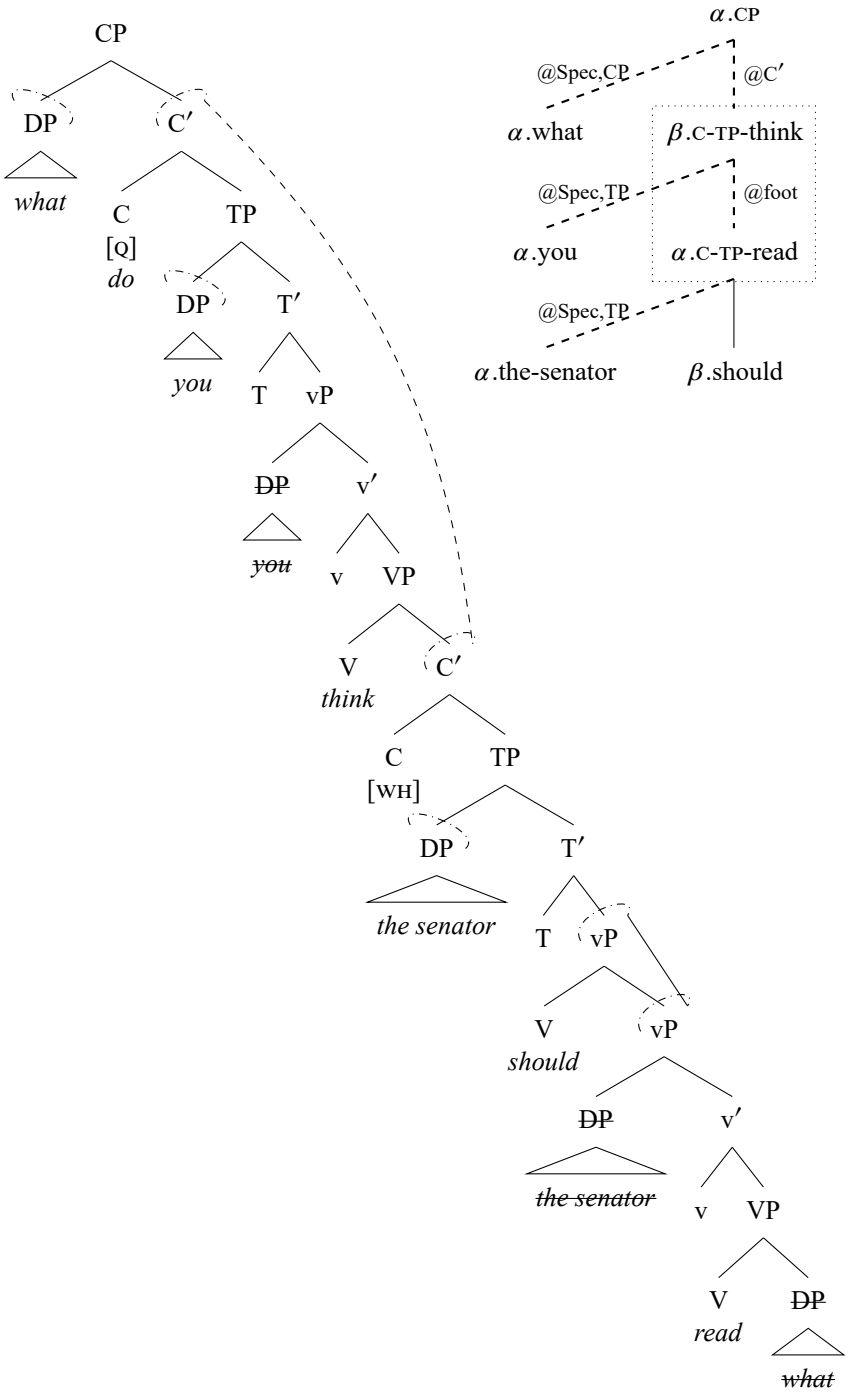
Still, there are some interrelated problems that have emerged since the first account in (155) and that have been dragged on for the sake of consistency with Frank (2002, 2004, 2006, 2010, 2013), Longenbaugh (2014), and Georgi (2019). They will be discussed shortly by relying on the representation on the next page in (267), which applies some changes on the derived tree first shown in (156) on page 218, so as to make it consistent with the derivation tree in (261), also reproposed in (267) for convenience.[178]

On the one hand, the lower T head is no longer lexicalized by *should* in (267), as the modal is included in a dedicated auxiliary tree as per the arguments in section 5.4.3. On the other hand, all the added arcs are placed at the root of the elementary trees involved in the derivation, in order to indicate their respective domains of locality. The only exception in this respect is represented by the foot of β*.should*, which is connected to the arc at the root of β*.should* by means of a solid edge. As opposed to the foot of β*.should*, the foot of β*.*C*-*TP*-think* coincides with the root of another elementary tree (i.e. α*.*C*-*TP*-read*), and the dashed line connecting the two C′ nodes highlights that this property follows from having β*.*C*-*TP*-think* involved in a chain of substitutions.

Starting from derived tree in (267), it becomes possible to highlight two problems of the TAG assumptions endorsed so far. The first *prima facie* seems to be a small one, as it is must be formulated in X-bar-theoretic terms. In (267), either of the two verbal extended projections must eventually fail to project up to CP. Although the derived tree in (267) suggests that it is the verb *read* that stops at at C′, the corresponding derivation tree reveals that it is the verb *think* that actually fails to both project and select for a CP node, so that the auxiliary tree it anchors be recursive at C′ as desired. Although this issue would not emerge under Bare Phrase Structure (Chomsky 1993 *et seq*.), an asymmetry of this sort among distinct verbal extended projections will emerge again later in the present section when raising predicates will be addressed once more.

[178] In the derived trees from (261) on, initial trees are taken to substitute in the position where they are overtly realized, as it is commonly argued in the TAG literature. Lower unpronounced copies are shown as per the copy theory of movement (discussed a.o. in Abels 2015, Gärtner 2022 and references therein), but their actual presence in the structure depends on how one assumes that the derivations are ultimately carried out, as discussed in sections 5.1.5, 5.3.2 and 5.4.3.

(267)



The second issue is instead twofold. In English, matrix questions trigger a T-to-C raising while the embedded ones do not, and the derivation in (267) accounts for this asymmetry by arguing that the higher C head in (267) bears a [Q] feature triggering such a movement on the one hand, while the lower C head lacks it on the other hand.[179] However, assuming

[179] As pointed out by Georgi (2019: 82-87) and references mentioned there, the absence of a [WH] feature in the higher C head in (267), if necessary for independent theoretical reasons, can be fixed by means of

that the C head in the extended projection of *think* bears a [Q] feature accounts for the matrix T-to-C raising at the cost of jeopardizing the desired interpretation. A person asking some speakers *What do you think the senator should read?* takes for granted that they think that the senator should read something. On the contrary, (267) suggests that such a presupposition is questioned as well, as the interrogative force is first projected within the elementary tree anchored by *think*.

The absence of a [Q] feature in the lower C head in (267) becomes instead problematic if one tries to factor out β.*C-TP.think*. In such a case, there would not be the formal premises to account for the sentence *What should the senator read?*, where the modal raises past the subject. To correctly derive the latter matrix question, one must assume the availability of a distinct (set of) elementary tree(s) where the relevant C head bears both [WH] and [Q]. Moreover, the latter factor should arguably prevent the shoehorning of the C′ node by another clause, so as to prevent the generation of an ungrammatical sentence where T-to-C raising occurs in the embedded clause(s).

At this point, it is striking to observe that (267) raises a problem related to the factoring out of an auxiliary tree, even though the latter operation is supposed to be the hallmark of TAGs. Moreover, the just mentioned way of circumventing the problem multiplies the number of elementary trees in a stipulative and inelegant way, and runs afoul of a principle that has been adopted since page 279 in section 5.3.2, namely that any assumed multi-component must correspond to a legitimate structure in a singleton set. Such a principle does not hold for α.*CP* and α.*C-TP.read* in (267), as their union does not derive a well-formed matrix question in English.

The solution that will be proposed shortly objects to an implicit assumption that has been endorsed since section 5.2.1 concerning the structure of each non-wrapping auxiliary tree. Given that a revised version of CETM has been adopted so far following Frank, the lexical head anchoring an elementary tree has to be at the base of such a structure. As a consequence, the foot node in each non-wrapping predicative auxiliary tree is arguably in a sisterhood relation with the anchor by being either the leftmost or the rightmost node of the spine, as exemplified by the sisterhood between *think* and the C′ node in (267). This property will now be challenged, though without altering the (non-)wrapping property of the auxiliary trees (as

the inter-clausal agreement discussed by Frank (2006: Sec. 5). Thus, this kind of issue for (267) is not as compelling as those discussed in the main text, as the [WH] feature in the lower C head can be assumed to percolate to the C′ node immediately dominating it and then along the spines of the auxiliary trees shoehorning the latter, like β.*C-TP.think*.

will be discussed further in section 5.5.2), by formulating the following speculative principle:

(268) The syncategorematic exception: In each auxiliary tree β, the sister of the foot node can be a syncategorematic expression selected by the anchor of β.

In (268), the term *syncategorematic* is understood along the lines of Schmerling (2018: Appendix A) and references discussed there. Krivochen (2023b) extensively explores the notion (see above all Krivochen 2023b: Sec. 2.2, 5.1.1), and Krivochen (2023d: 321-322) provides the following summary (italics in the original):

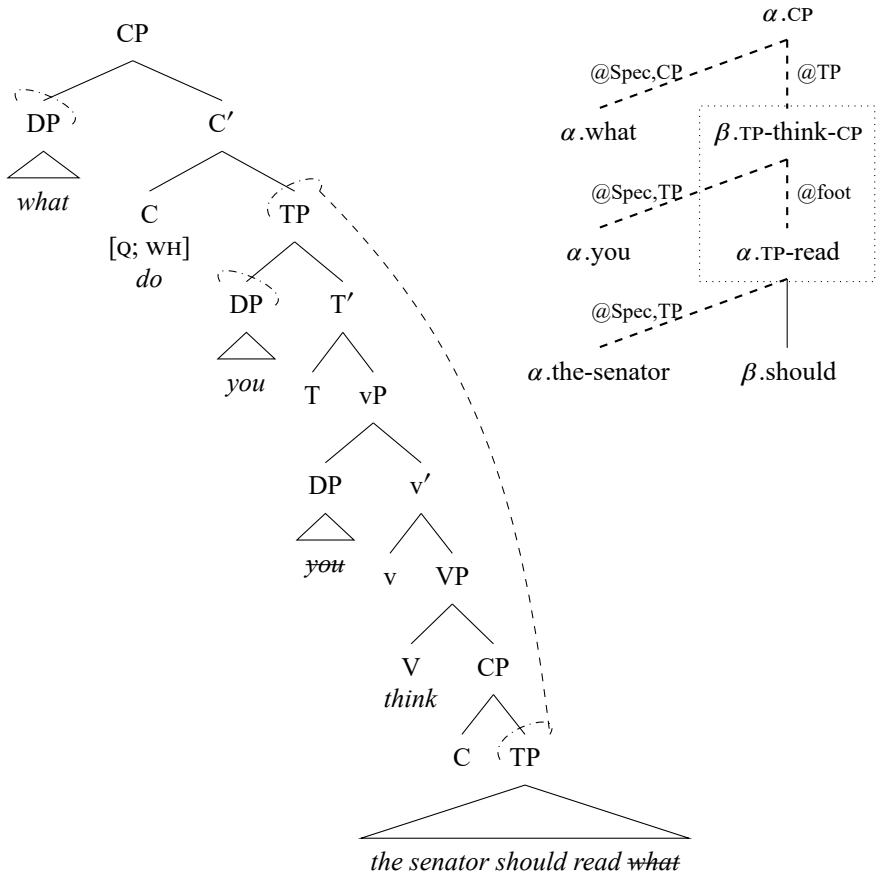
> Basic expressions can be either *categorematic* or *syncategorematic*: categorematic expressions have semantic interpretations which vary (or may vary) across models. Ns, Vs, As, Advs and contentful Ps are categorematic. Expressions like copulas, expletive pronouns, infinitival *to*, or case-marking prepositions [...], have *constant* interpretations across models. These expressions do not have meanings of their own, but every expression in which they occur does have a meaning: following the logical tradition [...], we will refer to them as *syncategorematic*.

Although Krivochen (2023b) discusses examples of syncategorematic elements in English and Spanish, a lot still remains to be done with respect the intra- and cross-linguistic variability of what counts as syncategorematic as per the quotation above.[180] Still, one can take advantage of (268) and rework the auxiliary tree anchored by *think* in (267) so as to make it recursive at TP, rather than at C′. This change is shown on the facing page in the derived tree in (269). The readers are also invited to observe that the derivation tree has been modified accordingly.

The derived tree in (269) improves the structural description in (267) in several welcome respects at the simple cost of complying with (268). This change can be applied once one follows Krivochen (2023b: Ch. 10) in arguing that the English complementizer *that* is syncategorematic. In (269), the lower C head in the derived tree no longer originates in the

[180] It is worth pointing out that the expressions that Krivochen (2023d: 321-322) encompasses as syncategorematic are a superset of those which Biberauer & Sheehan (2013: 40, 43) identify by means of the same term. In fact, the latter argue that only complementizers and auxiliary elements that are both head-final and FOFC-defying can be syntactegorematic. By following the empirical observations given by Dryer (2007) and previously mentioned at the end of section 2.3.2, Biberauer & Sheehan (2013: 40) apply syncategorematicity to elements that are deemed to be "acategorial" or "entirely lacking formal features", so that they "either do not project at all" or "project category-neutral structure" (Biberauer & Sheehan 2013: 43). However, the example of the English infinitival *to* illustrates that a projecting T head that bears no trouble to FOFC is nonetheless seen as syncategorematic by Krivochen (2023d: 321-322).

(269)



elementary tree anchored by the verb *think*. This means that it no longer needs to bear the features triggering some *wh*-movement or T-to-C raising. On the contrary, the higher C head in the derived tree bears both, and this would hold as expected even if the auxiliary tree *β.TP-think-CP* were not part of the derivation, as in such a scenario the sentence *What should the senator read?* would be derived as expected.

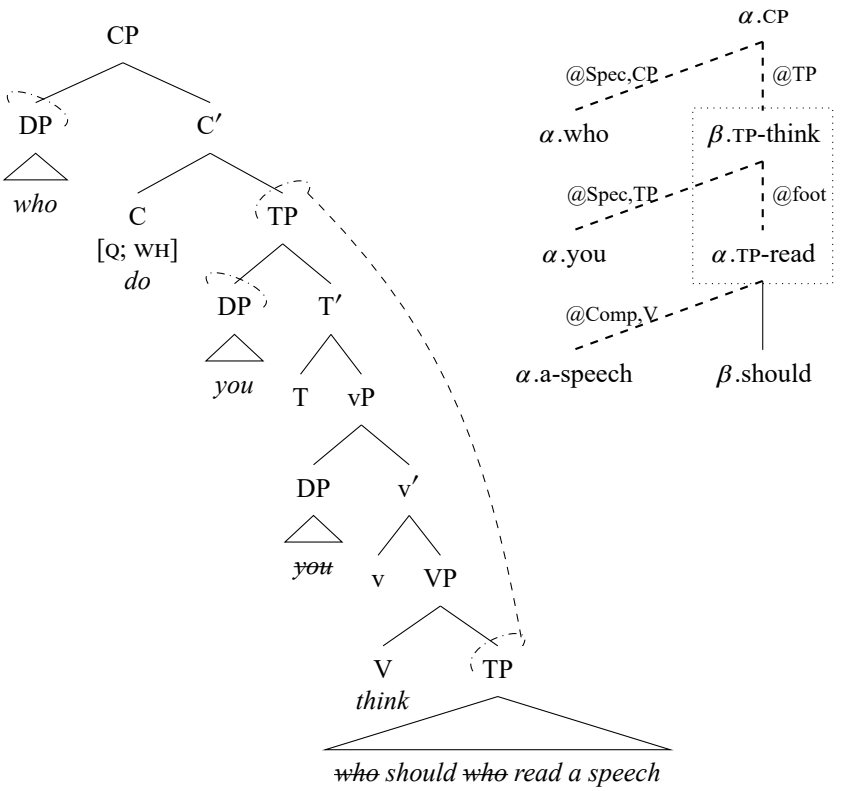
It is now important to stress that (268) states that a syncategorematic expression *can* be a sister of a foot node, and that the choice of such a modal in the definition is by no means accidental. In fact, a novel account for the *that-trace effect* (see a.o. Richards 2014: 4.6.1, Dikken 2018: Sec. 5, Hoot & Ebert 2021: Sec. 1 and references therein) can be proposed by hinging on such a detail. The basic case of this syntactic phenomenon can be represented by means of the following contrast, inspired by the examples given in Hoot & Ebert (2021: 1):

(270) a. What do you think (that) the senator should read?
b. Who do you think (*that) should read a speech?

If on the one hand (269) accounts for (270a), on the other hand (270b) can be derived by

having an auxiliary tree anchored by *think* that lacks the syncategorematic complementizer, as illustrated in (271) on the current page. By keeping in mind the important caveat made in footnote 178 on page 341, (269) shows the *wh*-movement of the object from its *in-situ* position to Spec,CP. In a similar fashion, (271) follows Moro (2000: 63 ff.) and related works in showing the raising of the *wh*-subject in the multi-component by means of an A-movement from Spec,*v*P to Spec,TP followed by an A′-movement to Spec,CP.

(271)



The readers who are familiar with cartographic syntax (Rizzi 1997, Rizzi & Cinque 2016, Rizzi & Bocci 2017) will surely point out that in both (269) and (271) there is no left periphery either in the elementary tree anchored by *think* or in the structure that the latter selects as its complement. Both are arguably problematic assumptions, which stem however from the need to make the auxiliary tree anchored by *think* not recursive at an internal node in the projection of C, so that the maximal projection of T emerges as the most plausible alternative in this respect. Put differently, the choice made in both (269) and (271) is inevitable within "an «impoverished syntax» approach assuming a single C head" (Rizzi 2013: 210).

One can rework (269) and (271) by resorting to the two left-peripheral heads that in cartographic works are deemed to lexicalize "*that*-like elements" crosslinguistically (Rizzi 2013:

210, italics in the original), i.e. ForceP and FinP, which are also the highest and the lowest heads in the split-CP proposed since Rizzi (1997), respectively. Given that FinP is usually deemed to mark the finiteness of a clause by which it is projected (Putnam 2020: 5) while ForceP is instead responsible for the lexicalization of the complementizer *that* in English (Rizzi & Bocci 2017: 4), one can legitimately formulate (272) as an adaptation of (269). In (272), it is FinP in the extended projection of *read* (and not TP) which is shoehorned by the auxiliary tree anchored by *think*.[181] The that-trace effect follows in (272) on the following page from the presence at the bottom of the latter auxiliary tree of a Force head, whereas its absence must be assumed for the corresponding counterpart of (271).[182]

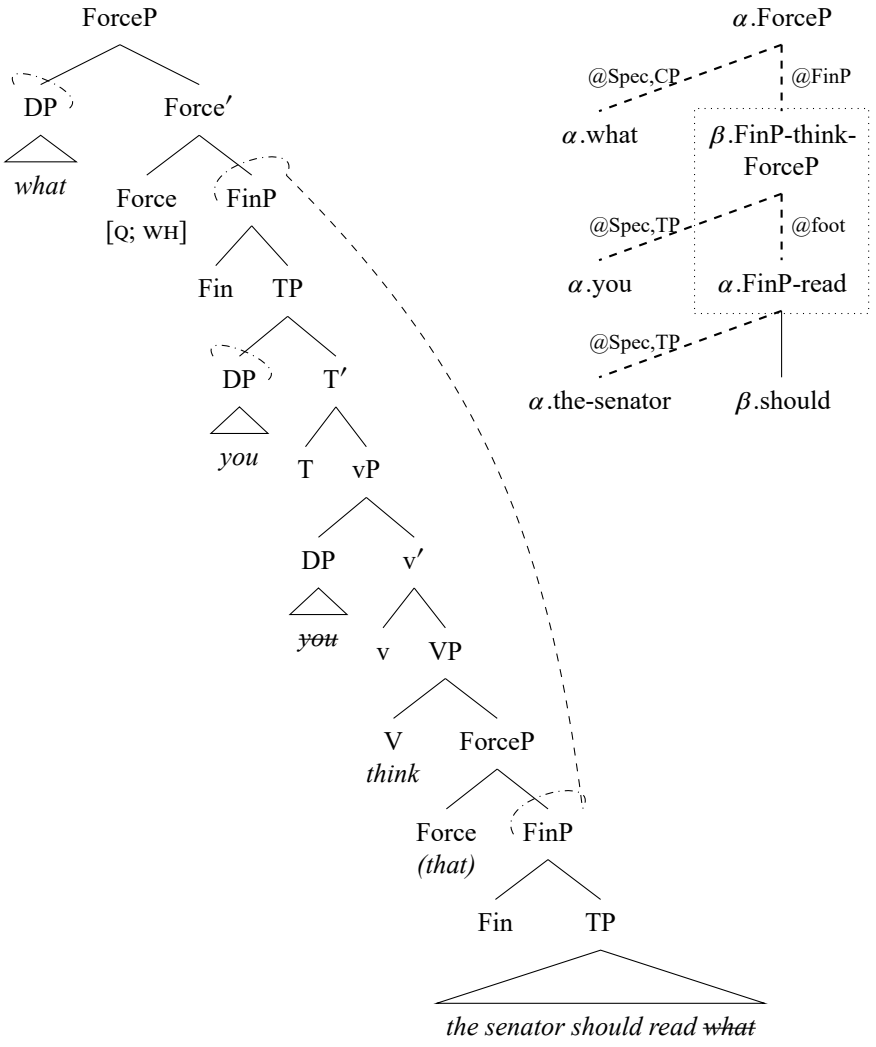
The derivations discussed in (269), (271) and (272) highlight something that the readers may have noted since the first 1-delayed 2-MC-TAG derivation proposed on page 309 (at the end of section 5.4.1) for the sentence *What do you think the senator should read?*, namely that the main clause anchored by *think* does not completely dominate the embedded clause anchored by *read*. Such a fact follows from two fundamental properties that the present approach assigns to what cartographic syntax (Rizzi & Cinque 2016, Rizzi & Bocci 2017) usually addresses as the left periphery of the main clause. On the one hand, it is always contained in the root of the TAG derivation tree. On the other hand, it is not necessarily part of the extended projection of the main clause, as it can rather originate as the periphery of an embedded clause like α*.FinP-read* in (272), from which it is then non-monotonically separated as discussed in section 5.3.2. Moreover, in (272) both α*.FinP-read* and β*.FinP-think-ForceP* are members of the delay that leads the former tree to α*.ForceP* (see again Chiang & Scheffler 2008: 19, previously quoted on page 297).

Derivations like (269), (271) and (272) further illustrate that, as long as the analysis is limited to the corresponding derived tree, one can mistakenly see the main left periphery as part

[181] The recursive nature of FinP assumed for (272) is further consistent with Putnam (2020: 7-8), who follows Ramchand & Svenonius (2014) in identifying Fin as the point of transition from the proximal zone for *situations* to the one for *propositions*. The connection between the Fin and the T head could also be investigated in the future with respect to the approach to clitic climbing adopted by Kayne (1989) and related works, where the extended projection of each bridge verb is argued to bear a left periphery of its own.

[182] A derivation like (272) is also in harmony with the core of the analysis that Douglas (2017: Sec. 2) proposes for the that-trace effect, although it dispenses with the anti-locality principle that he endorses. In fact, the latter is posited to ban some spec-to-spec movements and then explain why the corresponding successive-cyclic steps across phase edges are ruled out. A TAG approach like the one sketched for (272) does away with such a question by arguing that there is no successive cyclicity at all. Although the absence of successive cyclicity is in turn closely connected with the assumption of both [Q] and [WH] in Force (more on this in footnote 183 on page 349), this and other aspects of the analysis given by Douglas (2017) and related works need further exploration in the future.

(272)



of the extended projection of the verb *think*. This misunderstanding is further fostered by the semantic behavior discussed in section 5.4.3 for the auxiliary trees that are involved in a chain of susbstitutions. A raising predicate or a modal auxiliary that are control-flavored can get in the way between two other distinct elementary trees (i.e. a subject in Spec,TP and a lexical verb in the former case, a perfective auxiliary and a lexical verb in the latter) and interfere in a manner that obscures the semantic relation that would otherwise take place between them. An extension of such an analysis to (272) implies that the initial tree *α.FinP-read* eventually bears a loose semantic connection with the left periphery that roots the derivation tree because another verbal extended projection (anchored by *think*) interferes between the two.

As can be seen, a 2-delayed 2-MC-TAG account where FinP is a recursive node as in (272) seamlessly predicts truncated structures and asymmetries often known as main-clause phenomena (see a.o. Haegeman & Ürögdi 2010, Elordieta & Haddican 2018 and references

therein). It is also striking to observe that the approach being discussed here strikingly parallels the synthesis model of complementation proposed by Wurmbrand & Lohninger (2023: Sec. 3.2) in two relevant respects. First, the present defense of syncategorematic categories can be seen as an instance of the unique type of mismatch between syntax and semantics that is predicted by Wurmbrand & Lohninger (2023: 215), namely cases where syntactic structure has no consequence for interpretation.

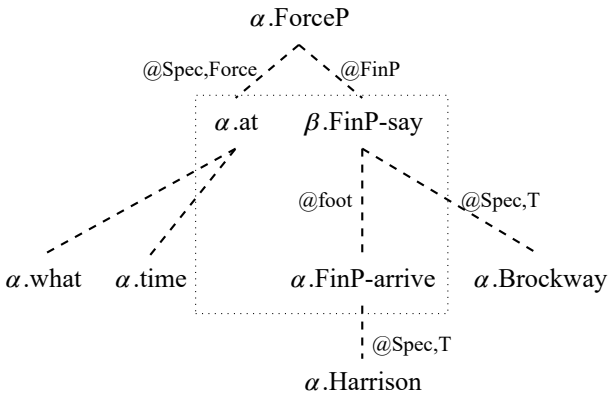
A second point on which the present TAG account converges with Wurmbrand & Lohninger (2023: 216) is "that in complementation the influence is mutual – a matrix verb can impose properties on the embedded clause, but properties of an embedded clause can also affect the matrix predicate". Despite this similarity between the two approaches, the TAG account proposed here stands out for proposing an entirely novel perspective on clausal complementation where it is the main clause, rather than the embedded one, that fails to project a complete extended projection. The latter point can be further witnessed by recasting the derivation trees proposed in (265a,b) and (266b) at the end of section 5.4.3, where the ambiguity of the sentence *At what time did Brockway say Harrison arrived?* was analyzed differently than Gardent & Narayan (2015: 54) did. In particular, it was observed that the substitution of the *wh*-phrase cannot be delayed to the left periphery at the root of the derivation tree, if the embedded clause also bears a left-peripheral head that can be endowed with the necessary [WH] feature.

The present discussion allows to tackle such an issue by adapting the major insights that Catasso (2024: Sec. 3) proposes for cases of truncation in German, so that the trees previously discussed at the end of section 5.4.3 can be revised as in (273) on the next page. In (273a), the truncated extended projection of a *verbum dicendi* fits in between another lexical verb and its left periphery, so that the desired reading follows from a delayed derivation. In (273c) it is the extended projection of *arrive* that is truncated instead. One can therefore argue that the C head belonging to α.*CP-TP-say* in (273b) "is in effect a composite force/finiteness head" (Radford 2004: 335) that is equivalent to the Force and Fin heads that in (273c) are part of α.*ForceP* and β.*FinP-say*, respectively. Therefore, the actual root of (273c) is arguably identical to the one in (273b), but (273c) is nonetheless accounted for by means of a tree-local delay to highlight the conceptual similarity between (273c) and (273a).[183]

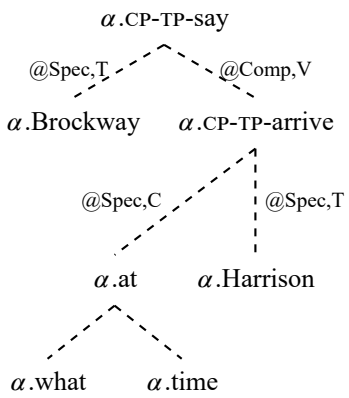
[183] As in section 5.4.3, the derivation trees in (273) accept uncritically the structure that Gardent & Narayan (2015: 54) assume for the *wh*-phrase *at what time*, as it is not a central concern. An arguably more important matter is that the Force head at the root of (272) and (273a,c) is deemed to bear (at least) two features, i.e. [Q] and [WH]. This assumption may appear problematic in a twofold respect. On the one hand,

(273)

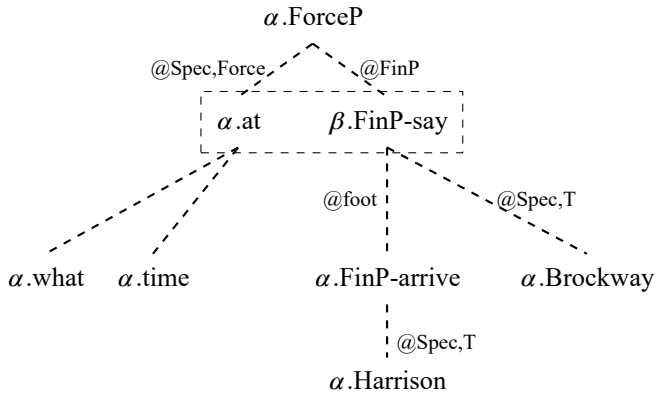
a. At what time did [Brockway say] Harrison arrived?



b. [Did Brockway say] at what time Harrison arrived?



c. [At what time did Brockway say] Harrison arrived?



The derivation tree in (273b) and those in (273a,c) illustrate that in the present analysis "the features of the C-domain may be located on a single (syncretic) syntactic head or scattered across a number of heads" (Douglas 2017: 4), which is in turn consistent with the *feature scattering* first proposed by Giorgi & Pianesi (1997) and more recently taken up a.o. by Hsu

---

it counters anti-lexicalist approaches like Putnam (2020: 4) and references therein, where "each functional head hosts and projects a singleton feature". However, section 5.1.5 showed that TAG approaches can be regarded as anti-lexicalist only in a very broad sense, because TAGs are inherently lexicalist. On the other hand, the presence in English of a [WH] feature in Force also goes against Rizzi (1997: 299) and subsequent works, where it is argued that in matrix questions the *wh*-phrase targets Spec,FocP, where Foc is a head between Force and Fin that hosts focalized constitutents. However, such a claim follows from successive-cyclic analyses of what happens in the left periphery of embedded clauses (see e.g. Radford 2004: Sec. 9.1-9.3), and phase-based approaches nonetheless need to assume that *wh*-phrases target either Spec,ForceP or Spec,FinP in the intermediate landing sites (see a.o. Douglas 2017: Sec. 4). In general, the need of other left-pheripheral heads besides Force and Fin in the TAG approach presented in the main text represents an issue that goes way beyond the scope of the present work.

(2017, 2021). However, this fact is yet another example of an aspect that some readers may have already noted in the derived trees from (269) on, namely that in the present account the classical head movement from V to T to C is not necessarily confined within a single elementary tree. On the one hand, in both (269) and (271) the T head originating in the auxiliary tree anchored by *think* combines with the C head in $\alpha.CP$ to ensure the correct inflectional features for the inserted *do* (see Radford 2004: Sec. 5.8, Gelderen 2017: 97, Diercks, Santorini & Kroch 2025: Sec. 12.4), while on the other hand the modal auxiliary *should* combines with the T head in $\alpha.TP$-*read* as explicitly shown in (267), as per Radford (2004: Sec. 5.6), Gelderen (2017: 23-25, 78-81), and Diercks, Santorini & Kroch (2025: Sec. 12.4).

Although both movements are in apparent contradiction with the extension of the Fundamental TAG Hypothesis assumed since page 270 following Frank, such an issue can easily be solved by slightly reworking the conditions under which Frank (2002: 161) claims that "inter-elementary-tree feature checking may take place". The latter (which was previously discussed from page 237 on, at the end of section 5.1.5) need not be confined to specifier-head configurations as argued by Frank (2002: 161) for subject-to-subject raising, but can arguably be extended to head-complement configurations. Such a move would not only justify all the debatable interactions among heads in the structures proposed from (269) on, but would also be consistent with common assumptions that have been made in the generative tradition since the X-bar schemata discussed in section 2.2.3 (see also Frank 2002: 46 ff.).

A further concern that the interactions among heads in (269) and (271) may raise is related to the point(s) in the derivation where they are supposed to take place. If the TAG literature (with the exception of the STAG approaches discussed at the end of section 5.2.1) is correct in linking the semantic interpretation of a sentence with its corresponding derivation tree, then interpretive effects are expected neither for (PF-oriented) rearrangements of heads taking place within a single extended projection nor for the inter-elementary-tree operations allowed by Frank (2002: 161) and Frank (2006: Sec. 5). Moreover, an adaptation of the thought experiment discussed in section 5.1.5 to the present discussion suggests that such rearrangements might not be a sufficient argument for a separated, post-syntactic component (which is a welcome result, at least according to Bruening 2022: 22). But of course, further research on this point is needed.[184]

[184] It would also be interesting to see whether and how the phonologically-oriented reflections in Frank (2010) and Richards (2016) can be implemented together with what is sketched in the present main text.

Another important asymmetry between the long movement of a *wh*-object in (269) and of a *wh*-subject in (271) is that, once the auxiliary tree anchored by *think* is factored out, the remainder of (271) must not require the insertion of *do*, contrary to what happens in a similar scenario for (269). In this respect, one can adopt the phonological mechanism defended by Longenbaugh (2014: 50-52) and references therein. In matrix questions where the *wh*-phrase in Spec,CP is the fronted subject, the absence of overtly realized material in Spec,TP makes the T-to-C movement unnecessary and thus "precluded by economy constraints" (Longenbaugh 2014: 51).

The appeal to a mechanism that Longenbaugh (2014) argues to be post-syntactic seems to be in contradiction with section 5.4.3, where many of the phenomena that he relegates to the post-syntactic component are instead accounted for by means of a 2-delayed 2-MC-TAG approach. However, this apparent inconsistency arises because of a false dichotomy that is (often implicitly) set up by discussions like those in Svenonius (2024: Sec. 3.2) and Longenbaugh (2014: Ch. 3), where *all* instances of head movement are taken to be either post-syntactic or not. The present 2-delayed 2-MC-TAG approach suggests that the problem can be faced in a different way. On the one hand, V-to-T-to-C movement as it is observable in English can be captured on the derived tree as discussed so far with respect to (269) and (271). On the other hand, the previous sections have shown that lexical heads anchoring their own elementary trees can be rearranged with one another by means of TAG operations.[185]

With respect to the operations typically allowed by TAGs, Joshi (1985, 1987) argues that they can capture the crossing dependencies in Swiss German discussed on page 218, and Kroch & Santorini (1991) and Krivochen & Padovan (2021: Sec. 3.2) extend this idea to the empirical data discussed by Haegeman & Riemsdijk (1986). Similarly, the previous sections have shown that reorderings among lexical heads by means of TAG operations can account for clitic climbing (discussed in section 5.3.3 following Chen-Main, Bleam & Joshi 2012) and also apply to the extended notion of derivation presented and reworked in section 5.2.1 (by hinging on Schabes & Shieber 1992, 1994). What matters with respect to such phenomena is that the semantically relevant dependencies be preserved on the derivation tree, and 2-delayed 2-MC-TAGs meet this desideratum by establishing agreeable dependencies where elementary trees are either tree-local on the tree they directly target or resort to legitimately

[185] As will be discussed shortly, Harizanov & Gribanova (2019: Sec. 2) and references therein highlight many cases where V moves to C without indepedent evidence of an intermediate step (i.e. V-to-T). The latter phenomenon is better captured by means of a TAG operation, rather than by the typical V-to-T-to-C movement.

formed delays so as to have a single common ancestor.

As a result, the 2-delayed 2-MC-TAG approach outlined here bears strong similarities with Harizanov & Gribanova (2019: 461), who argue what follows:

> [The notion of head movement] has been used to model two empirically distinct classes of phenomena. One class has to do with displacement of heads (fully formed morphological words) to higher syntactic positions, and includes phenomena like verb second and verb initiality. The other class has to do with the construction of complex morphological words and is involved in various types of word formation. Based on the very different clusters of properties associated with these two classes of phenomena, we argue that they each should be accounted for by distinct grammatical operations [...], rather than by the one traditional syntactic head movement operation. [... T]he operation responsible for upward displacement of heads [...] has the properties of syntactic phrasal movement, including the ability to affect word order, the potential to give rise to interpretive effects, and the locality associated with Internal Merge. On the other hand, word formation [...] is morphologically driven [...], it does not exhibit interpretive effects, and it has stricter locality conditions (the Head Movement Constraint [Travis 1984]).

The discussion in section 5.1.2 from page 219 on followed Frank (2013: Sec. 8.1) in acknowledging that the adjoining operation typical of TAGs resonates with the generalized transformations adopted in Chomsky (2002–1957) and later abandoned in Chomsky (1965). What has been said so far further strengthens the connection between TAGs and the model laid out by Chomsky (2002–1957).[186] In fact, the 2-delayed 2-MC-TAG approach that takes shape in the present work has the potential to agree with Chomsky (2002–1957: 114) in formulating three sets of rules, although with some necessary adaptations. As discussed in section 5.1.5, the first set handles the construction of elementary trees. In Chomsky (2002–1957), this was done by means of phrase-structure rules, while in Minimalism the latter are dispensed with because Merge takes care of structure building in their stead (as further discussed in section 5.3.2).

The exact collocation of Internal Merge in the architecture of a 2-delayed 2-MC-TAG

[186] Given that derivation trees built by 2-delayed 2-MC-TAGs seem to provide adequate semantic accounts, another aspect worth exploring in the future is the set of similarities that such trees share with the T-markers assumed in Chomsky (2002–1957, 1979). See Lasnik & Lohndal (2013: 32, 35-39) for further discussion.

needs further research in the future. In fact, IM should be part of the just mentioned first set of rules as long as one follows Chomsky (1995) and the schema discussed in section 5.1.5 based on Frank (2002: 24, 2006: 151, 2013: 241). However, in section 5.3.2 it was pointed out that Heck (2016, 2023) argues that IM is nothing but another consequence of the structure shrinking that is allowed by non-monotonic derivations. The latter take on IM would be further consistent with Chomsky (2002–1957: 114), who collocates simple transformations in the second set of rules.

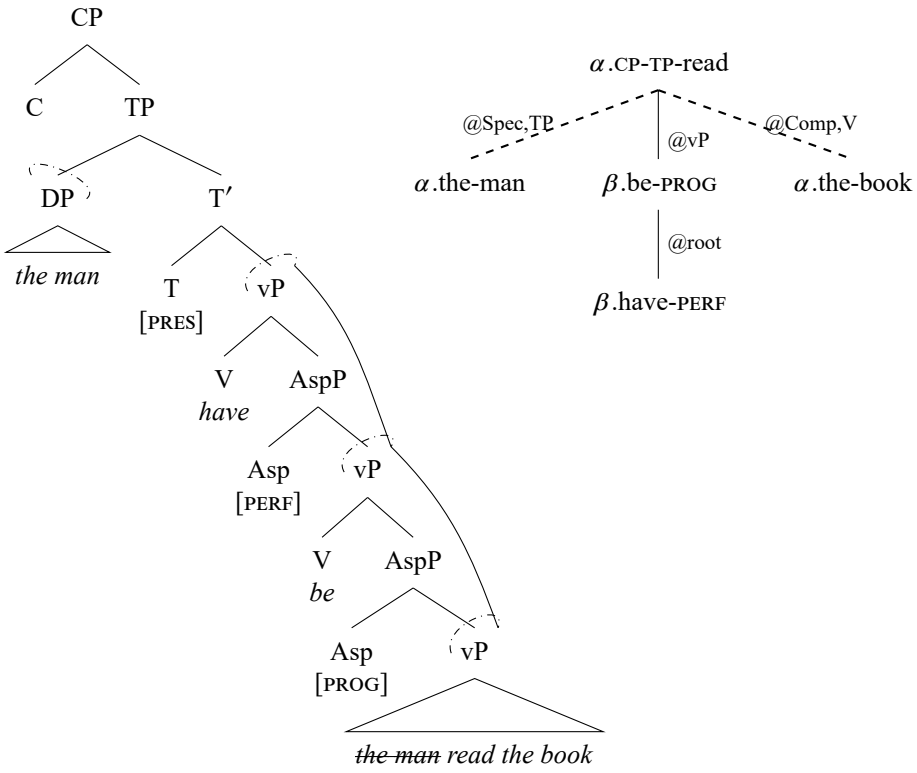
In Chomsky (2002–1957: 114), the second set of rules includes transformational rules more generally, so that generalized transformations also fit in there. Given that generalized transformations can be compared in turn to adjoining, one can argue that substitutions can be added to such a second set of rules by virtue of the similarities highlighted in section 5.1.5 and section 5.4.3 by following Frank (2002, 2006, 2013) and Kuhlmann & Satta (2012: 623), respectively.

Lastly, it is possible to observe a close similarity between the morphophonemic rules of Chomsky (2002–1957: 114) and the way the V-to-T-to-C movement has been accounted for in the present section. Such a resemblance becomes even more evident by combining what has been discussed in section 5.4.3 for the auxiliary chains of English together with on the syncategorematic exception in (268). The sentence *The man has been reading the book*, analyzed in Chomsky (2002–1957: 39-40), can be accounted for as shown on the facing page.

The derivation in (274) handles what is often addressed as *affix hopping* (as in the introduction by D.W. Lightfoot in Chomsky 2002–1957: ix, as well as in Radford 2004: Sec. 4.4) or *affix-hop* (Gelderen 2017: 122-123) by means of head-complement relations among elementary trees. One can argue that the aspectual features are realized by distinct syncategorematic heads by virtue of their frequently independent morphophonemic realizations in English (i.e. [PERF] as *-en* and [PROG] as *-ing*). Although this kind of account would need further discussion of its own, one can nonetheless observe that it already fares better than a TAG adaptation of Gelderen (2017: 123), where an analogous case of affix-hop is accounted for by "using uninterpretable features that check with the interpretable ones on a lower verb", as shown in (275) on page 356.

In a TAG perspective, the latter alternative represents another instance of derivation where the auxiliary trees cannot be factored out as desired, as the absence of β*.has* would leave the

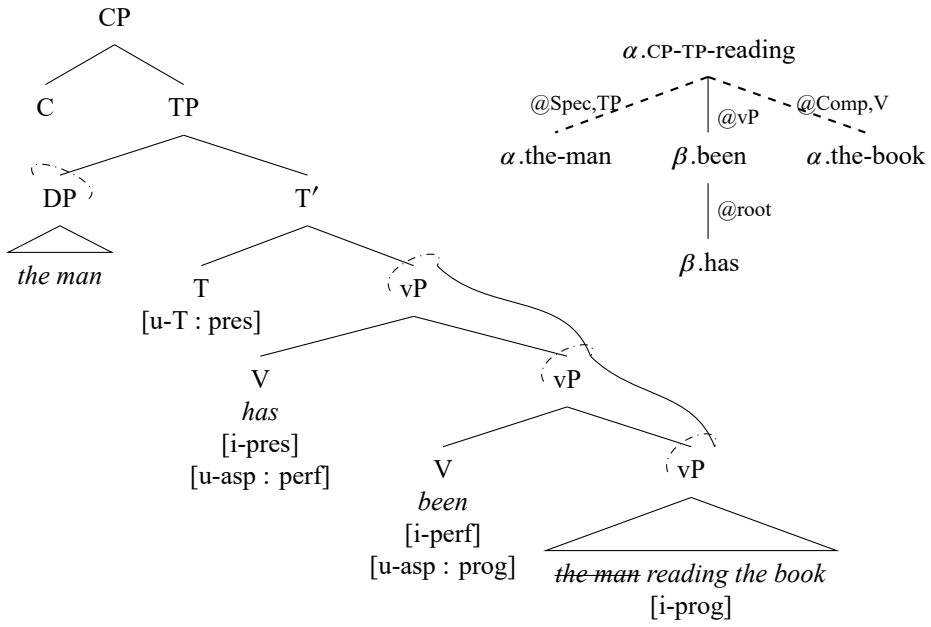
(274)



uninterpretable tense feature in T unchecked and unvalued. Such a problem actually descends from an issue that independently emerges in Gelderen (2017: Ch. 8), as in Gelderen (2017: 119) the T head is assumed to bear some interpretable tense feature when no auxiliary chain is present, *contra* Gelderen (2017: 123) and (275). One could defend the latter analysis by arguing once again that the presence/absence of interpretable tense features in T depends on the elementary tree (set) being derived, but this would lead to the questionable multiplication of elementary trees in the formalism that has already been criticized in this section.

Although (274) seems more promising than (275), it brings to the table two interrelated problems, which descend from the assumption that *β.have-PERF* and *β.be-PROG* are involved in the same chain of adjoining operations (i.e. a particular case of dependent derivation *à la* Schabes & Shieber 1992, 1994) and are both recursive at *v*P. As already discussed in section 5.2.1, the subsequent prediction is that one could legitimately swap *β.have-PERF* and *β.be-PROG* in the derivation (tree), so that both the semantic interpretation and the hierarchical relation between them are affected. However, such a swap in in (274) would derive the ungrammatical sentence **The man is having read the book* (where *read* is a participle). One first way to rule it out would consist in using indexes and applying the strong co-occurrence

(275)



constraint for predicative trees discussed on page 254.

One could alternatively posit that the ill-formed combination between *β.be-PROG* and *β.have-PERF* is ruled out by an independent principle that hinders the adjoining of a progressive auxiliary tree to a perfective one, e.g. the Auxiliary Selection Hierarchy that J. Baker (2019: Sec. 3.1) discusses following Sorace (2000).Although this would resonate with a similar assumption that was made for the L-TAG adaptation of Bruening (2018a) in section 5.2.2, holding that the perfective auxiliary *have* adjoins at *v*P would be still problematic in approaches that follow Chomsky (1995), where both passives and unaccusatives lack *v*P in their projection. If on the one hand the combination between *have* and a passive is correctly ruled out in favor of *be*, on the other hand the grammaticality of a sentence like *I have arrived* (to the detriment of **I am arrived*) would remain unaccounted for.

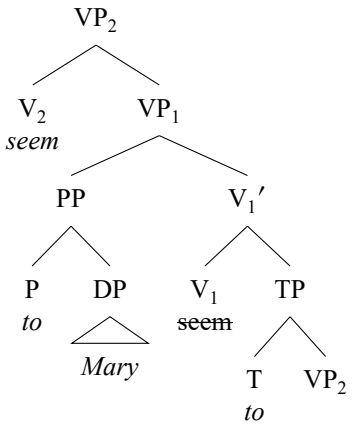
Before discussing a possible solution, it is worth observing that the same problem would also emerge when applying the syncategorematic exception in (268) to the elementary trees anchored by raising predicates. As discussed on page 249 (in section 5.2.1), two elementary trees anchored by distinct raising predicates can be swapped as in (276a,b), where the examples already discussed in (183a,b) are reported for convenience. If such auxiliary trees are taken to be recursive at T′ as in Frank (2002, 2004, 2006, 2010, 2013), the alternation between (276a) and (276b) must be accompanied by a stipulative exchange of the T heads in the two auxiliary trees, so as to always have the highest T head in the derived tree be the only

finite one. This additional refinement, done to avoid the generation of the ungrammatical sentence in (276c), highlights that raising predicates whose elementary trees are recursive at T′ are just as problematic as the verbal extended projections whose elementary trees are recursive at C′.

(276) a. John seems to Mary to be likely to be happy.
b. John is likely to seem to Mary to be happy.
c. *John to be likely seems to Mary to be happy.

A straightforward solution to the problem in (276c) can be found by resorting to (268) in combination with the claim on page 344 that Krivochen (2023d: 321-322) makes in favor of the syncategorematicity of infinitival *to*. From these premises, it becomes possible to formulate auxiliary trees anchored by raising predicates that are recursive at a node below T, as in (277).

(277)



As anticipated, the refinement in (277) raises however the same problem that has just been presented for (274) and (275), as it implies that the auxiliary tree anchored by *seem* should target a VP_2 label, whose correspondence with other typologies of verbs is not better specified by Storoshenko (2006). But even if it were, it would be likely to raise the same issue discussed for the auxiliary chains in English, as raising predicates do not impose restrictions on the verb classes they adjoin to. In fact, *John seems to be happy* is as legitimate as *Mary is likely to arrive* or *Luke appears to give a book to Mary*.

A solution to the problem raised by (274), (275) and (277) can be sketched by following two alternate approaches which end up pointing in the same direction. The first is proposed among others by Travis (2010: Sec. 4.4), who argues that the VP shells are always projected

by any type of verb (so that *v*P is always present in the structure), *contra* Chomsky (1995) and related works. The consequence of this assumption is that "there is a flexibility in the specification of [verbal] classes that is now determined more by features than by structure" Travis (2010: 121).

The second alternative makes features as pivotal as in Travis (2010: Sec. 4.4) without disavowing the verbal projections in Chomsky (1995). This is done by Koeneman & Zeijlstra (2014: 597-599), who argue that certain features act like a single *superfeature* when bundled together, so that under such a circumstance the (sub-)features trigger a certain syntactic phenomenon that would not take place otherwise.[187] Koeneman & Zeijlstra (2014: 599 fn. 20) give the central idea that is relevant for the present concerns:

> We take the assumption of superfeatures to be a more general strategy through which learners create functional categories, which can subsequently play a role in the clausal spine. Take for instance C, a syntactic position that can be filled by a verb or a complementizer. Whatever the precise nature of C is, it must be a feature that abstracts over verbs and complementizers.

Similarly to what has just been quoted, the specifier of CP (or ForceP) can be filled by a DP or a PP, as was shown respectively on pages 289 and 324 by focusing on the sentences *Which car does Sally wonder how to fix?* and *To whom does John seem to be happy?*. If the relevant features (such a [WH]), rather than the matching label, are taken to be the crucial factor that feeds any TAG operation, it becomes possible to circumvent the look-ahead problem that was pointed out (in sections 5.3.2 and 5.4.2) with respect to the *wh*-extractions just mentioned.

So far in this chapter, labels have been playing a crucial role in determining how and when a non-terminal node can be rewritten by a TAG operation. However, the superfeatures just introduced following Koeneman & Zeijlstra (2014: 597-599) can be used instead of labels to suit the same purpose. Such an alternate approach is in fact rather common in the TAG literature, as it dates back to Vijay-Shanker (1987) and is known as Feature-Based TAGs, henceforth FB-TAGs (see Gardent & Narayan 2015 and Frank 2002: Sec. 4.6 for formally and linguistically oriented discussions of this approach, respectively). Vijay-Shanker (1987) shows that if the features are bounded in size, "the resulting system has the same formal prop-

[187] Koeneman & Zeijlstra (2014) propose the notion of *superfeature* to make head movement be a consequence of rich agreement morphology, so that the former is bled whenever the desired bundle is not utterly formed. However, the motivations and the details of their proposals are not relevant in the present context, and can thus be set aside.

erties" of the TAGs described so far (Abeillé & Rambow 2000: 12), so that the conversion to be sketched shortly does not increase the generative capacity of the formalism.

The crucial premises of an FB-TAG alternative are summarized by Abeillé & Rambow (2000: 12) in the following way (italics in the original):

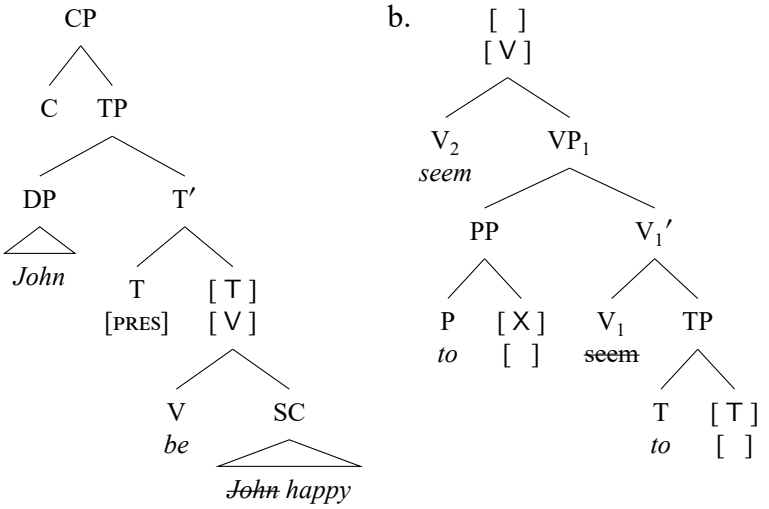
> (Vijay-Shanker 1987) proposes to annotate each [non-terminal] node [...] with feature structures, or more specifically, with two feature structures, called the *top feature structure* and *the bottom feature structure*. In linguistic applications, the top feature structure of a node describes (roughly speaking) the relation of the node to the tree above it, while the bottom feature structure describes the relation of the node to the subtree below it.

To illustrate this, one can resort to an elementary tree that has been used in section 5.2.1 as well as throughout section 5.4.2, namely a structure that would account for the sentence *John is happy*. It is reported in (278a) on the following page, with the VP node replaced by a pair of superfeatures/feature structures as per Abeillé & Rambow (2000: 12). The superfeature T stands for the set of features and selectional requirements that the T head brings to its sister node, whereas V indicates a similar counterpart that is projected by the verbal head. Crucially, the node which is a siter of the T head in (278a) is the only one that bears T and V as its top and bottom feature structures, respectively. Although the exact content of such sets of features goes beyond the scope of the present work, what matters is that the V superfeature be percolated by the verb independently of the class to which the latter belongs (e.g. unaccusative vs. di-transitive). In this manner, it becomes possible to have a raising predicate adjoined at the sister of the T head independently of the label that the shoehorned node would have in a TAG approach that is not Feature-Based (e.g. VP vs. *v*P).

If on the one hand the internal nodes are structured in the way just described for (278a), on the other hand the auxiliary tree in (278b) shows that the feature structures of the root and of the non-terminal leaves of an elementary tree work in a different way. In fact, since there is no structure above the root, the latter arguably has an empty top feature structure (which is equivalent to say that it is absent, as Findlay 2023: 2083 fn. 7 points out), while non-terminal leaves have an empty bottom feature structure due to an analogous reason.

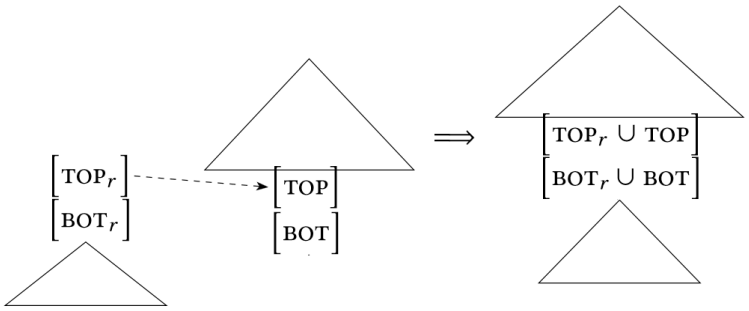
By following the original formulation in Vijay-Shanker (1987), substitution proceeds in such a way that the top features at the target node unify with the (empty) top features at the root of the targeting tree, and the same happens between the bottom features of the same

(278)



two nodes (see also Findlay 2023: 2083). This is shown on the current page by means of (279), which slightly modifies the illustration provided by Lichte (2015: 132) in the same respect.[188]

(279)

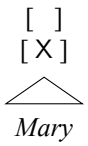


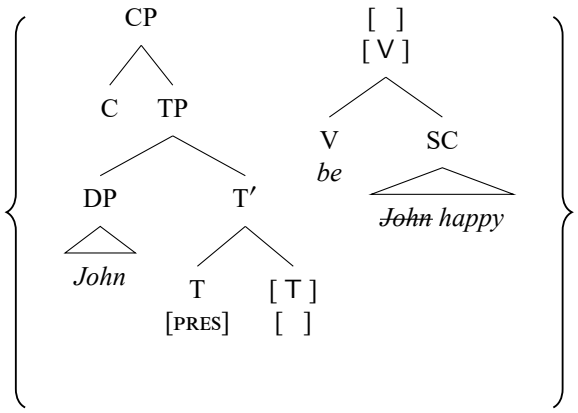
Before moving to the Feature-Based treatment of adjunction, it is convenient to point out that at the end of the derivation "the top and bottom features of a node" unify so as "to give a single description of the properties of that node" (Findlay 2023: 2082; see also Gardent & Narayan 2015: 49).[189] Such a procedure of unification further corroborates the

[188] Lichte (2015: 132) does not show the union of bottom features in the derived tree in (279). This is due to the fact that many FB-TAG approaches (e.g. XTAG Research Group 2001, Gardent & Narayan 2015: Sec. 4) assume a top feature structure for the root of initial trees, so that substitution is performed by unifying top features while ignoring bottom features (which are still deemed to be absent in non-terminal leaves, as acknowledged by Lichte 2015: 132 himself). Footnote 189 on this page will lay out an argument in favor of the substantial equivalence of the two approaches. However, the alternative presented here has the advantage of making the feature structures in the trees belonging to a non-monotonically derived multi-component consistent with the others elementary trees of the formalism, as will be discussed for (280).

[189] The fact pointed out in the main text can be put together with the discussion in footnote 188 on the current page to argue that the FB approach to substitution presented here is substantially equivalent to the one

inter-elementary tree operations previously elaborated in this section following Frank (2002: 161), especially if one further remarks that Frank (2002: 166) hinges on this very procedure to claim that "the features present in these sets may undergo checking and identification" (see also Kallmeyer & Scheffler 2004: Sec. 2). Moreover, both (278b) and (280a) concur in suggesting that the substitution of nominal arguments leads to the checking (by unification) of a set of nominal features X. This is another aspect that is relevant for the TAG analysis of U20 that will be discussed in section 5.5, as it essentially allows TAGs to remain agnostic with respect to a label-determined decision in favor of either the Lexical or the Functional Head Hypotheses of the nominal projection (Corver 2013). In fact, if any TAG can be converted into a Feature-Based counterpart where labels are a marginal aspect of the account, such a concern becomes no longer as central as in the works discussed in sections 2.2.3 and 2.2.4.

(280) a. 

b. 

The multi-component in (280b) shows that, whenever an elementary tree is non-monotonically split at a certain node, its top and bottom feature structures must also be severed accordingly, so that the top features are assigned to a leaf node in what Chen-Main & Joshi (2008: 11) call the «top» tree of the multi-component, while the bottom features remain at the root of the «bottom» tree. If this were not the case, such nodes would not bear the feature properties of other roots and non-terminal leaves, so that it would be necessary to posit a distinct mechanisms to have them take part in the chain of substitutions introduced in section 5.4.2.

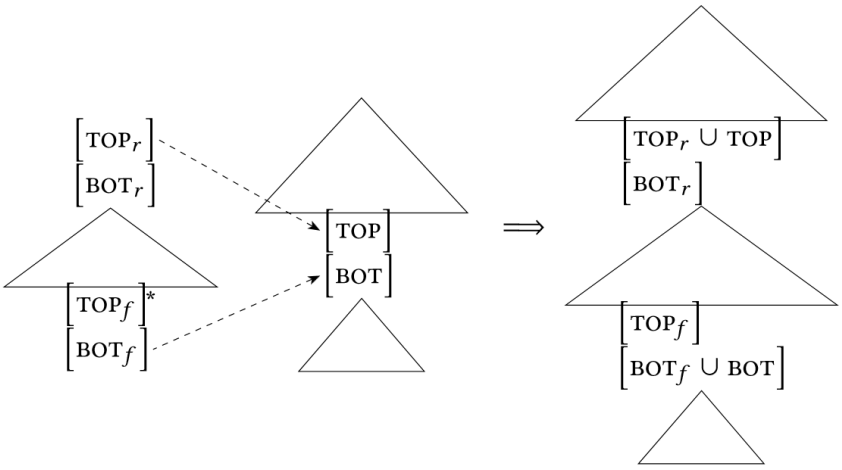
Adjoining in FB-TAGs proceeds instead in the way stated by Gardent & Narayan (2015:

assumed by XTAG Research Group (2001) and Gardent & Narayan (2015: Sec. 4). In such works, a non-terminal leaf bears a top feature T_0, while the initial tree α substituting there bears the top and bottom features T_1 and B_1. T_0 and T_1 are unified first, and at the end of the derivation they are unified with B_0, thus obtaining the following: $[T_0 \cup T_1 \cup B_1]$. The present approach claims instead that in an initial tree α' the top feature structure T_2 is an empty set. However, if its bottom features B_2 are such that $B_2 = [B_1 \cup T_1]$, the eventual unification with the top features in T_0 leads to the same result, i.e. $[T_0 \cup T_1 \cup B_1]$.

49), which is illustrated in (281) based on Lichte (2015: 133):

> The adjunction of an auxiliary tree β to a tree node η_o unifies the top and bottom feature structures of η_o with the top feature structure of the root node of β and the bottom feature structure of its foot node, respectively.

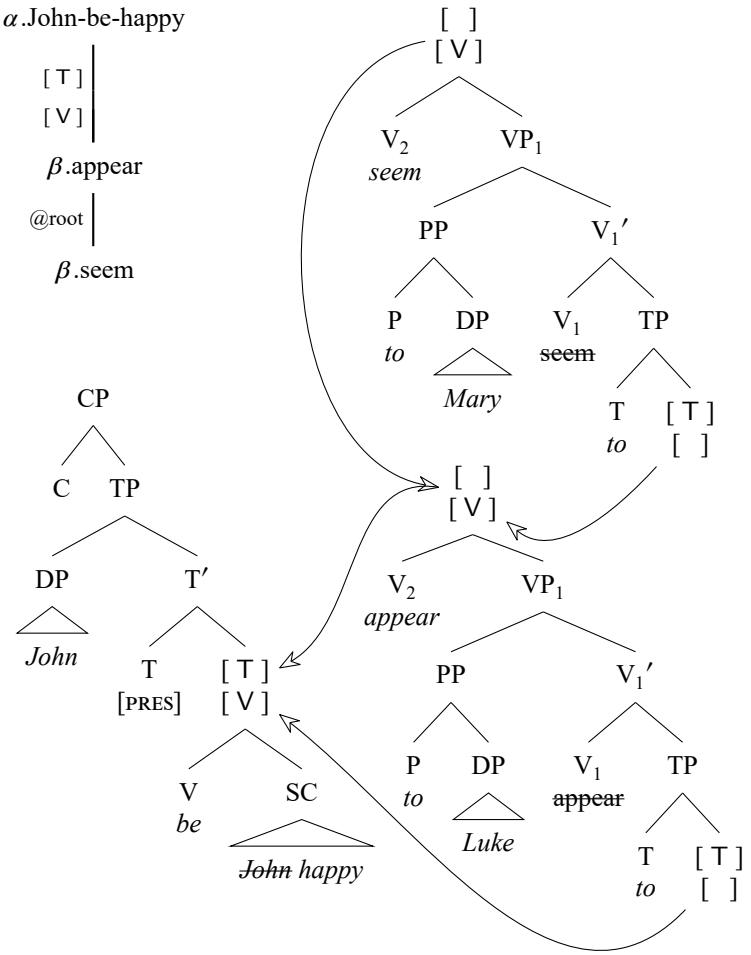
(281)



The chain of adjoining operations first presented in section 5.2.1 on page 248 to account for the sentence *John seems to Mary to appear to Luke to be happy* is proposed (in a simplified version) in (282) with a twofold purpose. On the one hand, it represents a concrete example of what is schematized in (281), while on the other hand it helps providing a definition of foot nodes in an FB-TAG approach. Given that in TAGs using labels the foot node is defined as the non-terminal leaf that has the same label of the root of the elementary tree it belongs to, FB-TAGs must recast this requirement on a different ground. The key property that (282) makes emerge is that the foot node of an auxiliary tree β is a non-terminal leaf that must target a node (of another elementary tree γ) whose bottom features coincide (in the nature, not necessarily in the actual values) with the bottom features in the root of β. In (282), the foot nodes of both β*.appear* and β*.seem* meet the requirement just stated by targeting an internal node that bears a bottom feature structure V, which is also the same that is present at the roots of both auxiliary trees.

The adjoining of β*.seem* to β*.appear* further highlights that, although the foot node is defined based on the properties of the root, the converse does not hold. Although one could argue that the root of β*.appear* targets a node whose top feature structure coincides with the one in the foot of the same auxiliary tree, the same line of reasoning cannot be extended to

(282)



the root of *β.seem*. In fact, the latter node targets the root of *β.appear*, whose top feature structure is not T as in the foot of *β.seem*, but is rather as empty as the root of *β.seem* itself. Such an asymmetry in the relation between the roof and foot nodes of the same auxiliary trees also manages to prevent the latter from adjoining at the foot node of another auxiliary tree, as it is often assumed in the TAG literature. In fact, if *β.appear* were adjoined at the root of *β.seem*, the bottom feature structure in the foot of *β.appear* would remain empty, *contra* the desideratum stated above.

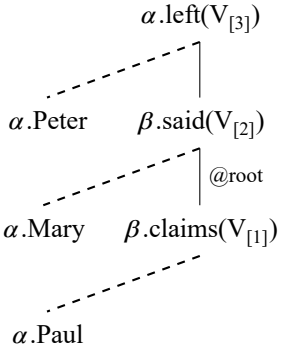
5.4.5 Some final remarks on the clausal domain

The discussion so far has set the premises to eventually frame the main topic of this dissertation, i.e. Greenberg's Universal 20, in a TAG approach that does not suffer from some well-known problems of L-TAGs, namely that in certain context they undergenerate or pro-

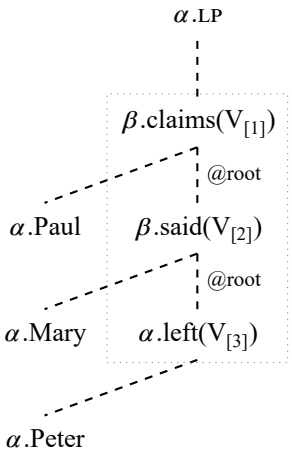
vide unsatisfactory derivation trees (or both). Given that chapter 4 has set the empirical basis for some novel parallels between the nominal and the clausal domains, one can argue that a good formal account of U20 should not discard the insights that can be taken from the analysis of clause structures. And before getting to section 5.5, it is in fact worth making a couple of final remarks on the latter, as they will eventually be relevant for the discussion in section 5.5.2.

On page 256 (in section 5.2.1) an analysis for a sentence discussed by Candito & Kahane (1998: 23) was provided, i.e. *Paul claims Mary said Peter left*. Such an account followed the referenced works (a.o. Joshi, Becker & Rambow 2000) by formulating a dependent derivation among clauses where the most embedded one (i.e. *Peter left*) corresponds to the root of the related derivation tree, reported for convenience once again in (283a) on this page.

(283) a.



b.



However, the subsequent discussion in sections 5.3 and 5.4 has demonstrated that a better analysis of such a sentence is (283b). In fact, in such an alternative the dependencies among clauses are correctly reversed as a consequence of a chain of substitutions (first presented in section 5.4.2). As was argued in section 5.4.4, the root of the recast derivation tree corresponds to the left periphery of the main clause, indicated as $\alpha.LP$ (as it can be a CP or a ForceP, depending on one's preferences). Given that Catasso (2024) argues that *verba dicendi* can foster truncated structures, it can be argued that both $\beta.claims$ and $\beta.said$ are incomplete extended projections and that $\alpha.LP$ originally belongs to the multi-component set of trees that also contains $\alpha.left$. As a result, the delay that the derivation in (283b) requires ends up including all elementary trees but the nominal arguments and the root $\alpha.LP$.

It is pretty straghtforward to fomulate a derived tree from either derivation tree in (283) because all the auxiliary trees involved are non-wrapping predicative trees of the left type (as first mentioned in section 5.2.1 following Schabes & Waters 1995). This implies that the derivation is tail-recursive because the foot node is always the rightmost node in the spine. However, works like Joshi, Becker & Rambow (2000), Chen-Main & Joshi (2008), and Lichte (2015) focus on clausal embedding in German, where the situation becomes more complex for a twofold reason. On the one hand, the sentence adapted from Kroch & Santorini (1991: 271) in (284a) shows that the clauses are not tail-recursive. On the other hand, the alternative adapted from S. Müller (2023b: 440) in (284b) shows that the order among the arguments selected by each verb can vary by having some of them undergo long-distance displacement (see also Abels 2015: 1408 ff.).

(284) Tamara sagt, dass... ('Tamara says that...')

a. ... [$_1$ *niemand* [$_2$ *den* *Bademeister* [$_3$ *den*
... nobody.NOM.SG DET.ACC.M.SG pool.attendant.M.SG DET.ACC.M
Mann *schwimmen* $_3$] *lassen* $_2$] *sah* $_1$]
man.M.SG swim.INF let.INF saw.3SG

b. ... *den* *Mann* *den* *Bademeister* [$_1$*niemand*
... DET.ACC.M.SG man DET.ACC.M.SG pool.attendant.M.SG nobody.NOM.SG
[$_2$ ~~den Mann~~ [$_3$ ~~den Bademeister~~ *schwimmen* $_3$] *lassen* $_2$] *sah* $_1$]
swim.INF let.INF saw.3SG

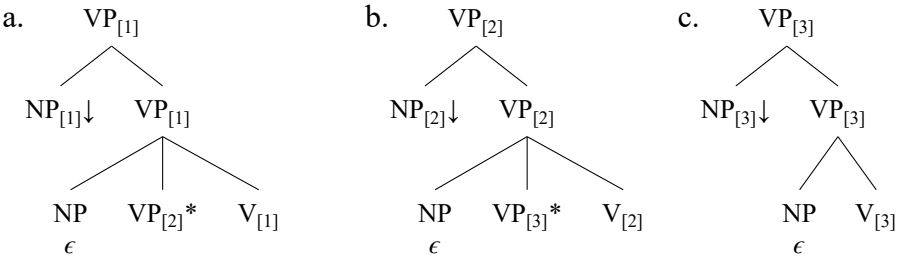
'... nobody saw the pool attendant let the man swim.'

Any TAG proposal that tries to capture (284) must find a compromise between two aspects that are in apparent contrast. On the one hand, each verb and the argument(s) it selects must belong to the same elementary tree, but on the other hand the latter must be enabled to move in such a manner that the resulting displacement appears to be non-local. What is more, the nature of the embedding shown in (284) leads to argue that in German (as opposed to what was seen for English) auxiliary trees anchored by a verb are of the wrapping type. This is due in turn to the fact that the foot node of each auxiliary tree that is anchored by a verb is arguably neither the leftmost nor the rightmost of the spine it belongs to.

A TAG work that abstracts over the data in (284) is Chen-Main & Joshi (2008: 12), where ternary-branching structures similar to those reported in (285) are proposed. In order to ac-

count for the alternatives in (284), one must assume that (285a), (285b) and (285c) provide the structures for *niemand sah*, *den Bademeister lassen* and *den Mann schwimmen*, respectively.

(285)



In (285), each tree follows the indexing rules described in section 5.2.1 while discussing the strong co-occurrence constraint for predicative auxiliary trees (on page 254). Furthermore, the nodes marked with an asterisk (e.g. $VP_{[2]}$*) are the foot nodes of the auxiliary trees, whereas the nodes indicated with a down arrow (e.g. $NP_{[1]}\downarrow$) indicate the nodes at which the arguments must be substituted. The terminal nodes lexicalized by a verb are instead indicated as $V_{[1]}$, $V_{[2]}$ and $V_{[3]}$ in (285a), (285,b) and (285c), respectively.

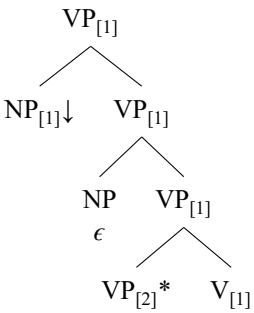
Each tree in (285) bears a ϵ symbol, which (as discussed at the end of section 5.3.2) is commonly used in formal language theory to indicate an empty terminal node. Chen-Main & Joshi (2008: 12) assume the latter to propose that the long displacements in (284b) are the consequence of arguments being moved to the left within the very elementary trees that include them. Given that in each tree in (285) the only argument does not raise over some other overt material during such a process, what Chen-Main & Joshi (2008: 12) propose is a case of raising that is as string-vacuous as those proposed by Kroch & Santorini (1991: Sec. 4.4) for the internal reordering of the verb cluster in West Germanic (see also Krivochen & Padovan 2021: Sec. 3.2). This idea permits to "reduce variation in the word order [...] to the node at which auxiliary trees are adjoined" (Kroch & Santorini 1991: 317), because another potential adjoining site (i.e. a VP node that is different from the VP root) is added for each elementary tree as a consequence of such a raising.

Both sentences in (284) can thus be accounted for by means of what Schabes & Waters (1995: 483) call *wrapping adjunction*, namely by having the wrapping trees like those in (285) combined by means of adjoining. In particular, (284a) can be generated by having (285a) adjoined at the root of (285b), which in turn adjoins at (285c). The alternative in (284b) can be obtained instead by combining the three trees in the same order, with the crucial difference that this time (285a) and (285b) target the non-root VP nodes in (285b) and (285c),

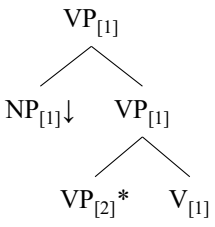
respectively.

Based on what has been said so far, the analysis provided by Chen-Main & Joshi (2008) suffers from two problems that go beyond the non-adoption of a C-T-v-V system discussed from section 5.4.1. On the one hand, it is yet another case of the inadequate dependent derivation in (283a). On the other hand, it forces the learners to postulate for each displaceable argument a link to an empty category (as per Kroch & Santorini 1991: 310) or a trace or a copy despite the "absence of positive string evidence" (Kroch & Santorini 1991: 323). In the case of the phenomenon which Chen-Main & Joshi (2008) focus on, the latter matter can be solved by means of binary branching counterparts to (285). Chen-Main & Joshi (2008: 14) provide such a conversion, but they recast a ternary-branching structure like (285a) as in (286a), which worsens the situation by providing a third redundant adjoining site. If instead one reformulates (285a) as in (286b), it becomes possible to retain the two desired adjoining sites with no need of the string-vacuous hypothesis, so that the analysis which Chen-Main & Joshi (2008) can easily be ameliorated at least in this respect.

(286) a.

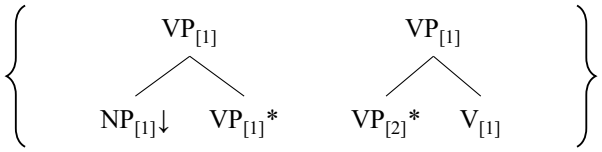


b.



Although the need of wrapping trees in TAGs has already been pointed out in section 5.2.1, what has just been discussed allows to observe another fact. Both Kroch & Santorini (1991) and Chen-Main & Joshi (2008) discuss the word order variation in German under the string-vacuous raising hypothesis, but the former do so to capture the reordering within the verb cluster (already mentioned in section 5.4.3), while the latter to account for the reordering among the related arguments. The fact that the two concomitant phenomena can be analyzed separately suggests that the wrapping trees involved can be prone to some non-monotonic severance that allows the lexical verb to recombine with the other verbs in the derived tree structure, or the nominal argument to be reordered with respect to the other arguments, or both. Leaving for future work the proper conversion into a C-T-v-V system (or a feature-based equivalent) where trees like those in (285) are combined by means of a chain of substitutions, the idea just sketched can be illustrated by proposing a multi-component set whose

members correspond to the structure in (286b), namely (287).

(287) 

Albeit sketchy, the representation in (287) sets the bases for two suggestions that will be somewhat relevant in the discussion of U20 in section 5.5. On the one hand, all wrapping auxiliary trees can be better analyzed by decomposing them into non-wrapping subparts (an idea to keep in mind in anticipation of section 5.5.2). On the other hand, the reordering among nominal arguments is a separate phenomenon from the reordering among verbal lexical heads because the trees involved in the former group are left auxiliary trees, whereas those involved in the latter are right auxiliary trees.

The consistent directionality among the trees to be recombined would not only be a decisive factor in the aforementioned genus-specific (i.e. West Germanic) facts, but can also be seen as some specific instance of the harmonic regularities proposed in many works discussed in chapters 2 and 3. Crucially, this TAG adaptation of the notion of Harmony refers to elementary structures that can correspond to heads as well as to more complex syntactic objects. Provided that the local properties of such structures are well-defined, the unidirectionality among the combinable spines (in the sense of Nesson 2009: 4, Kuhlmann & Satta 2012: 623) can thus be the crucial ingredient that underlies both the restrictions captured by FOFC with respect to heads (see sections 2.3.2 and 2.3.3) as well as those involving phrasal objects (for examples, see a.o. Feldcamp 2022: 2-3, 49, Cinque 2023: Sec. 2.2-2.7 and references therein). Although the discussion of U20 in section 5.5 represents an initial exploration in this direction, it is needless to say that further research on this point is needed.[190]

[190] Any inquiry that aims to explore further the potential relevance of the unidirectionality among non-wrapping auxiliary trees cannot help but consider a tree-rewriting formalism that has not been mentioned yet, namely Tree Insertion Grammars (TIGs). Any TIG "prohibits elementary wrapping auxiliary trees" as well as adjoining "on the roots of auxiliary trees" (Schabes & Waters 1995: 485), but even more importantly the adjoining of a left auxiliary tree is excluded on an internal non-root node in the spine of another right auxiliary tree; ditto for the right counterparts, which cannot adjoin to a left tree (Schabes & Waters 1995: 481). Due to these severe constraints, Schabes & Waters (1995: 493 ff.) show that TIGs are restricted instances of TAGs that are weakly equivalent to CFGs and strongly lexicalize them. However, TIGs are not closed under strong lexicalization (Kuhlmann & Satta 2012: Sec. 5) and, as soon as the multiple adjoining of both left and right trees is allowed at a certain node as per Schabes & Shieber (1992, 1994), the system becomes equivalent to an L-TAG (Joshi 2003: 254 fn. 4). More generally, Boullier & Sagot (2011) show that any MC-TAG with arity *h* can be converted into an equivalent *2h*-MC-TIG, although the class of 2h-MC-TIGs is overall more powerful than that of *h*-MC-TAGs.

5.5 A consistent (MC-)TAG account of U20

The problems raised by the two alternate TAG analyses in section 5.2 are crucially connected to the fact that both are still set on L-TAG terms. On the one hand, section 5.2.2 has shown the limits emerging from an L-TAG account that allows extended derivations *à la* Schabes & Shieber (1992, 1994). On the other hand, section 5.2.3 has shown those linked to an adaptation of Abels & Neeleman (2009, 2012) that hinges on a notion that is central to the analyses of U20 in Cinque (2005, 2023) and related works, i.e. pied-piping.

The discussion made so far in this chapter has set the premises to finally show how MC-TAGs can capture Greenberg's U20 in an effective way. In particular, section 5.5.1 will illustrate that the 2-delayed 2-MC-TAG approach supported in the previous sections has the twofold merit of providing an analysis of U20 that is both consistent and capable of capturing the facts discussed in chapter 4. Section 5.5.2 will conclude the chapter by shortly exploring how TAGs allow to reframe some problems left open in section 3.1, even though much must be left for future work.

5.5.1 A 2-delayed 2-MC-TAG adaptation of pied-piping for U20

The 2-delayed 2-MC-TAG analysis to be discussed starts from the same conditions on pied-piping first provided in section 5.2.3, and reproposed below in (288) for convenience.

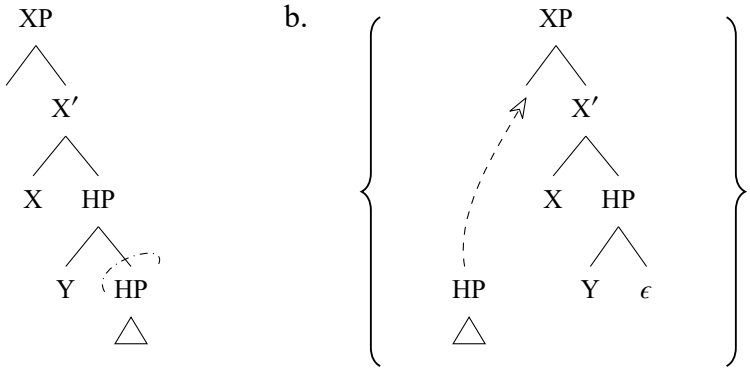
(288) Pied-piping applies under the following conditions:

a. A projection XP and a projection HP are first-merged so that the former dominates the latter.
b. The left specifier of X is free and can host HP.
c. The structural portion that intervenes between XP and HP must contain some overt material Y that c-commands the first-merged position of HP without being involved in the pied-piping of HP.

For the purposes of the present section, the pied-piping of a projection HP can be seen as consisting of the non-monotonic severance of HP from the structure that also includes Y and XP, as shown in (289a) on the next page. A multi-component is then formed, and HP substitutes at Spec,XP by means of a 1-delayed 2-MC-TAG derivation, as in (289b). The representation in (289) shows the case where Y is a head and, although the head-initiality of

both X and Y is suggested, nothing prevents them to be head-final without further stipulations (e.g. FOFC, discussed in sections 2.3.2 and 2.3.3). The representation in (289) further takes Y as a non-projecting head immediately dominated by a node labeled HP. However, both the non-projecting nature of Y and the label of the node immediately dominating it come without a strong commitment, as the use of a HP label simply stands for the possibility that some subset of features [H] be percolated from the lower HP projection (which is pied-piped) to the higher HP node (which is not pied-piped), as per the discussion in section 5.4.4.

(289) a. b.



In the present context, what truly matters is that either the HP label or the subset of features [H] contained in both nodes be what allows other auxiliary trees to adjoin to either of them. Hence, a TAG account of U20 can start by replacing each HP occurrence in (289) as NP (or alternatively by making the [N] features an instantiation of [H]). The exact nature of the X and Y heads, on the other hand, can be left unspecified. In fact, the pied-piping of NP is never string-vacuous and always compliant with anti-locality as soon as there are two distinct morphophonological realizations in the nominal extended projection that do not undergo pied-piping (with the proviso that X c-commands Y).

Thus, X and Y can be instantiated by heads like an inherent case marker K and a determiner D, or a D and a classifier Clf, or any other conceivable combination of any two empirically-grounded heads proposed in the literature. This is generally consistent with a premise mentioned in section 5.1, namely that the crosslinguistic variation descends from differences in the elementary trees (Frank 2013: 238).

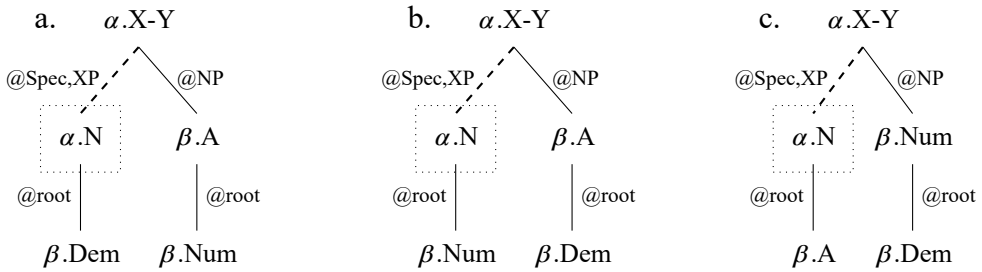
Given that there are two HP projection in (289), the modifiers of U20 can be adjoined in two different positions. It will be further assumed that all modifiers anchor non-wrapping predicative auxiliary trees (though see section 5.5.2 for discussion of the possible wrapping

counterparts), and that their combination is always constrained by the independent principle in (290), which was previously proposed on page 257. In the L-TAG adaptation of Bruening (2018a) discussed in section 5.2.2, (290) was adopted to regulate the combination among left auxiliary trees. In the present context, it is assumed to apply instead to any kind of auxiliary tree.

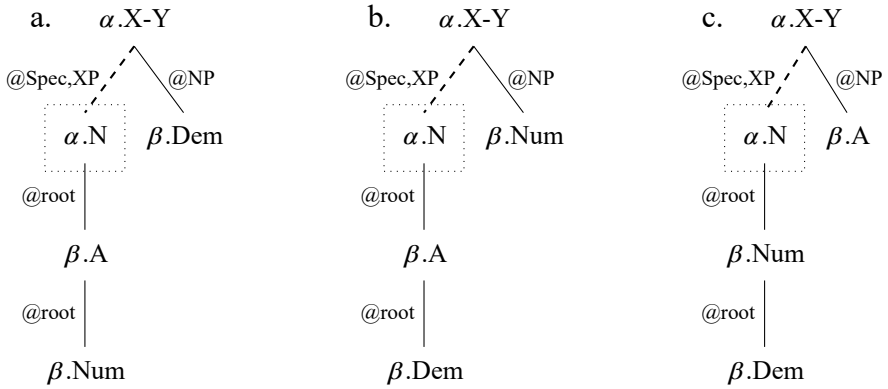
(290) *Dem* > *Num* > *A* > *N*

At this point, a notion that was introduced on page 259 becomes important once more, namely that of derivation tree families. In fact, both (289) and (290) determine those reported in (291), (292) and (293). In contrast with the L-TAG adaptation of Bruening (2018a) shown in section 5.2.2, the auxiliary trees involved in such derivations do not have to comply with a leftwardness condition. Thus, all the derivation tree families in (291) and (292) derive four permuations each, just like (293a) does. All the eight orders that Abels & Neeleman (2009, 2012) argue to be base-generated can be derived from (293b), and the same holds for (293c). The latter further stands out for being the only derivation tree family where no pied-piping occur.

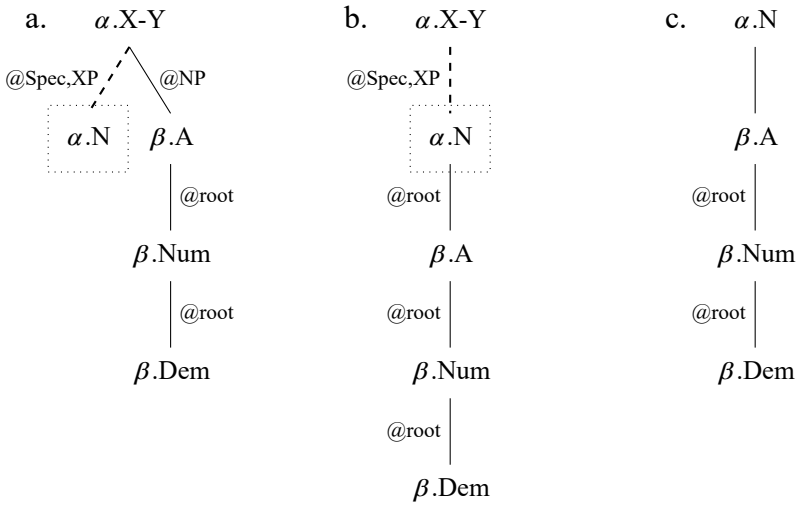
(291)



(292)

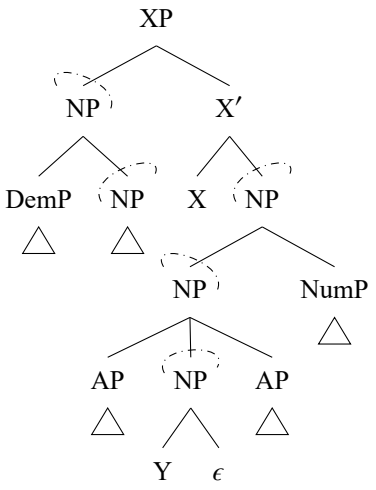


(293)



In order to discuss what the derivation tree families just shown can account for, it is first necessary to introduce the notation that will be adopted from now on to illustrate the related linear orders. The rationale behind it can be understood by looking at how the Dem-N-A-Num order can be mapped from the derived tree in (294), which represents in turn a possibility that is allowed by the dervation-tree-family in (291a).

(294)



The derived tree in (294) on this page shows that, as soon as the Noun is extracted from the position in which it is first-merged, two partitions can be established.[191] The former partition is formed by whatever fills Spec,XP, namely the Noun and all the modifiers that adjoin at is

[191] The idea that there are two partitions in (294) can be seen as a novel application of the notions of *prefield* and *postfield*, first proposed by Rijkhoff (2002: 251-252) and already mentioned at the end of section 4.4. However, in Rijkhoff these two terms were defined by using N as a reference point, so that prenominal and postnominal elements were part of the prefield and the postfield, respectively. In (294), the reference point is instead represented by the X head, so that N is also included in one of the two partitions.

root, e.g. Dem in (294). The latter partition is formed instead by the modifiers that adjoin at the node dominating Y, such as NumP and AP in (294). The ternary branching in (294) further shows that the structurally lower modifier that adjoins in the second partition (e.g. AP) can indifferently be a right or a left auxiliary tree, as its linear position with respect to the other elements of U20 would not be affected by either option. All these things can be captured for (294) as follows:

(295) Dem^{L} N^{PP} | $A^{R/L}$ Num^{R}

In (295), a vertical bar divides the two aforementioned partitions, whereas each modifier is superscripted with either *L* or *R* depending on whether it must be a left or a right modifier auxiliary tree, respectively. If the directionality of the spine of an auxiliary tree is irrelevant, as it is for AP in (294), the chosen superscript is *R/L*. Lastly, *PP* (for *pied-piping*) and *B* (for *base*) are the superscripts for N, depending on whether it is or it is not severed from its first-merged position, respectively. Whenever N does not raise (thus being a N^{B}), the vertical bar is not reported, as a XP hosting the leftmost partition can, though need not, be present in the structure (ditto for Y).

The attention can now be directed towards (291b) and (292c), as they derive the two orders that Bruening (2018a) expects *contra* Cinque (2005, 2023) and related works, i.e. Num-N-Dem-A and N-Num-Dem-A, as shown below:

(296) Orders that can be mapped from (291b):

a. Num^{L} N^{PP} | $A^{R/L}$ Dem^{R}
b. N^{PP} Num^{R} | $A^{R/L}$ Dem^{R}
c. Num^{L} N^{PP} | Dem^{L} $A^{R/L}$ [ruled out in Cinque (2005) *et seq.*]
d. N^{PP} Num^{R} | Dem^{L} $A^{R/L}$ [ruled out in Cinque (2005) *et seq.*]

(297) Orders that can be mapped from (292c):

a. Dem^{L} N^{PP} Num^{R} | $A^{R/L}$
b. Dem^{L} Num^{L} N^{PP} | $A^{R/L}$
c. Num^{L} N^{P} Dem^{R} | $A^{R/L}$ [ruled out in Cinque (2005) *et seq.*]
d. N^{P} Num^{R} Dem^{R} | $A^{R/L}$ [ruled out in Cinque (2005) *et seq.*]

Without any additional assumption, the derivation trees shown in (291)-(293) account for the same sixteen orders that can be generated by the approach described in section 5.2.2. However, one can effectively reduce the number of allowed permutations to the fourteen discussed by Cinque (2005, 2023) and related works. What one needs is a ban on the property that unites (291b) with (292c), so that the orders in both (296) and (297) are ruled out:

(298) a. Num cannot adjoin to the root of N^{PP} directly without the intervention of A. Put differently,

b. Num cannot be pied-piped by N without A.

The alternate formulation in (298b) helps to understand that the exclusion dictated by (298) boils down to something that is not unheard of. In fact, it is a staple in Cinque (2005, 2023) and Roberts (2017b), namely in any approach that accounts for U20 by means of successive-cyclic raisings of N within a structure that is constrained by antisymmetry (Kayne 1994).

Although some readers might hinge on the works just mentioned to object that a principle akin to (298) should also apply to Dem, it must be observed that Cinque (2005, 2023) has to assume an extension of (298) to Dem because he cannot do without it (though see Steddy & Samek-Lodovici 2011 for an alternate account of Cinque 2005 that diverges on this point). In fact, the assumed successive-cyclic raisings of N leave no alternative in this respect, and other works try to fix the wrong predictions that this stipulation makes by positing a low position for Dem (see Roberts 2017b) or for Num (see Ouwayda & Shlonsky 2017).

In addition to that, it is worth pointing out another relevant observation that Shlonsky & Bocci (2019: Sec. 5) make emerge. By focusing on the crucial factors by means of which the ten unattested orders in Cinque (2005) are ruled out, Shlonsky & Bocci (2019: Sec. 5) observe that Num-N-Dem-A and N-Num-Dem-A are the only two where the offending violation does not consist of the independent raising of some modifier(s), but rather of the movement of a non-constituent containing N, i.e. [Num N] or [N Num].

Since TAGs dispense with successive-cyclicity, it is up to the linguist to decide whether it is appropriate to assert that a ban like (298) also applies to Dem. As it turns out, (298) suffices not only due to the theoretical reasons just mentioned, but also because nothing but each of the fourteen attested orders of U20 is accounted for once (291b) and (292c) are excluded from the families of derivation trees in (291)-(293).

Given that (293b) and (293c) map eight linear orders each while (291a), (291c), (292a),

(292b) and (293a) map four orders each, the total number of strings mapped from the present account amounts to thirty-six. What has just been stated trivially implies that certain permutations of U20 can be generated by two or more distinct derivation trees.[192] Thus, instead of grouping together the linear orders of U20 that are mapped from a certain derivation tree, it is arguably more profitable to pick a certain permutation of U20 and show what derivation trees map it, and then move to the next permutation. Such a strategy bears a twofold advantage. On the one hand, it allows to easily check how consistent the superscripts are in the alternatives for the same permutation that are generated by the present account. On the other hand, is permits to create separate groups for the permutations of U20 that in chapter 4 turned out to bear distinct correlations with the clausal domain.

The permutations that the present approach accounts for can now be listed. (299) lists the two attested permutations that negatively correlate with an OV linear order in main declarative mono-transitives, while (302) lists those that can be argued to negatively correlate with VS under the same conditions. The remaining orders are instead listed in (300) and (301). The former lists the permutations of U20 that occur with a relatively high number of V-initial languages even though they do not negatively correlate with OV, whereas (301) lists the two orders that have been shown to include few V-initial languages, which can be regarded either as exceptions that can be explained away in some other manner or as data that an adequate theory must capture.

(299)	a.	Num^{L}	N^{PP}	A^{R}	| $Dem^{R/L}$	[mapped from (292a)]
	b.	Num^{L}	N^{PP}	A^{R}	Dem^{R} |	[m. f. (293b)]
	c.	Num^{L}	N^{B}	A^{R}	Dem^{R}	[m. f. (293c)]
	d.	Num^{L}	A^{L}	N^{PP}	| $Dem^{R/L}$	[m. f. (292a)]
	e.	Num^{L}	A^{L}	N^{PP}	Dem^{R} |	[m. f. (293b)]
	f.	Num^{L}	A^{L}	N^{B}	Dem^{R}	[m. f. (293c)]

[192] Due to the exclusion of (291b) and (292c) dictated by the principle in (298), Dem-N-Num-A and N-Num-A-Dem are exclusively obtained from (291a) and (293a), respectively. As a consequence, these two permutations are different from all other orders listed in (299)-(302), which are instead mapped from two or more of the derivation tree families in (291)-(293). If one thus wanted to reduce the redundancies to the minimum, (291a) and (293a) would be indispensable, whereas the exclusion of some other derivation tree family would be theoretically possible. But of course, a choice in this respect would need further theoretical or empirical support.

(300)	a.	Dem^L	Num^L	N^{PP}	A^R |	[mapped from (293b)]
	b.	Dem^L	Num^L	N^B	A^R	[m. f. (293c)]
	c.	Dem^L	Num^L	A^L	N^{PP} |	[m. f. (293b)]
	d.	Dem^L	Num^L	A^L	N^B	[m. f. (293c)]
	e.	N^{PP}	| Dem^L	Num^L	$A^{R/L}$	[m. f. (293a)]
	f.	N^{PP}	Dem^R	| Num^L	$A^{R/L}$	[m. f. (291a)]

(301)	a.	N^{PP}	A^R	| Dem^L	$Num^{R/L}$	[mapped from (291c)]
	b.	N^{PP}	A^R	Dem^R	| $Num^{R/L}$	[m. f. (292b)]
	c.	N^{PP}	| $A^{R/L}$	Num^R	Dem^R	[m. f. (293a)]
	d.	N^{PP}	A^R	| $Num^{R/L}$	Dem^R	[m. f. (291c)]
	e.	N^{PP}	A^R	Num^R	| $Dem^{R/L}$	[m. f. (292a)]
	f.	N^{PP}	A^R	Num^R	Dem^R |	[m. f. (293b)]
	g.	N^B	A^R	Num^R	Dem^R	[m. f. (293c)]

(302)	a.	Dem^L	N^{PP}	| $A^{R/L}$	Num^R	[mapped from (291a)]
	b.	Dem^L	N^{PP}	A^R	| $Num^{R/L}$	[m. f. (292b)]
	c.	Dem^L	N^{PP}	A^R	Num^R |	[m. f. (293b)]
	d.	Dem^L	N^B	A^R	Num^R	[m. f. (293c)]
	e.	Dem^L	A^L	N^{PP}	| $Num^{R/L}$	[m. f. (292b)]
	f.	Dem^L	A^L	N^{PP}	Num^R |	[m. f. (293b)]
	g.	Dem^L	A^L	N^B	Num^R	[m. f. (293c)]
	h.	Dem^L	N^{PP}	| Num^L	$A^{R/L}$	[m.f. (291a)]
	i.	N^{PP}	| Dem^L	$A^{R/L}$	Num^R	[m. f. (293a)]
	j.	N^{PP}	Dem^R	| $A^{R/L}$	Num^R	[m. f. (291a)]
	k.	N^{PP}	| Num^L	$A^{R/L}$	Dem^R	[m. f. (293a)]
	l.	A^L	N^{PP}	| Dem^L	$Num^{R/L}$	[m. f. (291c)]
	m.	A^L	N^{PP}	Dem^R	| $Num^{R/L}$	[m. f. (292b)]
	n.	A^L	N^{PP}	| $Num^{R/L}$	Dem^R	[m. f. (291c)]
	o.	A^L	N^{PP}	Num^R	| $Dem^{R/L}$	[m. f. (292a)]

p. A^{L} N^{PP} Num^{R} Dem^{R} | [m. f. (293b)]

q. A^{L} N^{B} Num^{R} Dem^{R} [m. f. (293c)]

Based on the data collected in appendix A and discussed in chapter 4, the orders that negatively correlate with an OV order in mono-transitive clauses are grouped in (299). Under the present account, what unites them is that Num anchors a left auxiliary tree while Dem either anchors a right one, or is the only modifier adjoining to a different partition. By adopting C as in section 4.5.2 to state a ban on OV and D to indicate a ban on VS, the cluster in (299) can be captured as follows:

(303) $Num^{L}\ N\ <Dem^{R}> \mid <Dem^{R/L}> \rightarrow (C \wedge \neg D)$

As opposed to what discussed in section 4.5.2, the options listed in (299) do not seem to bear a common property that can be used as a generalization of their banning of VS. However, one can alternatively regard the permutations of U20 listed in (300) and (301) as the complement set of those properties that, in addition to (303), may lead to a permission of VS. Hence, one can formulate the following:

(304) a. $\mid Dem^{L}\ Num^{L} \mid \rightarrow \neg D$

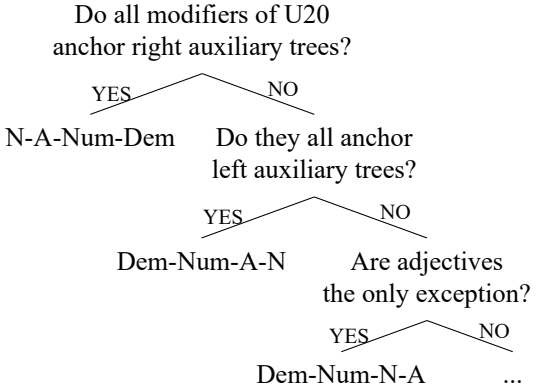
b. $A^{R} \mid Dem^{L} \rightarrow \neg D$

The premise in (304a) is what unites Dem-Num-N-A, Dem-Num-A-N and N-Dem-Num-A in (300a)-(300e), i.e. both Num and Dem anchor left auxiliary trees belonging to the same partition. The exclusion of (300f) from the generalization can be seen as an advantage as soon as one recalls that the data in appendix A.8 suggest that, among the languages that have N-Dem-Num-A as the unmarked order of U20, the V-final and V-initial ones belong to distinct families and genera.

One can further observe that (304a) can potentially apply to (301a), thus capturing the V-initial languages that have been gathered for N-A-Dem-Num in appendix A.7. However, the insightfulness of (304a) for N-A-Dem-Num and of (304b) for both N-A-Dem-Num (301a) and N-A-Num-Dem (301e) largely depends on one's take with respect to the V-initial languages that are listed in appendices A.2 and A.7. As just argued for N-Dem-Num-A, the fact that the statements in (304) apply only to (301a) and (301e) can be a possible explanation for the fact V-initial languages are possible, though relatively infrequent.

The latter take on N-A-Num-Dem fosters the discussion of another point that the present analysis has in common with Abels & Neeleman (2009, 2012) and Georgi & Müller (2010), namely that a broader range of alternate derivations for a single permutation of U20 does not necessarily imply a higher crosslinguistic frequency for such an order (see again section 3.2.4). In section 3.2.6, it has been observed that the higher frequency of certain orders of U20 has been linked to the fact that they are easier to acquire, either because of the iso- or homo-morphic property of the structure (Martin et al. 2019) or due to intracategorial harmony (Dryer 2018a: 801). As anticipated at the end of section 5.4.5, these intuitions can be captured by sketching a heuristics for the acquisition of the structures in (291)-(293) that goes along the lines of (305) on the current page (see DelBusso 2020: Sec. 5.2 and references therein for further discussion):

(305)



Thus, even though (299) shows that Num-N-A-Dem and Num-A-N-Dem can be accounted for by means of three distinct derived trees each, their degraded frequency with respect to Dem-Num-A-N and Dem-Num-N-A can be due to the fact that the latter occupy higher nodes in (305) than the former. The formulation of (305) also allows to remark that the non-pied-piped nature of the derived trees mapped from (293c) is not a direct predictor of their higher crosslinguistic frequency, contrary to what is argued in many works following Abels & Neeleman (2009, 2012) (e.g. Culbertson & Adger 2014, Martin et al. 2020).

In opposition to the L-TAG accounts of U20 provided in sections 5.2.2 and 5.2.3, the analysis that has just been presented manages to establish consistent dependencies among the derivation trees in (291)-(293), no matter if N raises (with some of its arguments) or not. Thus, (291)-(293) are coherent with the desiderata that are generally met by 2-delayed 2-MC-TAGs (sections 5.4.3 and 5.4.4).

Still, the accounts presented in section 5.2.3 this section share the core idea, namely that an adequate TAG adaptation of U20 can benefit from the insights given by a pied-piping approach *à la* Cinque (2005, 2023) so as to generate the non-homomorphic permutations. Although it has been argued for the eventual MC-TAG alternative that all modifiers of U20 always adjoin to a projection of the lexical N as per Bruening (2018a), such an account still needs to extend the nominal projection by means of some functional structure, as in the L-TAG adaptation of Abels & Neeleman (2009, 2012).

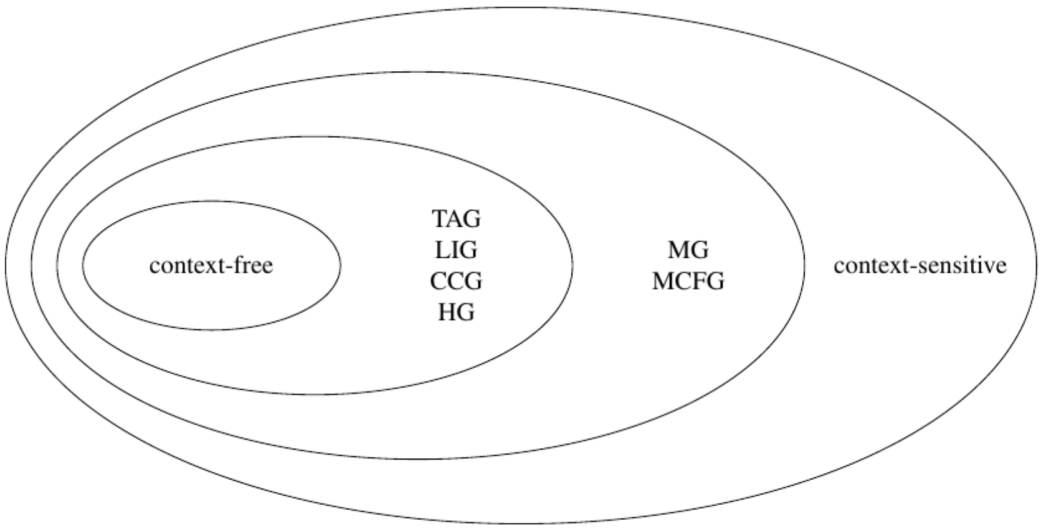
Thus, neither of the two accounts relying on pied-piping is entirely compliant with a Lexical Head Hypothesis to the nominal domain (Corver 2013), but only the one presented in this section remains consistent regardless of whether N raises or not. This follows from the fact that in the second analysis the functional structure is never targeted by the modifiers of U20. Put differently, no problem arises if N does not raise, as in such a scenario the presence vs. absence of XP and of some material Y is irrelevant as far as one's only focus is represented by the modifiers of U20.

The aspects just pointed out thus counteract the rejection of the Functional Head Hypothesis, and find interesting parallels with other recent works which similarly argue for the need of some additional functional extension of the lexical NP projection. To adequately account for the demonstratives that are found in the German dialects of Rhine and Moselle Franconian, Rauth & Speyer (2021: Sec. 5) claim that a functional projection above NP is necessary (see again section 3.1.1 for related discussion). A similar point is raised by Krivochen (2022: 356), whose analysis of Spanish relative clauses requires a projection headed by D so as to overcome the shortcomings that would arise under an alternate account that exclusively hinges on the tenets of the Lexical Head Hypothesis.

In conclusion, it is worth observing that the TAG analysis of U20 just presented generates the sixteen orders predicted by Bruening (2018a), but provides at the same time a principled reason to set aside the two permutations that are not predicted by Cinque (2005). In this respect, the statement in (298) can be seen either as a factor that excludes Num-N-Dem-A and N-Num-Dem-A altogether, or as a motivation for their crosslinguistic rarity. The latter view would be in line with the Zipfian distribution of U20 that is argued for in Croft (2003: Sec. 3.1) and Merlo & Ouwayda (2018: Sec. 4.2, 7.1), but would raise the question on how to capture the remaining permutations of U20 (two of which attested, according to Dryer 2018a).

A solution for the latter problem is given by Chen-Main & Joshi (2008), whose work has already been presented in section 5.4.5. Chen-Main & Joshi (2008: Sec. 3-5) show that, as far as scrambling of verbal arguments in German is concerned, an L-TAG can capture a certain number of permutations, but that more can be accounted for as soon as one assumes that the elementary trees can be combined in novel ways. Their argument thus shows that the power of any TAG can be increased not only to make the latter match the exact formal power of natural language syntax, but also to provide formal grounds for the intra- and cross-linguistic rarity of certain syntactic phenomena. In fact, any two mildly context-sensitive formalisms must not be necessarily weakly equivalent. Even in that case, their generative powers can fit within the range that is strictly included between context-freeness proper on the one hand and context-sensitivity proper on the other. This concept can be illustrated with the picture in (306) on this page, which is reported from the comparison that Frank & Hunter (2021: 3) do among (L-)TAGs and other mildly context-sensitive formalisms.[193]

(306)



5.5.2 On the elementary tree that each modifier of U20 anchors

In the previous subsection, a promising account of Greenberg's Universal 20 has been eventually proposed by resorting to the 2-delayed 2-MC-TAG approach introduced and defended

[193] The acronyms in (306) stand for the following formalisms: (Lexicalized) Tree Adjoining Grammar, Linear-Indexed Grammar, Combinatory Categorial Grammar, Head Grammar, Minimalist Grammar (in the sense of Stabler 1997 and subsequent works), Multiple Context-Free Grammars. See Frank & Hunter (2021: Sec. 1) and references therein for discussion.

throughout the present chapter. Such an analysis of U20 has been grounded on a twofold premise. On the one hand, all modifers of U20 are included in distinct non-wrapping predicative auxiliary trees. On the other hand, the latter elementary trees are combined with one another by means of adjoining, so as to consistently participate in dependent derivations. However, these assumptions overlook many theoretical and empirical challenges that have been discussed in sections 3.1 and 5.2.1. The aim of the present subsection is to take them up.

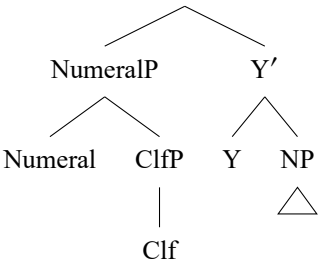
In section 5.2.1 it was argued that the extended notion of derivation provided by Schabes & Shieber (1992, 1994) has the potential to provide a more fine-grained analysis with respect to the semantic and syntactic variation of attributive adjectives. As long as the use of modifier auxiliary trees is not extended to cardinal numerals and demonstratives, admitting mixed derivations does not affect the adequacy of the analysis of U20 that has been previously discussed. In fact, the outermost predication rule posited by Schabes & Shieber (1994: 102) states that in mixed derivations modifier auxiliary trees are always lower in the derived tree than predicative trees participating in dependent derivations. Put differently, whenever a unique node is targeted by several auxiliary trees adjoining there, adjectives included in modifier trees are always closer to the noun than adjectives included in predicative trees, cardinal numerals and demonstratives. This fact perfectly resonates with the principle that was independently established on page 257 to determine legitimate dependent derivations with respect to U20, as the latter similarly posits that numerals and demonstratives must dominate adjectives in any chain of adjoining operations.

At this point, the focus can shift on the internal structure of the auxiliary trees that have been assumed in sections 5.2.2, 5.2.3 and 5.5.1. In this respect, it is worth highlighting once more that the assumption of a distinct elementary tree for each modifier of U20 says little about the internal complexity of the modifiers themselves. This is a welcome aspect of TAGs, because intra- and cross-linguistic variation with respect to the (non-)phrasal nature of nominal modifiers is expected (see again Frank 2013: 238). Moreover, this flexibility neither complicates nor prevents a uniform analysis that is concerned with the manners in which the respective elementary trees can be combined with one another.

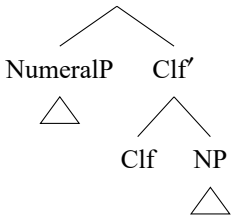
Having structures that can vary internally is an appealing aspect of TAGs that can be observed by taking up once again the analysis given by Little, Moroney & Royer (2022) (see section 3.1.2). The latter assume two distinct structures, reported once again in (307) for con-

venience, to account for the crosslinguistic syntactic variation of numeral-noun constructions where sortal classifiers are involved. In the present work, they have been first presented on page 98, and in the following pages it has been argued that assuming either alternative or both does not help explain typological facts like the Greenberg-Sanches-Slobin Generalization (GSSG, Greenberg 1990–1972, Sanches & Slobin 1973) nor the fact that the noun is hardly ever attested between the cardinal numeral and the sortal classifier.

(307) a.

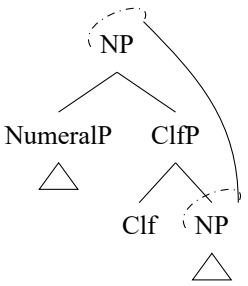


b.

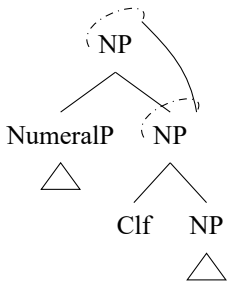


TAGs allow to recast (307a) and (307b) as shown in (308a) and (308b), respectively. Readers can observe that (308) allows to improve the analysis given by Little, Moroney & Royer (2022) in one major respect, i.e. a single derived tree needs to be assumed to capture a single syntactic phenomenon. Contrary to (307a), the classifier-for-numerals analysis in (308a) no longer needs a functional projection whose specifier be filled by a constituent formed by the numeral and the sortal classifier. In fact, the local relation between the two is already captured by having them included within the same elementary tree. On a similar ground, in the classifier-for-nouns counterpart the sortal classifier no longer needs to be a projecting head that selects for both the noun and the cardinal numeral as in (307b). In (308b), the sortal classifier is part of the initial tree that is anchored by the noun. This predicts the co-occurrence of the two in absence of the cardinal numeral, as the presence of the latter depends on whether a separate auxiliary tree containing it is adjoined or not.

(308) a.



b.



At this point, one can speculate on how the GSSG and the numeral-classifier adjacency can be captured by means of the derived trees in (308). With respect to the GSSG, one can assume that the presence of a Clf head prevents the merger of a distinct Num head *à la* Ritter (1991) in both (308a) and (308b). Alternatively, one may argue for the merger of a distinct Num head, with Clf and Num becoming one indistinguishable unit as a consequence of head movement or feature attraction (recall that these operations may also take place across elementary trees in the ways discussed in section 5.4.4).

Similarly, there are at least two possible manners to account for the numeral-classifer adjacency in (308). A first way is to argue that the branching directionality of the elementary trees involved is constrained as follows. In (308b), a left auxiliary tree containing a numeral can only adjoin to initial trees where the sortal classifier is prenominal; if the latter is postnominal, the auxiliary tree containing the numeral must be of the right type. The restrictions are even simpler for (308a), as in such a case it is sufficient to ban wrapping auxiliary trees containing both a numeral and a sortal classifier. However, on page 100 (at the end of section 3.1.2) the following Welsh numeral-noun construction was reported from Willis (2006: 1828):

(309) *y tair merch ddawnus ar hugain hyn*
DET three girl talented on twenty DEM.PROX
'These twenty-three talented girls'

TAGs can easily capture examples like (309) by assuming that complex numerals framing the noun are included in wrapping auxiliary trees, but this apparently jeopardizes the constraints that have just been sketched for (308). However, cases like (309) may not be problematic as they first seem. In section 5.4.5, facts from the German clausal domain have been presented to argue that each wrapping auxiliary tree can also be analyzed as the graph union of two non-wrapping trees that bear opposite directionalities. Applied to (309), this would mean that the wrapping tree containing the numeral is actually made of a left and a right auxiliary tree, where the former includes *tair* 'three' and the latter the PP *ar hugain* 'on twenty'. Getting back to (308), the constraints previously mentioned for (308b) can thus be left unchanged, as the requirement of either a left or a right auxiliary tree is always trivially satisfied by one of the two subparts of a wrapping tree.

A bit more must be said with respect to (308a). In this case, the adjacency between the

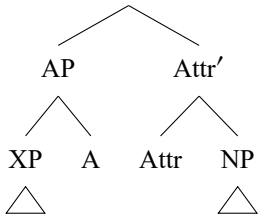
numeral and the sortal classifier can be accounted for by positing that only a complex numeral can give rise to a wrapping auxiliary tree. Put differently, one must state that a sortal classifier cannot be the unique element included in one of the two non-wrapping subparts that form a wrapping auxiliary tree, so that it is never the decisive factor that determines the wrapping nature of the tree it appears in. In section 5.4.4, the morphosyntactic behavior just assumed for sortal classifiers has been linked to syncategorematicity, and in fact this would not be the only point supporting their syncategorematic nature. On the one hand, the classifier is a sister of the foot node of the auxiliary tree in (308a), which is in line with the syncategorematic exception stated on page 344. On the other hand, one can argue that the interpretation of sortal classifiers remains constant across models (Krivochen 2023d: 321-322) because, as soon as they are understood as multipliers, sortal classifiers represent an identity (or neutral) element in a mathematical sense. In fact, Wu & Her (2021: 43) hinge on Greenberg (1990–1972: 172) in observing the following constrast between sortal (C) and mensural (M) classifiers (italics in the original):

> A long-neglected mathematical perspective offers a unique insight to the relation between Num and C/M, as well as the precise distinction between C and M. The earliest source of such a view we know of is Greenberg [1990: 172], where he states with unmistakable clarity that «all the classifiers are from the referential point of view merely so many ways of saying 'one', or more accurately, 'times one'». Thus, *san tiao yu* [three C fish] [...] is seen as [[*3*×*1*] fish]. [...] Under this precise formulation, [...] the one unit denoted by Cs has the precise numerical value of *one*, but the one unit denoted by Ms does not [...].

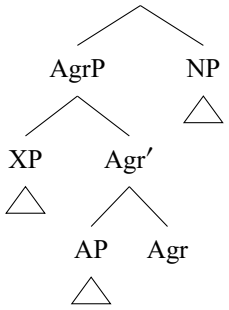
The second aforementioned way to account for the numeral-classifer adjacency in (308) does not necessarily conflict with the first one discussed so far, as it consists in (further) constraining the mapping from syntax to phonology. This can be done by resorting to principles like those discussed in Alexeyenko & Zeijlstra (2021a: Sec. 3.2) and Zeijlstra (2023: 197), which play a central role in the theoretical implementation of the Modifier-noun Adjacency Generalization (MAG) discussed on page 111 in section 3.1.3. At the end of that subsection, it was observed that Alexeyenko & Zeijlstra (2021b,a) share with Little, Moroney & Royer (2022) the idea of proposing a dual source for the single syntactic phenomenon they examine, i.e. the fact that in certain languages no phrasal projection can intervene between a noun and an attributive adjective. The representations discussed on page 113 can be reported in (310)

so as to make it easier to see that TAGs allow once again to recast the analysis by making the two derived trees in (311) less different than they would appear otherwise.

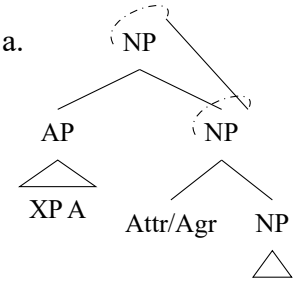
(310) a.



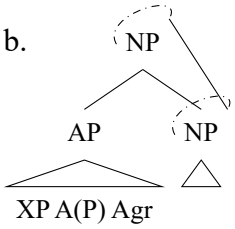
b.



(311) a.



b.



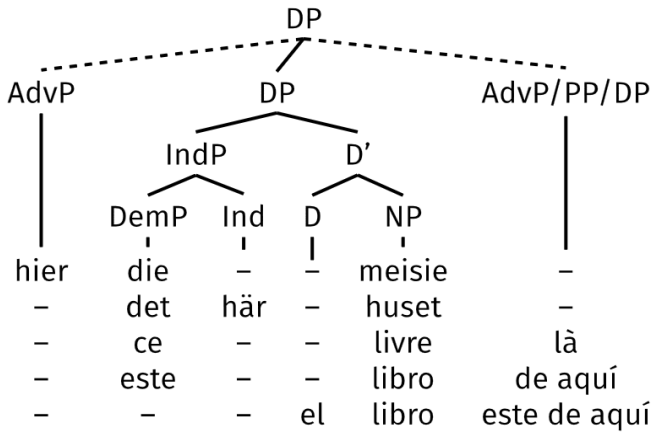
Both structures in (311) *prima facie* seem to be modifier auxiliary trees, rather than predicative auxiliary trees. As stated at the beginning of the present section, this take on such elementary trees would not represent a theoretical problem for the analysis of U20 given in the previous subsection, but it predicts a flexible order of these complex attributive adjectives with respect to others (whenever present). However, the representations in (311) raise a problem already discussed in section 5.2.2, i.e. they simplify the internal structure of the auxiliary trees up to a point where they no longer appear as predicative trees. However, this does not imply that they are not. This aspect is particularly relevant with respect to the exact structural position of the Agr head in (311b), as a TAG analysis of MAG may work even without assuming that Agr projects and selects PP as shown in (310b) by following Alexeyenko & Zeijlstra (2021b: 23, 2021a: 55). In this respect, it is worth recalling that the exact status of Agr in both (310b) and (311b) must be left for future work, as it "might even be introduced on the lexical A-head itself" (Alexeyenko & Zeijlstra 2021b: 22).

The derived tree in (311a) recasts (310a) by making two major changes. First, the derived tree in (310a) is no longer rooted at AttrP, but rather at NP. This fact is not in sharp contrast with the original account as it fist seems. In fact, Alexeyenko & Zeijlstra (2021b: 23-24)

argue that Attr selects for an NP and then projects its own categorial |N| feature. Second, the functional head in (311a) is indicated as Attr/Agr to recall that, among the other possibilities to be investigated further, Attr "may in fact be categorially identical with Agr" (Alexeyenko & Zeijlstra 2021a: 55). Such a theoretical option would make the derived trees in (311) parallel to those proposed in (308) for numerals and classifiers, as in both cases the crucial factor to examine is whether Clf or Attr/Agr is merged in the initial tree or in the auxiliary tree.

Depending on whether and how MAG is applied in a language, a PP in (311) might be extraposed, i.e. realized on the opposite side of the noun in comparison with the adjective. Thus, the theoretical appeal of wrapping trees discussed for numerals arguably extends to adjectives and, in a similar fashion, to demonstratives. In fact, the discussion with respect to the latter in section 3.1.1 has shown that Rauth & Speyer (2021: 17) account for the morphosyntax of demonstratives by resorting to the structure first reported on page 90, and reproposed in (312) on the current page for convenience. As soon as a postnominal locative reinforcer co-occurs with a prenominal demonstrative, the assumption of a wrapping auxiliary tree becomes a worthwhile option.

(312)



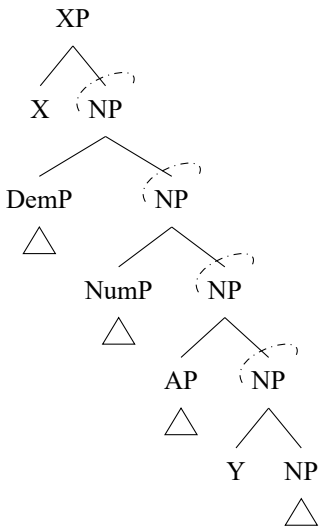
However, if the entire DP projection shown in (312) is to be understood as an auxiliary tree adjoined at NP, one must conclude that the D head is not part of the initial tree anchored by the noun. Given that the analysis of U20 provided in the previous subsection has argued in favor of the FHH (Corver 2013) on independent reasons, it is better not to exclude a priori the role that D is often argued to play with respect to reference and argumenthood (Giusti in press). In this respect, seeing D as something that is optionally adjoined may turn out to be problematic, depending on one's theoretical assumptions.

One first possible solution to this issue consists in arguing that D is not part of the auxiliary tree containing the rest of the demonstrative as per (312). In this scenario, D can be an instance of the X head that was assumed in sections 5.2.3 and 5.5.1 to provide TAG adaptations of pied-piping. To illustrate the case, (313b) shows the derived tree that is obtained if all modifiers of U20 are left auxiliary trees that adjoin to the non-raised N, as summarized in (313a). If D is realized in X, (313b) is at first sight in line with the analysis of demonstratives that is proposed by Roehrs (2010), Baggio (2019) and Dékány (2021: 148).

Contrary to what has been discussed in section 3.1.1, the eventual realization of Dem in Spec,D is however not a possibility in (313b), given that phrasal movement across elementary trees is banned in TAGs. Thus, one should claim that in (313b) the often argued spec-head relation between D and Dem is the result of D-lowering (Harizanov & Gribanova 2019: Sec. 3) as a consequence of the inter-elementary-tree operations allowed by Frank (2002: 161) and Frank (2006: Sec. 5) (see again section 5.4.4).

(313) a. Dem^{L} Num^{L} A^{L} N^{B}

b.



An alternate take to solve the apparently problematic status of D consists in arguing that it is part of the same auxiliary tree that also includes the demonstrative, but that the latter is not combined to the rest of the nominal domain by resorting to adjoining. Instead, it is introduced in the derivation by means of (a chain of) substitution(s). In this respect, it is worth concluding the present section by observing that such an operation has been fundamental in solving many issues discussed in section 5.4, but that it turned out to be unnecessary in order to come up with the account for U20 eventually presented in section 5.5.1. The exploration

of the theoretical implications and advantages that chains of substitutions can provide to the analysis of the nominal domain are left for future work.

Chapter 6

Conclusions

Among the various universals proposed in Greenberg (1966–1963), his Universal 20 stands out for two reasons. First, it tries to generalize over the order of four elements belonging to a single domain without resorting to any implicational relation. Second, it soon turned out to exclude some orders that are actually attested. Still, some principle along those lines stands and asks for a reformulation, as some orders remain unattested even if more and more data are collected crosslinguistically (Dryer 2018a,b, Cinque 2021). These facts have fostered subsequent research aimed at assessing the empirical breadth of the attested permutations on the one hand, and at providing a potential explanation of the patterns on the other hand.

The two lines of research have become intertwined over time, as two main ideas have gradually matured. The first is that it is not possible to understand the phenomenon by adopting a perspective that is either exclusively theoretical or exclusively empirical. The second is that, no matter the theoretical premises, an adequate account of U20 is inevitably related to the verification of another hypothesis, namely that there is some compelling parallel between the clausal and the nominal domains. The present dissertation has shown (from chapter 4 on) that there are both empirical and theoretical grounds to argue that the two domains affect one another. The observations that support this claim have been presented in a rather pre-theoretic sense in section 4.4, whereafter section 4.5.2 and section 5.5.1 have tried to give possible formalizations to capture them in a more meaningful way.

These facts have brought to the conclusion that the relation between the nominal and the clausal domains can be captured (once some due theoretical abstraction is applied) neither by assuming that their respective structures are completely independent of one another, nor by arguing that the two are completely paralleling each other. Instead, the adoption for the

nominal domain of what Corver (2013) addresses as the Functional Head Hypothesis (see sections 2.2.3 and 2.2.4) has been suggested. The FHH is better formulation of the DP-hypothesis, as it is not committed to the label and the formal properties of a specific functional projection (i.e. DP), but still backs up the intuitions that motivate and support the rejection of the alternative, namely the Lexical Head Hypothesis (a.k.a. the NP-hypothesis).

Chapter 5 has dealt with Tree Adjoining Grammars, which turned out to be a good theoretical approach to grasp the complexity of what is analyzed in the present work. In fact, TAGs provide a local complication of some clearly determined and constrained structure in favor of a global simplification of the whole system (Bangalore & Joshi 2010). These properties of the formalism are intertwined with an approach to both the nominal and the clausal domains where successive-cyclicity is largely (if not utterly) dispensed with (see again section 5.1).

As discussed in sections 3.2, 4.4, 4.5, 5.2 and 5.5, the focus on both domains is motivated by the relevance that the debate between the LHH and the FHH plays in determining the possible ways to formalize Greenberg's Universal 20 (Greenberg 1966–1963: 87, 111). Thus, many detours were needed to frame U20 into a broader picture that encompasses the disputed presence of parallels between the nominal and the clausal domains on the one hand (see chapter 2), and an assessment of the premises to adopt for a novel analysis on the other hand.

Two steps have been necessary for the assessment of the latter premises. First, section 4.1 has discussed the methodology that has led to the data collection in appendix A, which represents in turn the empirical backbone of the analysis. Second, section 3.2 has presented the various theoretical proposals that have inspired the sampling and that have subsequently been put to test against the gathered evidence (see the end of section 4.3.4, as well as sections 4.4 and 4.5). The limits that have been observed have prompted the adoption of TAGs. However, the latter have been shown to suffer in turn from a number of theoretical problems of their own, which were surveyed and then solved (see in particular sections 5.3 and 5.4).

The eventually endorsed 2-delayed 2-MC-TAG approach bears a twofold merit. On the one hand, section 5.5.2 shows that the tree-internal structural variability allowed by the formalism can be exploited to formulate a single derived tree per morphosyntactic phenomenon (e.g. sortal classifiers in numeral-noun constructions and complex adjectival phrases used attributively), as opposed to competing analyses where two unrelated structures must be assumed for the same purpose (Little, Moroney & Royer 2022, Alexeyenko & Zeijlstra 2021b,a). On the other hand, sections 5.4 and 5.5 have shown that this formal flexibility

does not lead to the formulation of an account for U20 that contradicts the data or that is not consistent with its own premises. This eventual success goes against what was seen for both the competing analyses surveyed in section 3.2 and the L-TAG adaptations proposed in sections 5.2.2 and 5.2.3.

In addition to the findings mentioned so far, an important contribution of this dissertation lies in the annotated commentary on previous works. This has been carried out in chapters 2 and 3 by enumerating and comparing several distinct formal approaches that have been related to insights from the typological tradition. As a result of this operation, some specific observations have emerged that might be relevant for the interested practitioners. A first example in this sense was discussed on page 61, where it is pointed out that a movement-triggering feature like the one proposed by Biberauer, Holmberg & Roberts (2014) makes contradictory predictions with respect to what they call *spines* (i.e. an uninterrupted sequence of multiple categorically-identical extended projections) whenever the latter have mixed directionality as in Basque (Duguine 2022). Another example came out in section 3.1.3, where the analyses provided by Haider (2022b) and Alexeyenko & Zeijlstra (2021b,a) demonstrably struggle to capture syntactic restrictions applying to phrasal adjectives in Basque and adverbials in the Romance and Germanic languages, respectively.

Other specific observations were made in section 3.2.6, where a novel analysis has been given for the significatively stronger preference that Culbertson & Adger (2014) and Martin et al. (2019, 2020) observe for N-A-Dem as opposed to other homomorphic patterns with postnominal modifiers. In addition to that, section 3.2.6 has also highlighted potential counterevidence to the formal universal proposed by Steedman (2020), as well as a potential explanation of Greenberg's Universal 18 (Greenberg 1966–1963: 87, 111). With respect to the latter, it has been argued that all one would need to assume is that (at least) some orders that are base-generated in Abels & Neeleman (2009, 2012) are constrained by FOFC (on the latter, see sections 2.3.2 and 2.3.3).

The survey of past works has also fostered a series of observations that can be of interest to those working in the field of formal language theory. In particular, a novel definition for the Markovian property of TAGs has been given (see the end of section 5.1.4), as well as a novel adaptation of independent and mixed derivations to capture the peculiar properties that Svenonius (1994) points out with respect to a phrase like *a dead dangerous animal* (section 5.2.1). Then, non-monotonic derivations *à la* Heck (2016, 2023) have been adopted in

section 5.3.2 to justify the formation of multi-component sets of elementary trees. This has fostered in turn the introduction of chains of substitutions (section 5.4.2), which represent a crucial alternative to the adjoining operation and are possible only in k-delayed h-MC-TAGs where $k \geq 1$ and $h \geq 2$ (see page 295 ff. for discussion of the terminology). More specifically, a 2-delayed 2-MC-TAG has been adopted in the present work, due to the theoretical and empirical motivations put forth in sections 5.3.3 and 5.4.2. Still, this TAG alternative has been argued to be weakly (though, crucially, not strongly) equivalent to other formalisms listed on page 335, so that further analyses and comparisons are expected in this respect.

Finally, as discussed in section 5.4.3, one should not follow Longenbaugh (2014) in rejecting MC-TAGs, because chains of substitutions provide a solution for a twofold problem of (L-)TAGs. On the one hand, such chains allow to account for clausal complementation and long extractions in a uniform way that circumvents the problems raised by Georgi (2019), while on the other hand they provide a solution for the missing link problem on derivation trees (Kallmeyer & Kuhlmann 2012, Chen-Main & Joshi 2014). In addition to that, certain well-known crosslinguistic differences can be captured under a raising-flavored and control-flavored dichotomy that resonates with Wurmbrand (2001: 158) and references therein. The latter distinction has been first discussed in section 5.4.3, and finds even broader applications as soon as one adopts what sections 5.4.4 and 5.5.2 address, namely a Feature-Based TAG whose auxiliary trees are allowed to contain syncategorematic heads as sisters of the foot nodes.

Appendix A

Typological data

This appendix is made of tables which list the relevant properties of the sample presented and motivated in chapter 4. As discussed in section 4.2.1, glottocodes are the best suited tool to identify a specific language that might be addressed with different names across different sources. All the tables thus feature two columns that respectively list for each entry a glottocode and a plain name used by Dryer (2018b) and Cinque (2021) or Grambank (Skirgård, Haynie, Blasi, et al. 2023, Skirgård, Haynie, Hammarström, et al. 2023).

Contrary to what one might think, providing both a glottocode and a plain name is not redundant for the following reasons. First, latter is reported in roman type when the order of U20 in such a language is classified exclusively by Cinque (2021), while it is italicized when it is present in both Dryer (2018b) and Cinque (2021).[194] A second reason is that Cinque (2021) and Dryer (2018b) may report a linear order of U20 for a specific variety of a language, while Grambank reports in one single entry the clausal word order found across different varieties of that language. In these fourteen cases, the entry in Grambank is thus a superset of the former, and is indicated as such between brackets.[195] Also the opposite takes place, i.e. Cinque (2021) and Dryer (2018b) provide the order of U20 for a language, while Grambank records two or more varieties of the language separately. In this case, the language was kept in the sample only if Grambank's classification was consistent across varieties (see section 4.1), and all individual glottocodes used by Grambank were assigned to the single language found in Cinque (2021) and Dryer (2018b).

The template of the table in appendix A.1 is different from the others showed from ap-

[194] It is important to provide this information also because the languages classified by Dryer (2018b) do not necessarily overlap with the ones whose relevant values are recorded on WALS (Dryer 2013f,b,d,c).

[195] The glottocode chosen in these situations is the one of the 'superset language' found on Grambank.

pendix A.2 on. The columns with language names and glottocodes are the third and the fourth (from the left), respectively, while they are the leftmost and the rightmost in the tables in the remainder of the appendix. This difference in appendix A.1 is done to respect the *Family* > *Genus* > *Language* hierarchy. Families are listed in alphabetical orders, ditto for genera belonging to the same family and for languages belonging to the same genus.[196] The entry of a language isolate is distinguishable for having the same name in the *Family*, *Genus* and *Language* cells.

Each section from appendix A.2 focuses on a single order of U20, by providing a table where the languages featuring it are reported. Sections are ordered from the most to the least frequent in the present sample (see section 4.2.2). The second column from the left assigns for each entry one of the types defined in section 4.1.4. This is established by following the data from APiCS and WALS (Huber & the APiCS Consortium 2013 and Dryer 2013f, respectively), which are reported in the column *WALS*, and by hinging on the data from Grambank (Skirgård, Haynie, Blasi, et al. 2023, Skirgård, Haynie, Hammarström, et al. 2023). The *V-I*, *V-M* and *V-F* columns respectively report whether a language is verb-initial according to GB131, verb-medial according to GB132, or verb-final according to GB133.

A.1 All the languages in the sample, listed for family and genus

Family	Genus	Language	Glottocode	Macroarea
Abkhaz-Adyge	Abkhaz-Adyge	Kabardian	kaba1278	Eurasia
Abkhaz-Adyge	Abkhaz-Adyge	Ubykh	ubyk1235	Eurasia
Abun	Abun	*Abun*	abun1252	Papunesia
Afro-Asiatic	Berber	Berber (Figuig)	tazn1238	Africa
Afro-Asiatic	Berber	Berber (Rif)	tari1263	Africa
Afro-Asiatic	Berber	Siwi	siwi1239	Africa
Afro-Asiatic	Berber	Tamashek	tama1365	Africa
Afro-Asiatic	Biu-Mandara	Buduma	budu1265	Africa
Afro-Asiatic	Biu-Mandara	Buwal	buwa1243	Africa

[196] Macroareas could not fit into the hierarchy because languages from the same genus (and genera from the same family) can span across different macroareas, so that no consistent order can be defined on that base.

Family	Genus	Language	Glottocode	Macroarea
Afro-Asiatic	Biu-Mandara	Gemzek (Zulgo-Gemzek superset)	zulg1242	Africa
Afro-Asiatic	Biu-Mandara	Gude	gude1246	Africa
Afro-Asiatic	Biu-Mandara	Lamang	lama1288	Africa
Afro-Asiatic	Biu-Mandara	*Malgwa*	malg1251	Africa
Afro-Asiatic	Biu-Mandara	Mpade	mpad1242	Africa
Afro-Asiatic	Biu-Mandara	*Uldeme*	wuzl1236	Africa
Afro-Asiatic	Biu-Mandara	Zulgo (Zulgo-Gemzek superset)	zulg1242	Africa
Afro-Asiatic	Creoles and Pidgins	Nubi	nubi1253	Africa
Afro-Asiatic	Cushitic	Alaba-K'abeena	alab1254	Africa
Afro-Asiatic	Cushitic	Arbore	arbo1245	Africa
Afro-Asiatic	Cushitic	Dahalo	daha1245	Africa
Afro-Asiatic	Cushitic	Elmolo	elmo1238	Africa
Afro-Asiatic	Cushitic	Gorowa	goro1270	Africa
Afro-Asiatic	Cushitic	Hadiyya	hadi1240	Africa
Afro-Asiatic	Cushitic	Iraqw	iraq1241	Africa
Afro-Asiatic	Cushitic	Konso	kons1243	Africa
Afro-Asiatic	Cushitic	*Oromo (Harar)*	east2652	Africa
Afro-Asiatic	Cushitic	*Oromo (Waata)*	waat1238	Africa
Afro-Asiatic	Cushitic	Qafar	afar1241	Africa
Afro-Asiatic	Cushitic	*Rendille*	rend1243	Africa
Afro-Asiatic	Cushitic	*Sidaama*	sida1246	Africa
Afro-Asiatic	Cushitic	Somali	soma1255	Africa
Afro-Asiatic	Cushitic	Xamtanga	xamt1239	Africa
Afro-Asiatic	Cushitic	*Yaaku*	yaak1241	Africa
Afro-Asiatic	East Chadic	Baraïn (Barayin)	bare1279	Africa
Afro-Asiatic	West Chadic	Goemai	goem1240	Africa
Afro-Asiatic	West Chadic	Hausa	haus1257	Africa
Afro-Asiatic	West Chadic	*Miya*	miya1266	Africa

Family	Genus	Language	Glottocode	Macroarea
Afro-Asiatic	West Chadic	Ngas	ngas1240	Africa
Afro-Asiatic	West Chadic	*Ngizim*	ngiz1242	Africa
Afro-asiatic	West Chadic	Nyam	nyam1285	Africa
Afro-Asiatic	West Chadic	Pa'anci	paaa1242	Africa
Afro-Asiatic	West Semitic	Amharic	amha1245	Africa
Afro-Asiatic	West Semitic	Ancient Hebrew	anci1244	Eurasia
Afro-Asiatic	West Semitic	Arabic (Egyptian)	egyp1253	Africa
Afro-Asiatic	West Semitic	Arabic (Modern Standard)	stan1318	Eurasia
Afro-Asiatic	West Semitic	Argobba	argo1244	Africa
Afro-Asiatic	West Semitic	Baggara Arabic	chad1249	Africa
Afro-Asiatic	West Semitic	Chaha	seba1251, chah1248	Africa
Afro-Asiatic	West Semitic	*Gulf Arabic*	gulf1241	Eurasia
Afro-Asiatic	West Semitic	Hebrew (Modern)	hebr1245	Eurasia
Afro-Asiatic	West Semitic	Maltese	malt1254	Eurasia
Afro-Asiatic	West Semitic	Neo-Aramaic (Arbel Jewish)	lish1245	Eurasia
Afro-Asiatic	West Semitic	Wolane	wola1253	Africa
Ainu	Ainu	Ainu	ainu1240	Eurasia
Algic	Algonquian-Blackfoot	Arapaho	arap1274	North America
Algic	Algonquian-Blackfoot	*Ottawa*	otta1242	North America
Algic	Yurok	Yurok	yuro1248	North America
Anêm	Anêm	*Anêm*	anem1249	Papunesia
Angan	Angan	Angaataha	anga1290	Papunesia
Angan	Angan	*Menya*	meny1245	Papunesia
Araucanian	Araucanian	*Mapudungun*	mapu1245	South America
Arawakan	Caribbean Arawakan	Garifuna	gari1256	North America

Family	Genus	Language	Glottocode	Macroarea
Arawakan	Caribbean Arawakan	*Goajiro*	wayu1243	South America
Arawakan	Central-Eastern Maipuran	*Palikur*	pali1279	South America
Arawakan	Japura-Colombia	Achagua	acha1250	South America
Arawakan	Japura-Colombia	Baniwa	bani1255	South America
Arawakan	Japura-Colombia	Curripaco	curr1243	South America
Arawakan	Japura-Colombia	*Resígaro*	resi1247	South America
Arawakan	Japura-Colombia	Warekena	ware1255	South America
Arawakan	Medio Rio Negro	Baré	bare1276	South America
Arawakan	Southern Maipuran	Asháninka	asha1243	South America
Arawakan	Southern Maipuran	Ashéninka Perené	ashe1272	South America
Arawakan	Southern Maipuran	*Baure*	baur1253	South America
Arawan	Arawan	*Kulina*	culi1244	South America
Athabaskan-Eyak-Tlingit	Athabaskan	Navajo	nava1243	North America
Athabaskan-Eyak-Tlingit	Athabaskan	*Sarcee*	sars1236	North America
Athabaskan-Eyak-Tlingit	Athabaskan	*Slave*	slav1253	North America
Atlantic-Congo	Akpes-Edoid	Degema	dege1246	Africa
Atlantic-Congo	Akpes-Edoid	Emai	emai1242	Africa
Atlantic-Congo	Akpes-Edoid	*Engenni*	enge1239	Africa
Atlantic-Congo	Akpes-Edoid	Esan	esan1238	Africa
Atlantic-Congo	Bantoid	*Aghem*	aghe1239	Africa
Atlantic-Congo	Bantoid	*Akoose*	akoo1248	Africa

Family	Genus	Language	Glottocode	Macroarea
Atlantic-Congo	Bantoid	Baba	baba1264	Africa
Atlantic-Congo	Bantoid	Babanki	baba1266	Africa
Atlantic-Congo	Bantoid	*Babungo*	veng1238	Africa
Atlantic-Congo	Bantoid	Bafia	bafi1243	Africa
Atlantic-Congo	Bantoid	Bafut	bafu1246	Africa
Atlantic-Congo	Bantoid	*Bakueri*	mokp1239	Africa
Atlantic-Congo	Bantoid	Bebele	bebe1248	Africa
Atlantic-Congo	Bantoid	*Chichewa*	nyan1308	Africa
Atlantic-Congo	Bantoid	Chimakonde	mako1251	Africa
Atlantic-Congo	Bantoid	Chimiini	chim1312	Africa
Atlantic-Congo	Bantoid	Chimpoto	mpot1240	Africa
Atlantic-Congo	Bantoid	Cuwabo	chuw1238	Africa
Atlantic-Congo	Bantoid	*Digo*	digo1243	Africa
Atlantic-Congo	Bantoid	Ekegusii	gusi1247	Africa
Atlantic-Congo	Bantoid	Eton	eton1253	Africa
Atlantic-Congo	Bantoid	Gîtigania (Meru superset)	tiga1244, meru1245	Africa
Atlantic-Congo	Bantoid	Gyeli	gyel1242	Africa
Atlantic-Congo	Bantoid	Ha	haaa1252	Africa
Atlantic-Congo	Bantoid	Igembe (Meru superset)	igem1238, meru1245	Africa
Atlantic-Congo	Bantoid	Ikalanga (Kalanga superset)	ikal1242, kala1384	Africa
Atlantic-Congo	Bantoid	Ikihehe	hehe1240	Africa
Atlantic-Congo	Bantoid	Jukun	juku1254	Africa
Atlantic-Congo	Bantoid	Kagulu	kagu1239	Africa
Atlantic-Congo	Bantoid	Kemezung	keme1240	Africa
Atlantic-Congo	Bantoid	Kîîtharaka	thar1283	Africa
Atlantic-Congo	Bantoid	Kikamba	kamb1297	Africa
Atlantic-Congo	Bantoid	Kikuyu	kiku1240	Africa
Atlantic-Congo	Bantoid	Kiuru	moch1256	Africa
Atlantic-Congo	Bantoid	Kivunjo	vunj1238	Africa

Family	Genus	Language	Glottocode	Macroarea
Atlantic-Congo	Bantoid	Kiyaka	yaka1269	Africa
Atlantic-Congo	Bantoid	Kol	kolc1235	Africa
Atlantic-Congo	Bantoid	Kom (1)	komc1235	Africa
Atlantic-Congo	Bantoid	Langi	lang1320	Africa
Atlantic-Congo	Bantoid	Lingala	ling1263	Africa
Atlantic-Congo	Bantoid	Lubukusu	buku1249	Africa
Atlantic-Congo	Bantoid	Lwena	luva1239	Africa
Atlantic-Congo	Bantoid	Makaa	maka1304	Africa
Atlantic-Congo	Bantoid	Makhuwa-Enahara	makh1264	Africa
Atlantic-Congo	Bantoid	Mmen	mmen1238	Africa
Atlantic-Congo	Bantoid	Mundabli	mund1328	Africa
Atlantic-Congo	Bantoid	*Mungbam*	abar1238	Africa
Atlantic-Congo	Bantoid	Mwani	mwan1247	Africa
Atlantic-Congo	Bantoid	*Nchane*	ncan1245	Africa
Atlantic-Congo	Bantoid	Ndengeleko	nden1248	Africa
Atlantic-Congo	Bantoid	*Ndumu*	ndum1239	Africa
Atlantic-Congo	Bantoid	*Nkore-Kiga*	chig1238	Africa
Atlantic-Congo	Bantoid	Nomaande	noma1260	Africa
Atlantic-Congo	Bantoid	Nsenga	nsen1242	Africa
Atlantic-Congo	Bantoid	Nyakyusa	nyak1260	Africa
Atlantic-Congo	Bantoid	Nyoro	nyor1246	Africa
Atlantic-Congo	Bantoid	*Runyankore*	nyan1307	Africa
Atlantic-Congo	Bantoid	Setswana	tswa1253	Africa
Atlantic-Congo	Bantoid	Shimwela (Mwera)	mwer1248	Africa
Atlantic-Congo	Bantoid	Shona	shon1251	Africa
Atlantic-Congo	Bantoid	Tiene	tien1242	Africa
Atlantic-Congo	Bantoid	Tiv	tivv1240	Africa
Atlantic-Congo	Bantoid	Tuki	tuki1240	Africa
Atlantic-Congo	Bantoid	Tunen	tune1261	Africa
Atlantic-Congo	Bantoid	Xhosa	xhos1239	Africa
Atlantic-Congo	Bantoid	Xitsonga	tson1249	Africa

Family	Genus	Language	Glottocode	Macroarea
Atlantic-Congo	Bantoid	Yasa	yasa1242	Africa
Atlantic-Congo	Bantoid	Zulu	zulu1248	Africa
Atlantic-Congo	Benue-Congo Plateau	*Birom*	bero1242	Africa
Atlantic-Congo	Benue-Congo Plateau	*Migili*	liji1238	Africa
Atlantic-Congo	Bua-Kim-Day	*Day*	dayy1236	Africa
Atlantic-Congo	Cameroun-Ubangian	*Baka (in Cameroon)*	baka1272	Africa
Atlantic-Congo	Cameroun-Ubangian	*Doyayo*	doya1240	Africa
Atlantic-Congo	Cameroun-Ubangian	Kolbila	kolb1240	Africa
Atlantic-Congo	Cameroun-Ubangian	*Linda (Banda-Bambari superset)*	band1344	Africa
Atlantic-Congo	Cameroun-Ubangian	Mambai	mamb1294	Africa
Atlantic-Congo	Cameroun-Ubangian	*Mayogo*	mayo1261	Africa
Atlantic-Congo	Cameroun-Ubangian	*Mbum*	mbum1254	Africa
Atlantic-Congo	Cameroun-Ubangian	Mɔnɔ	mono1270	Africa
Atlantic-Congo	Cameroun-Ubangian	Mundang	mund1325	Africa
Atlantic-Congo	Cameroun-Ubangian	*Munzombo*	monz1249	Africa
Atlantic-Congo	Cameroun-Ubangian	Ngbaka (Ma'bo)	ngba1284	Africa
Atlantic-Congo	Cameroun-Ubangian	Ngbandi (Northen variety)	nort2774	Africa
Atlantic-Congo	Cameroun-Ubangian	*Samba Leko*	samb1305	Africa
Atlantic-Congo	Cameroun-Ubangian	*Sango*	sang1328	Africa

Family	Genus	Language	Glottocode	Macroarea
Atlantic-Congo	Cameroun-Ubangian	Zande	zand1248	Africa
Atlantic-Congo	Defoid	Arigidi	arig1246	Africa
Atlantic-Congo	Defoid	*Yoruba*	yoru1245	Africa
Atlantic-Congo	Delta Cross	Eleme	elem1253	Africa
Atlantic-Congo	Delta Cross	Gokana	goka1239	Africa
Atlantic-Congo	Delta Cross	*Ibibio*	ibib1240	Africa
Atlantic-Congo	Delta Cross	*Leggbó*	legb1242	Africa
Atlantic-Congo	Delta Cross	Ogbronuagum	ogbr1243	Africa
Atlantic-Congo	Ebira-Nupoid	*Gwari*	gbag1258	Africa
Atlantic-Congo	Ebira-Nupoid	*Nupe*	nupe1254	Africa
Atlantic-Congo	Gbaya-Manza-Ngbaka	Gbaya (Northwestern)	nort2775	Africa
Atlantic-Congo	Gbe	*Ewe*	ewee1241	Africa
Atlantic-Congo	Gbe	*Fongbe*	fonn1241	Africa
Atlantic-Congo	Gola	Gola	gola1255	Africa
Atlantic-Congo	Gur	*Chakali*	chak1271	Africa
Atlantic-Congo	Gur	Dagbani	dagb1246	Africa
Atlantic-Congo	Gur	Farefare (Gurune)	fare1241	Africa
Atlantic-Congo	Gur	*Koromfe*	koro1298	Africa
Atlantic-Congo	Gur	*Kulango*	boun1243	Africa
Atlantic-Congo	Gur	*Sisaala*	tumu1242	Africa
Atlantic-Congo	Idomoid	*Igede*	iged1239	Africa
Atlantic-Congo	Igboid	*Igbo*	nucl1417	Africa
Atlantic-Congo	Igboid	Izi	izie1238	Africa
Atlantic-Congo	Kainji	Ut-Ma'in	kagf1238	Africa
Atlantic-Congo	Ka-Togo	Avatime	avat1244	Africa
Atlantic-Congo	Ka-Togo	*Tafi*	tafi1243	Africa
Atlantic-Congo	Ka-Togo	Tuwuli	tuwu1238	Africa
Atlantic-Congo	Mel	*Baga Sitemu*	baga1272	Africa
Atlantic-Congo	Mel	Sherbro	sher1258	Africa
Atlantic-Congo	Na-Togo	Sekpele	sekp1241	Africa

Family	Genus	Language	Glottocode	Macroarea
Atlantic-Congo	Na-Togo	*Selee*	sele1249	Africa
Atlantic-Congo	Na-Togo	Siwu	siwu1238	Africa
Atlantic-Congo	North Volta-Congo	Kam	kamm1249	Africa
Atlantic-Congo	North-Central Atlantic	*Diola-Fogny*	jola1263	Africa
Atlantic-Congo	North-Central Atlantic	*Ndut*	ndut1239	Africa
Atlantic-Congo	North-Central Atlantic	*Noon*	noon1242	Africa
Atlantic-Congo	North-Central Atlantic	Pular	pula1262	Africa
Atlantic-Congo	North-Central Atlantic	*Wolof*	nucl1347	Africa
Atlantic-Congo	Nyo	*Akan*	akan1250	Africa
Atlantic-Congo	Nyo	*Anufo*	anuf1239	Africa
Atlantic-Congo	Nyo	Ga (Gã)	gaaa1244	Africa
Atlantic-Congo	Nyo	Nawuri	nawu1242	Africa
Atlantic-Congo	Nyo	*Nkonya*	nkon1248	Africa
Atlantic-Congo	Oko-Eni-Osayen	Oko	okoe1238	Africa
Atlantic-Congo	Senufo	Sucite Senoufo	sici1249	Africa
Atlantic-Congo	Senufo	Supyire Senoufo	supy1237	Africa
Atlantic-Congo	Volta-Congo	*Ega*	egaa1242	Africa
Atlantic-Congo	Volta-Congo	Fali (South)	sout2782	Africa
Austroasiatic	Aslian	*Jahai*	jeha1242	Eurasia
Austroasiatic	Aslian	Jah-Hut	jahh1242	Eurasia
Austroasiatic	Aslian	Jedek	jede1239	Eurasia
Austroasiatic	Aslian	Kensiu	kens1248	Eurasia
Austroasiatic	Aslian	Maniq	tong1308	Eurasia
Austroasiatic	Aslian	Semaq Beri	sema1265	Eurasia
Austroasiatic	Aslian	Semelai	seme1247	Eurasia
Austroasiatic	Aslian	*Temiar*	temi1246	Eurasia
Austroasiatic	Bahnaric	*Brao*	lave1249	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Austroasiatic	Bahnaric	Chrau	chra1242	Eurasia
Austroasiatic	Bahnaric	Jeh	jehh1245	Eurasia
Austroasiatic	Bahnaric	Laven	lave1248	Eurasia
Austroasiatic	Bahnaric	*Mnong (Central)*	cent1992	Eurasia
Austroasiatic	Bahnaric	*Rengao*	reng1252	Eurasia
Austroasiatic	Bahnaric	Sapuan	sapu1248	Eurasia
Austroasiatic	Bahnaric	Sedang	seda1262	Eurasia
Austroasiatic	Bahnaric	*Sre (Koho superset)*	koho1244	Eurasia
Austroasiatic	Bahnaric	*Stieng*	bulo1242, bude1236	Eurasia
Austroasiatic	Katuic	Brôu (Eastern variety)	east2332	Eurasia
Austroasiatic	Katuic	Brôu (Western variety)	west2397	Eurasia
Austroasiatic	Katuic	Kuy	kuyy1240	Eurasia
Austroasiatic	Katuic	*Pacoh*	paco1243	Eurasia
Austroasiatic	Khasi-Palaung	*Khasi*	khas1269	Eurasia
Austroasiatic	Khasi-Palaung	Pnar	pnar1238	Eurasia
Austroasiatic	Khmeric	*Khmer*	cent1989	Eurasia
Austroasiatic	Khmuic	Khmu	khmu1256	Eurasia
Austroasiatic	Khmuic	Mlabri	mlab1235	Eurasia
Austroasiatic	Mangic	Bugan	buga1247	Eurasia
Austroasiatic	Monic	Mon	monn1252	Eurasia
Austroasiatic	Mundaic	Korku	kork1243	Eurasia
Austroasiatic	Mundaic	Mundari	mund1320	Eurasia
Austroasiatic	Mundaic	*Santali*	sant1410	Eurasia
Austroasiatic	Nicobaric	Nicobarese (Car)	carn1240	Eurasia
Austroasiatic	Palaungic	Blang	blan1242	Eurasia
Austroasiatic	Palaungic	*Lawa (Eastern)*	east2330	Eurasia
Austroasiatic	Palaungic	*Palaung*	ruch1235, ruma1248, shwe1236	Eurasia
Austroasiatic	Palaungic	*Parauk*	para1301	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Austroasiatic	Pearic	*Chong*	cent2314	Eurasia
Austroasiatic	Pearic	*Kasong*	chon1285	Eurasia
Austroasiatic	Vietic	Cuoi (Hung variety)	hung1275	Eurasia
Austroasiatic	Vietic	*Thavung*	aheu1239	Eurasia
Austroasiatic	Vietic	*Vietnamese*	viet1252	Eurasia
Austronesian	Basap-Greater Barito	*Bajau (West Coast)*	west2560	Papunesia
Austronesian	Basap-Greater Barito	Maanyan	maan1238	Papunesia
Austronesian	Basap-Greater Barito	Sama Balangingi	bala1311	Papunesia
Austronesian	Basap-Greater Barito	*Yakan*	yaka1277	Papunesia
Austronesian	Bilic	*Tboli*	tbol1240	Papunesia
Austronesian	Bima-Lembata	Alorese	alor1247	Papunesia
Austronesian	Bima-Lembata	Dhao	dhao1237	Papunesia
Austronesian	Bima-Lembata	Hawu (Sawu)	sabu1255	Papunesia
Austronesian	Bima-Lembata	Kéo	keoo1238	Papunesia
Austronesian	Bima-Lembata	*Manggarai*	mang1405	Papunesia
Austronesian	Bunun	(Takituduh) Bunun	bunu1267	Papunesia
Austronesian	Celebic	Balantak	bala1315	Papunesia
Austronesian	Celebic	Barang-Barang (Laiyolo superset)	laiy1246	Papunesia
Austronesian	Celebic	*Muna*	muna1247	Papunesia
Austronesian	Celebic	*Pendau*	pend1242	Papunesia
Austronesian	Celebic	Tajio	taji1246	Papunesia
Austronesian	Celebic	*Tukang Besi*	tuka1247, tuka1248	Papunesia
Austronesian	Celebic	*Wolio*	woli1241	Papunesia
Austronesian	Celebic	Wotu	wotu1240	Papunesia
Austronesian	Central Malayo-Polynesian	Lun Bawang (Lundayeh superset)	lund1271	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Austronesian	Central Malayo-Polynesian	Rejang Kayan	reja1241	Papunesia
Austronesian	Chamorro	Chamorro	cham1312	Papunesia
Austronesian	Creoles and Pidgins	Ambonese Malay	ambo1250	Papunesia
Austronesian	Creoles and Pidgins	Larantuka Malay	lara1260	Papunesia
Austronesian	Creoles and Pidgins	Papuan Malay	papu1250	Papunesia
Austronesian	East Formosan	*Amis*	amis1246	Papunesia
Austronesian	East Formosan	Kavalan	kava1241	Papunesia
Austronesian	Eastern Malayo-Polynesian	Big Nambas	bign1238	Papunesia
Austronesian	Eastern Malayo-Polynesian	Bugawac	buga1250	Papunesia
Austronesian	Eastern Malayo-Polynesian	Dakaka	daka1243	Papunesia
Austronesian	Eastern Malayo-Polynesian	Luang	luan1263	Papunesia
Austronesian	Eastern Malayo-Polynesian	Pingelapese	ping1243	Papunesia
Austronesian	Greater Central Philippine	Cebuano	cebu1242	Papunesia
Austronesian	Greater Central Philippine	Cotabato Manobo	cota1241	Papunesia
Austronesian	Greater Central Philippine	Maranao	mara1404	Papunesia
Austronesian	Greater Central Philippine	Masbatenyo	masb1238	Papunesia
Austronesian	Greater Central Philippine	Talaandig-Binukid	binu1244	Papunesia
Austronesian	Lampungic	*Lampung*	lamp1243	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Austronesian	Land Daiak	Mateq	kemb1249	Papunesia
Austronesian	Malayo-Chamic	Balinese	bali1278	Papunesia
Austronesian	Malayo-Chamic	*Eastern Cham*	east2563	Eurasia
Austronesian	Malayo-Chamic	Iban	iban1264	Papunesia
Austronesian	Malayo-Chamic	*Indonesian*	indo1316	Papunesia
Austronesian	Malayo-Chamic	Jambi Malay	jamb1236	Papunesia
Austronesian	Malayo-Chamic	Jarai	jara1266	Eurasia
Austronesian	Malayo-Chamic	Kupang Malay	kupa1239	Papunesia
Austronesian	Malayo-Chamic	Madurese	nucl1460	Papunesia
Austronesian	Malayo-Chamic	Manado Malay	mala1481	Papunesia
Austronesian	Malayo-Chamic	Manadonese	mala1481	Papunesia
Austronesian	Malayo-Chamic	Northern Roglai	nort2994	Eurasia
Austronesian	Malayo-Chamic	Phan Rang Chamb	east2563	Papunesia
Austronesian	Malayo-Chamic	Sasak	sasa1249	Papunesia
Austronesian	Malayo-Chamic	*Sundanese*	sund1252	Papunesia
Austronesian	Malayo-Chamic	Urak Lawoi’	urak1238	Eurasia
Austronesian	Malayo-Chamic	Western Cham	west2650	Eurasia
Austronesian	Moken-Moklen	*Moken*	moke1242	Eurasia
Austronesian	North Borneo Malayo-Polynesian	*Begak-Ida’an*	idaa1241	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Austronesian	North Borneo Malayo-Polynesian	Bonggi	bong1289	Papunesia
Austronesian	North Borneo Malayo-Polynesian	Kimarang	kima1244	Papunesia
Austronesian	North Borneo Malayo-Polynesian	*Tatana*	tata1257	Papunesia
Austronesian	North Borneo Malayo-Polynesian	Timugon Murut	timu1262	Papunesia
Austronesian	Northern Luzon	Kankanaey	kank1243	Papunesia
Austronesian	Northwest Formosan	Saisiyat	sais1237	Papunesia
Austronesian	Northwest Sumatra-Barrier Islands	*Batak (Karo)*	bata1293	Papunesia
Austronesian	Northwest Sumatra-Barrier Islands	Batak (Toba)	bata1289	Papunesia
Austronesian	Northwest Sumatra-Barrier Islands	*Gayo*	gayo1244	Papunesia
Austronesian	Northwest Sumatra-Barrier Islands	*Nias*	nias1242	Papunesia
Austronesian	Nunusaku	*Nuaulu*	sout2895	Papunesia
Austronesian	Oceanic	*'Are'are*	area1240	Papunesia
Austronesian	Oceanic	Adzera	adze1240	Papunesia
Austronesian	Oceanic	*Amara*	amar1272	Papunesia
Austronesian	Oceanic	*Ambae (Lolovoli Northeast)*	east2443	Papunesia
Austronesian	Oceanic	Anejom	anei1239	Papunesia
Austronesian	Oceanic	Apma	apma1240	Papunesia
Austronesian	Oceanic	*Araki*	arak1252	Papunesia
Austronesian	Oceanic	*Arop-Lokep*	arop1243	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Austronesian	Oceanic	*Arop-Sissano*	arop1242	Papunesia
Austronesian	Oceanic	Arosi	aros1241	Papunesia
Austronesian	Oceanic	*Avava*	katb1237	Papunesia
Austronesian	Oceanic	*Awad Bing*	awad1244	Papunesia
Austronesian	Oceanic	*Bali-Vitu*	unea1237	Papunesia
Austronesian	Oceanic	Bariai	bari1286	Papunesia
Austronesian	Oceanic	Barok	baro1253	Papunesia
Austronesian	Oceanic	Bierebo	bier1244	Papunesia
Austronesian	Oceanic	Bola	bola1250	Papunesia
Austronesian	Oceanic	*Buang (Mapos)*	mapo1242	Papunesia
Austronesian	Oceanic	Buhutu	buhu1237	Papunesia
Austronesian	Oceanic	*Buma*	tean1237	Papunesia
Austronesian	Oceanic	Cheke Holo	chek1238	Papunesia
Austronesian	Oceanic	East Uvean	wall1257	Papunesia
Austronesian	Oceanic	Erromangan (Sye)	siee1239	Papunesia
Austronesian	Oceanic	*Fijian*	fiji1243	Papunesia
Austronesian	Oceanic	Gela	gela1263	Papunesia
Austronesian	Oceanic	Gumawana	guma1254	Papunesia
Austronesian	Oceanic	*Halia*	hali1244	Papunesia
Austronesian	Oceanic	*Hawaiian*	hawa1245	Papunesia
Austronesian	Oceanic	*Hoava*	hoav1238	Papunesia
Austronesian	Oceanic	Hote	hote1245	Papunesia
Austronesian	Oceanic	*Iduna*	idun1242	Papunesia
Austronesian	Oceanic	*Ifira-Mele*	mele1250	Papunesia
Austronesian	Oceanic	*Jabêm*	yabe1254	Papunesia
Austronesian	Oceanic	*Kairiru*	kair1263	Papunesia
Austronesian	Oceanic	*Kara (in Papua New Guinea)*	kara1486	Papunesia
Austronesian	Oceanic	Karnai	karn1252	Papunesia
Austronesian	Oceanic	*Kaulong*	kaul1240	Papunesia
Austronesian	Oceanic	*Kele*	kele1258	Papunesia
Austronesian	Oceanic	Kiribati	gilb1244	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Austronesian	Oceanic	*Kokota*	koko1269	Papunesia
Austronesian	Oceanic	*Koluwawa*	kolu1245	Papunesia
Austronesian	Oceanic	*Kosraean*	kosr1238	Papunesia
Austronesian	Oceanic	Kove	kove1237	Papunesia
Austronesian	Oceanic	Kubokota (Ghanongga superset)	ghan1242	Papunesia
Austronesian	Oceanic	*Kwaio*	kwai1243	Papunesia
Austronesian	Oceanic	Kwaraqae	kwar1239	Papunesia
Austronesian	Oceanic	*Labu*	labu1248	Papunesia
Austronesian	Oceanic	Lele	lele1270	Papunesia
Austronesian	Oceanic	Lelepa	lele1267	Papunesia
Austronesian	Oceanic	Lemerig	leme1238	Papunesia
Austronesian	Oceanic	*Lenakel*	lena1238	Papunesia
Austronesian	Oceanic	Lengo	leng1259	Papunesia
Austronesian	Oceanic	*Lewo*	lewo1242	Papunesia
Austronesian	Oceanic	*Longgu*	long1395	Papunesia
Austronesian	Oceanic	*Loniu*	loni1238	Papunesia
Austronesian	Oceanic	Lote	lote1237	Papunesia
Austronesian	Oceanic	Lou	louu1245	Papunesia
Austronesian	Oceanic	*Lusi*	lusi1240	Papunesia
Austronesian	Oceanic	*Madak*	mada1285	Papunesia
Austronesian	Oceanic	*Maisin*	mais1250	Papunesia
Austronesian	Oceanic	*Maleu*	male1289	Papunesia
Austronesian	Oceanic	Malua Bay	malu1245	Papunesia
Austronesian	Oceanic	Manam	mana1295	Papunesia
Austronesian	Oceanic	Mandara	mand1440	Papunesia
Austronesian	Oceanic	Mangga Buang	mang1404	Papunesia
Austronesian	Oceanic	Mangseng	mang1403	Papunesia
Austronesian	Oceanic	Maori	maor1246	Papunesia
Austronesian	Oceanic	Marquesan (northern variety)	nort2845	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Austronesian	Oceanic	Marshallese	mars1254	Papunesia
Austronesian	Oceanic	Maskelynes	mask1242	Papunesia
Austronesian	Oceanic	Mato	mato1252	Papunesia
Austronesian	Oceanic	*Mavea*	mafe1237	Papunesia
Austronesian	Oceanic	*Mekeo*	meke1243	Papunesia
Austronesian	Oceanic	Meramera	mera1242	Papunesia
Austronesian	Oceanic	*Mokilese*	moki1238	Papunesia
Austronesian	Oceanic	Mono-Alu	mono1273	Papunesia
Austronesian	Oceanic	Motu	motu1246	Papunesia
Austronesian	Oceanic	Mussau	muss1246	Papunesia
Austronesian	Oceanic	*Mwotlap*	motl1237	Papunesia
Austronesian	Oceanic	*Nalik*	nali1244	Papunesia
Austronesian	Oceanic	Namakir	nama1268	Papunesia
Austronesian	Oceanic	Natqgu	natu1246	Papunesia
Austronesian	Oceanic	Navahaq	sout2857	Papunesia
Austronesian	Oceanic	Nehan	neha1247	Papunesia
Austronesian	Oceanic	Nese	nese1235	Papunesia
Austronesian	Oceanic	Neve'ei	vinm1237	Papunesia
Austronesian	Oceanic	Neverver	ling1265	Papunesia
Austronesian	Oceanic	Ninde	labo1244	Papunesia
Austronesian	Oceanic	*Niuean*	niue1239	Papunesia
Austronesian	Oceanic	Nochi (Notsi)	nots1237	Papunesia
Austronesian	Oceanic	Owa	owaa1237	Papunesia
Austronesian	Oceanic	Paamese	paam1238	Papunesia
Austronesian	Oceanic	Paluai	balu1257	Papunesia
Austronesian	Oceanic	*Port Sandwich*	port1285	Papunesia
Austronesian	Oceanic	Pukapuka	puka1242	Papunesia
Austronesian	Oceanic	Puluwatese	pulu1242	Papunesia
Austronesian	Oceanic	*Ramoaaina*	ramo1244	Papunesia
Austronesian	Oceanic	*Rapanui*	rapa1244	Papunesia
Austronesian	Oceanic	*Roro (Waima superset)*	waim1251	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Austronesian	Oceanic	*Rotuman*	rotu1241	Papunesia
Austronesian	Oceanic	*Roviana*	rovi1238	Papunesia
Austronesian	Oceanic	Sa'a	saaa1240	Papunesia
Austronesian	Oceanic	Sakao	saka1289	Papunesia
Austronesian	Oceanic	Satawalese	sata1237	Papunesia
Austronesian	Oceanic	*Seimat*	seim1238	Papunesia
Austronesian	Oceanic	Siar (Siar-Lak)	siar1238	Papunesia
Austronesian	Oceanic	Sie	siee1239	Papunesia
Austronesian	Oceanic	Sikaiana	sika1261	Papunesia
Austronesian	Oceanic	*Sinaugoro*	sina1266	Papunesia
Austronesian	Oceanic	*Sio*	sioo1240	Papunesia
Austronesian	Oceanic	*Sisiqa*	sisi1250	Papunesia
Austronesian	Oceanic	Sobei	sobe1238	Papunesia
Austronesian	Oceanic	South Efate	sout2856	Papunesia
Austronesian	Oceanic	Southwest Tanna	sout2869	Papunesia
Austronesian	Oceanic	*Sudest*	sude1239	Papunesia
Austronesian	Oceanic	Sunwadia	mari1426	Papunesia
Austronesian	Oceanic	Sursurunga	surs1246	Papunesia
Austronesian	Oceanic	Tahitian	tahi1242	Papunesia
Austronesian	Oceanic	*Taiof (Saposa superset)*	sapo1253	Papunesia
Austronesian	Oceanic	*Takia*	taki1248	Papunesia
Austronesian	Oceanic	*Tama(m)bo*	malo1243	Papunesia
Austronesian	Oceanic	Tami	tami1290	Papunesia
Austronesian	Oceanic	*Tawala*	tawa1275	Papunesia
Austronesian	Oceanic	*Tigak*	tiga1245	Papunesia
Austronesian	Oceanic	Tinrin	tiri1258	Papunesia
Austronesian	Oceanic	Tirax	maee1241	Papunesia
Austronesian	Oceanic	Tokelau	toke1240	Papunesia
Austronesian	Oceanic	Toqabaita	toab1237	Papunesia
Austronesian	Oceanic	*Tungak*	tung1290	Papunesia
Austronesian	Oceanic	Tuvaluan	tuva1244	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Austronesian	Oceanic	Ubir	ubir1237	Papunesia
Austronesian	Oceanic	Ughele	ughe1237	Papunesia
Austronesian	Oceanic	Ulithian	ulit1238	Papunesia
Austronesian	Oceanic	*Unua*	unua1237	Papunesia
Austronesian	Oceanic	Vaeakau-Taumako	pile1238	Papunesia
Austronesian	Oceanic	Vera'a	vera1241	Papunesia
Austronesian	Oceanic	*Vinmavis*	vinm1237	Papunesia
Austronesian	Oceanic	*Wala*	wala1266	Papunesia
Austronesian	Oceanic	Whitesands	whit1269	Papunesia
Austronesian	Oceanic	*Woleaian*	wole1240	Papunesia
Austronesian	Oceanic	*Wuvulu-Aua*	wuvu1239	Papunesia
Austronesian	Oceanic	*Xârâcùù*	xara1244	Papunesia
Austronesian	Rukai	*Rukai*	ruka1240	Papunesia
Austronesian	Sangiric	Ratahan	rata1244	Papunesia
Austronesian	South Halmahera - West New Guinea	*Ambai*	amba1265	Papunesia
Austronesian	South Halmahera - West New Guinea	*Biak*	biak1248	Papunesia
Austronesian	South Halmahera - West New Guinea	Dusner	dusn1237	Papunesia
Austronesian	South Halmahera - West New Guinea	*Irarutu*	irar1238	Papunesia
Austronesian	South Halmahera - West New Guinea	Matbat	matb1237	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Austronesian	South Halmahera - West New Guinea	*Mor*	morm1235	Papunesia
Austronesian	South Halmahera - West New Guinea	*Taba*	east2440	Papunesia
Austronesian	South Halmahera - West New Guinea	Wamesa	wand1267	Papunesia
Austronesian	South Halmahera - West New Guinea	*Warembori*	ware1253	Papunesia
Austronesian	South Halmahera - West New Guinea	Waropen	waro1242	Papunesia
Austronesian	South Halmahera - West New Guinea	Wooi	woii1237	Papunesia
Austronesian	South Sulawesi	Bugis (Buginese variety)	bugi1244	Papunesia
Austronesian	South Sulawesi	*Selayar*	sela1260	Papunesia
Austronesian	Timor-Babar	*Leti*	leti1246	Papunesia
Austronesian	Timor-Babar	Maku'a	maku1277	Papunesia
Austronesian	Timor-Babar	*Selaru*	sela1259	Papunesia
Austronesian	Timor-Babar	*Tugun*	tugu1245	Papunesia
Austronesian	Tsouic	Tsou	tsou1248	Papunesia
Austronesian	West Central Maluku	*Buru*	buru1303	Papunesia
Aymaran	Aymaran	Aymara (Southern)	sout2996	South America
Aymaran	Aymaran	*Jaqaru*	jaqa1244	South America

Family	Genus	Language	Glottocode	Macroarea
Baibai-Fas	Baibai-Fas	Momu-Fas	fass1245	Papunesia
Baining	Baining	Kaket (Qaqet)	qaqe1238	Papunesia
Barbacoan	Barbacoan	*Awa Pit*	awac1239	South America
Barbacoan	Barbacoan	Cayapa	chac1249	South America
Barbacoan	Barbacoan	Tsafiki	colo1256	South America
Basque	Basque	*Basque*	basq1248	Eurasia
Berta	Berta	Berta	bert1248	Africa
Blue Nile Mao	Blue Nile Mao	Hoozo	hozo1236	Africa
Bosavi	Bosavi	*Edolo*	edol1239	Papunesia
Bosavi	Bosavi	Kaluli	kalu1248	Papunesia
Burushaski	Burushaski	*Burushaski*	buru1296	Eurasia
Cariban	Guianan	Carib (De'kwana)	maqu1238	South America
Cariban	Guianan	Carib (Galibi)	gali1262	South America
Cariban	Guianan	Tiriyo	trio1238	South America
Cayubaba	Cayubaba	Cayuvava	cayu1262	South America
Central Sudanic	Lenduic	Lendu	lend1245	Africa
Central Sudanic	Moru-Madi	*Logoti*	logo1259	Africa
Central Sudanic	Sara-Bongo-Bagirmi	Bongo	bong1285	Africa
Central Sudanic	Sara-Bongo-Bagirmi	Kenga	keng1240	Africa
Central Sudanic	Sara-Bongo-Bagirmi	Mbay	mbay1241	Africa
Chibchan	Core Chibchan	(Costa Rica) Teribe	teri1250	North America
Chibchan	Core Chibchan	*Boruca*	boru1252	North America
Chibchan	Core Chibchan	*Bribri*	brib1243	North America

Family	Genus	Language	Glottocode	Macroarea
Chibchan	Core Chibchan	Buglere	bugl1243	North America
Chibchan	Core Chibchan	Cabécar	cabe1245	North America
Chibchan	Core Chibchan	Cuna (Border)	bord1248	South America
Chibchan	Core Chibchan	Cuna (San Blas varieties)	sanb1242	South America
Chibchan	Core Chibchan	Kogui	cogu1240	South America
Chibchan	Core Chibchan	*Muisca*	chib1270	South America
Chibchan	Core Chibchan	*Ngäbere*	ngab1239	North America
Chibchan	Core Chibchan	Panama (Naso)	teri1250	North America
Chibchan	Pech	Pech (Pesh)	pech1241	North America
Chimakuan	Chimakuan	Quileute	quil1240	North America
Chimariko	Chimariko	Chimariko	chim1301	North America
Choco	Choco	*Emberá (Northern)*	nort2972	South America
Chukotko-Kamchatkan	Chukotian	Chukchi	chuk1273	Eurasia
Cochimi-Yuman	Yuman	Hualapai	hava1248	North America
Dagan	Dagan	*Daga*	daga1275	Papunesia
Dizoid	Dizoid	Dizin	dizi1235	Africa
Dizoid	Dizoid	Sheko	shek1245	Africa
Dogon	Dogon	Ben Tey	bent1238	Africa
Dogon	Dogon	*Jamsay*	jams1239	Africa
Dogon	Dogon	*Tommo So*	tomm1242	Africa
Dravidian	Central Dravidian	Kolami	nort2699	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Dravidian	North Dravidian	Malto	kuma1274	Eurasia
Dravidian	South Dravidian	Dandami Maria	dand1238	Eurasia
Dravidian	South Dravidian	Gondi	nort2702	Eurasia
Dravidian	South Dravidian	*Kannada*	nucl1305	Eurasia
Dravidian	South Dravidian	Koya	koya1251	Eurasia
Dravidian	South Dravidian	*Kuvi*	kuvi1243	Eurasia
Dravidian	South Dravidian	*Malayalam*	mala1464	Eurasia
Dravidian	South Dravidian	*Tamil*	tami1289	Eurasia
Dravidian	South Dravidian	Telugu	telu1262	Eurasia
Dravidian	South Dravidian	Tulu	tulu1258	Eurasia
Dravidian	South Dravidian	Urali Irula	irul1244	Eurasia
Duna	Duna	*Duna*	duna1248	Papunesia
East Bird's Head	East Bird's Head	*Meyah*	meya1236	Papunesia
East Bird's Head	East Bird's Head	Sougb	mani1235	Papunesia
East Strickland	East Strickland	Konai	kona1242	Papunesia
Eleman	Eleman	Orokolo	orok1267	Papunesia
Eleman	Eleman	Tairuma (Toaripi)	toar1246	Papunesia
Eskimo-Aleut	Eskimo	Central Alaskan Yupik	cent2127	North America
Fulniô	Fulniô	*Fulniô*	fuln1247	South America
Furan	Furan	*Fur*	furr1244	Africa
Fuyug	Fuyug	Fuyug	fuyu1242	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Guató	Guató	Guató	guat1253	South America
Gumuz	Gumuz	Northern Gumuz	gumu1244	Africa
Gumuz	Gumuz	Southern Gumuz	sout3236	Africa
Gunwinyguan	Gunwinyguan	Gunbalang	kunb1251	Australia
Haida	Haida	*Haida*	nort2938, sout2956	North America
Hatam-Mansim	Hatam-Mansim	*Hatam*	hata1243	Papunesia
Heibanic	Heibanic	*Moro*	moro1285	Africa
Hibito-Cholon	Hibito-Cholon	Cholón	chol1284	South America
Hmong-Mien	Hmongic	Bunu (Younuo)	youn1235	Eurasia
Hmong-Mien	Hmongic	*Hmong Njua*	hmon1264	Eurasia
Hmong-Mien	Hmongic	She	shee1238	Eurasia
Hmong-Mien	Mienic	Mien	iumi1238	Eurasia
Hruso	Hruso	Hruso	hrus1242	Eurasia
Huitotoan	Huitotoan	Huitoto (Minica)	mini1256	South America
Ijoid	Ijoid	Defaka	defa1248	Africa
Ijoid	Ijoid	Ijo (Kolokuma)	izon1238	Africa
Ijoid	Ijoid	Nkoroo	nkor1239	Africa
Indo-European	Albanian	Albanian	tosk1239, gheg1238	Eurasia
Indo-European	Armenic	*Armenian (Eastern)*	nucl1235	Eurasia
Indo-European	Armenic	Armenian (Western)	homs1234	Eurasia
Indo-European	Balto-Slavic	*Bosnian-Serbian-Croatian*	sout1528	Eurasia
Indo-European	Balto-Slavic	*Bulgarian*	bulg1262	Eurasia
Indo-European	Balto-Slavic	Czech	czec1258	Eurasia
Indo-European	Balto-Slavic	Latvian	latv1249	Eurasia
Indo-European	Balto-Slavic	Macedonian	mace1250	Eurasia
Indo-European	Balto-Slavic	*Polish*	poli1260	Eurasia
Indo-European	Balto-Slavic	Slovenian	slov1268	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Indo-European	Balto-Slavic	Ukrainian	ukra1253	Eurasia
Indo-European	Celtic	*Breton*	bret1244	Eurasia
Indo-European	Celtic	Cornish	corn1251	Eurasia
Indo-European	Celtic	Gaelic (Scots)	scot1245	Eurasia
Indo-European	Celtic	Irish	iris1253	Eurasia
Indo-European	Celtic	*Welsh*	wels1247	Eurasia
Indo-European	Creoles and Pidgins	Berbice Dutch Creole	berb1259	South America
Indo-European	Creoles and Pidgins	Bislama	bisl1239	Papunesia
Indo-European	Creoles and Pidgins	Haitian	hait1244	North America
Indo-European	Creoles and Pidgins	Ho	sara1340	South America
Indo-European	Creoles and Pidgins	*Ndyuka*	suri1272	South America
Indo-European	Creoles and Pidgins	Tok Pisin	tokp1240	Papunesia
Indo-European	Germanic	Danish	dani1285	Eurasia
Indo-European	Germanic	Dutch	dutc1256	Eurasia
Indo-European	Germanic	*English*	stan1293	Eurasia
Indo-European	Germanic	Faroese	faro1244	Eurasia
Indo-European	Germanic	German	stan1295	Eurasia
Indo-European	Germanic	Icelandic	icel1247	Eurasia
Indo-European	Germanic	*Norwegian*	norw1258	Eurasia
Indo-European	Germanic	Swedish	swed1254	Eurasia
Indo-European	Graeco-Phrygian	*Greek (Modern)*	mode1248	Eurasia
Indo-European	Indo-Aryan	Bhojpuri	bhoj1244	Eurasia
Indo-European	Indo-Aryan	*Brokskat*	brok1247	Eurasia
Indo-European	Indo-Aryan	Bundeli	bund1253	Eurasia
Indo-European	Indo-Aryan	Darai	dara1250	Eurasia
Indo-European	Indo-Aryan	*Gujarati*	guja1252	Eurasia
Indo-European	Indo-Aryan	Hindi	hind1269	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Indo-European	Indo-Aryan	*Kashmiri*	kash1277	Eurasia
Indo-European	Indo-Aryan	*Konkani*	konk1267	Eurasia
Indo-European	Indo-Aryan	*Kumauni*	kuma1273	Eurasia
Indo-European	Indo-Aryan	Lamani	lamb1269	Eurasia
Indo-European	Indo-Aryan	Magahi	maga1260	Eurasia
Indo-European	Indo-Aryan	Maithili	mait1250	Eurasia
Indo-European	Indo-Aryan	*Marathi*	mara1378	Eurasia
Indo-European	Indo-Aryan	*Nepali*	nepa1254	Eurasia
Indo-European	Indo-Aryan	*Oriya*	oriy1255	Eurasia
Indo-European	Indo-Aryan	*Panjabi*	panj1256	Eurasia
Indo-European	Indo-Aryan	Shina	shin1264	Eurasia
Indo-European	Indo-Aryan	*Torwali*	torw1241	Eurasia
Indo-European	Indo-Aryan	*Urdu*	urdu1245	Eurasia
Indo-European	Iranian	Kumzari	kumz1235	Eurasia
Indo-European	Iranian	*Pashto*	cent1973	Eurasia
Indo-European	Iranian	Persian	west2369	Eurasia
Indo-European	Iranian	Tajik	taji1245	Eurasia
Indo-European	Iranian	*Wakhi*	wakh1245	Eurasia
Indo-European	Iranian	*Zazaki*	diml1238	Eurasia
Indo-European	Latinic	Latin	lati1261	Eurasia
Indo-European	Nuristani	Prasuni	pras1239	Eurasia
Indo-European	Romance	*Catalan*	stan1289	Eurasia
Indo-European	Romance	*French*	stan1290	Eurasia
Indo-European	Romance	Italian	ital1282	Eurasia
Indo-European	Romance	Portuguese	port1283	Eurasia
Indo-European	Romance	Romansch (Sursilvan)	roma1326	Eurasia
Indo-European	Romance	*Spanish*	stan1288	Eurasia
Iwaidjan Proper	Iwaidjan Proper	Iwaidja	iwai1244	Australia
Iwaidjan Proper	Iwaidjan Proper	Maung	maun1240	Australia
Japonic	Japonic	*Japanese*	nucl1643	Eurasia
Japonic	Japonic	Yuwan	sout2954	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Kadugli-Krongo	Kadugli-Krongo	Deiga	keig1242	Africa
Kaki Ae	Kaki Ae	*Kaki Ae*	kaki1249	Papunesia
Kamula-Elevala	Kamula-Elevala	Kamula	kamu1260	Papunesia
Kanoê	Kanoê	Kanoê	kano1245	South America
Kariri	Kariri	Kipea	kari1255	South America
Karok	Karok	Karok	karo1304	North America
Kartvelian	Kartvelian	Laz	lazz1240	Eurasia
Keresan	Keresan	Laguna Keres (Western Keres superset)	lagu1251 (west2632)	North America
Khoe-Kwadi	Khoe-Kwadi	*Nama*	nama1264	Africa
Kiowa-Tanoan	Kiowa-Tanoan	Kiowa	kiow1266	North America
Koiarian	Koiarian	*Barai*	nucl1630	Papunesia
Koiarian	Koiarian	*Koiali (Mountain)*	moun1252	Papunesia
Koiarian	Koiarian	*Koiari*	gras1249	Papunesia
Koiarian	Koiarian	*Koita*	koit1244	Papunesia
Koiarian	Koiarian	*Ömie*	omie1241	Papunesia
Koman	Koman	*Gwama*	kwam1249	Africa
Koman	Koman	Uduk	uduk1239	Africa
Korean	Korean	*Korean*	kore1280	Eurasia
Kuliak	Kuliak	*Ik*	ikkk1242	Africa
Kuliak	Kuliak	*So*	sooo1256	Africa
Kunama	Kunama	*Kunama*	kuna1268	Africa
Kunimaipan	Kunimaipan	Amam	amam1246	Papunesia
Kunimaipan	South East Papuan	Biangai	bian1252	Papunesia
Kunimaipan	South East Papuan	Weri	weri1253	Papunesia
Kuot	Kuot	*Kuot*	kuot1243	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Kusunda	Kusunda	Kusunda	kusu1250	Eurasia
Kwomtari-Nai	Kwomtari-Nai	Kwomtari	nucl1593	Papunesia
Kxa	Kxa	ǂHòã	hoaa1235	Africa
Laal	Laal	Laal	laal1242	Africa
Laragia	Laragia	Laragia	lara1258	Australia
Lavukaleve	Lavukaleve	Lavukaleve	lavu1241	Papunesia
Leco	Leco	Leco	leco1242	South America
Left May	Left May	Ama (Papa New Guinea)	amap1240	Papunesia
Lower Sepik-Ramu	Lower Sepik	Yimas	yima1243	Papunesia
Lower Sepik-Ramu	Ramu	Awar	awar1249	Papunesia
Maban	Maban	*Maba*	maba1277	Africa
Mairasic	Mairasic	*Mairasi*	nucl1594	Papunesia
Mande	Eastern Mande	*Busa*	busa1253	Africa
Mande	Eastern Mande	*Dan*	dann1241	Africa
Mande	Southeastern Mande	*Mona*	mwan1250	Africa
Mande	Southeastern Mande	Wan	wann1242	Africa
Mande	Western Mande	Bambara	bamb1269	Africa
Mande	Western Mande	Jalkunan	jalk1242	Africa
Mande	Western Mande	*Mandinka (Gambian)*	mand1436	Africa
Mande	Western Mande	Soninke	soni1259	Africa
Mande	Western Mande	*Vai*	vaii1241	Africa
Mande	Western Mande	*Xasonga*	xaas1235	Africa
Manubaran	Manubaran	Doromu-Koki	doro1266	Papunesia
Matacoan	Matacoan	Wichí	wich1264	South America
Mayan	Core Mayan	Cakchiquel	kaqc1270	North America

Family	Genus	Language	Glottocode	Macroarea
Mayan	Core Mayan	Chuj	chuj1250	North America
Mayan	Core Mayan	*Huastec*	huas1242	North America
Mayan	Core Mayan	Itzá	itza1241	North America
Mayan	Core Mayan	*Jakaltek*	kanj1263	North America
Mayan	Core Mayan	*Mam*	mamm1241	North America
Mayan	Core Mayan	Mayan Quiché	kich1262	North America
Mayan	Core Mayan	Popti’	popt1235	North America
Mayan	Core Mayan	Q’anjob’al	qanj1241	North America
Mayan	Core Mayan	*Sipakapense*	sipa1247	North America
Maybrat-Karon	Maybrat-Karon	*Maybrat*	maib1239	Papunesia
Misumalpan	Misumalpan	Miskito	misk1235	North America
Misumalpan	Misumalpan	Ulwa	ulwa1239	North America
Mixe-Zoque	Mixe-Zoque	Zoque (Chimalapa)	chim1300	North America
Mongolic-Khitan	Mongolic	Buriat	mong1330	Eurasia
Mongolic-Khitan	Mongolic	*Dagur*	daur1238	Eurasia
Mongolic-Khitan	Mongolic	Khalkha	halh1238	Eurasia
Mongolic-Khitan	Mongolic	*Mangghuer*	huzh1238	Eurasia
Mongolic-Khitan	Mongolic	Oirad-Kalmyk-Darkhat	kalm1243	Eurasia
Mpur	Mpur	Mpur	mpur1239	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Muskogean	Muskogean	*Chickasaw*	chic1270	North America
Muskogean	Muskogean	*Choctaw*	choc1276	North America
Muskogean	Muskogean	*Creek*	cree1270	North America
Naduhup	Naduhup	*Dâw*	daww1239	South America
Naduhup	Naduhup	*Hup*	hupd1244	South America
Nakh-Daghestanian	Daghestanian	*Archi*	arch1244	Eurasia
Nakh-Daghestanian	Daghestanian	Avar	avar1256	Eurasia
Nakh-Daghestanian	Daghestanian	Bagvalal	bagv1239	Eurasia
Nakh-Daghestanian	Daghestanian	Bezhta	bezh1248	Eurasia
Nakh-Daghestanian	Daghestanian	*Godoberi*	ghod1238	Eurasia
Nakh-Daghestanian	Daghestanian	Hinuq	hinu1240	Eurasia
Nakh-Daghestanian	Daghestanian	*Hunzib*	hunz1247	Eurasia
Nakh-Daghestanian	Daghestanian	Khwarshi-Inhoqwari	khva1239	Eurasia
Nakh-Daghestanian	Daghestanian	Kryz	kryt1240	Eurasia
Nakh-Daghestanian	Daghestanian	Lak	lakk1252	Eurasia
Nakh-Daghestanian	Daghestanian	Lezgian	lezg1247	Eurasia
Nakh-Daghestanian	Daghestanian	Rutul	rutu1240	Eurasia
Nakh-Daghestanian	Daghestanian	*Tsez*	dido1241	Eurasia
Nakh-Daghestanian	Daghestanian	Udi	udii1243	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Nakh-Daghestanian	Nakh	Tsova-Tush	bats1242	Eurasia
Narrow Talodi	Narrow Talodi	Asheron	ache1245	Africa
Ndu	Ndu	*Ambulas*	ambu1247	Papunesia
Ndu	Ndu	Hanga Hundi	hang1263	Papunesia
Ndu	Ndu	*Iatmul*	iatm1242	Papunesia
Ndu	Ndu	Manambu	mana1298	Papunesia
Nilotic	Eastern Nilotic	Arusa (Maasai superset)	masa1300	Africa
Nilotic	Eastern Nilotic	*Bari*	bari1284	Africa
Nilotic	Eastern Nilotic	*Turkana*	turk1308	Africa
Nilotic	Southern Nilotic	Kipsigis	kips1239	Africa
Nilotic	Western Nilotic	Belanda Bor	bela1256	Africa
Nilotic	Western Nilotic	Labwor	labw1238	Africa
Nilotic	Western Nilotic	*Lango*	lang1324	Africa
Nilotic	Western Nilotic	*Luo*	luok1236	Africa
Nilotic	Western Nilotic	Shilluk	shil1265	Africa
Nimboranic	Nimboranic	*Nimboran*	nucl1633	Papunesia
Nivkh	Nivkh	Nivkh	gily1242	Eurasia
North Halmahera	North Halmahera	Ternate	tern1247	Papunesia
North Halmahera	North Halmahera	*Tidore*	tido1248	Papunesia
North Halmahera	North Halmahera	Tobelo	tobe1252	Papunesia
North Halmahera	North Halmahera	West Makian	west2600	Papunesia
Nubian	Nubian	*Nubian (Dongolese)*	dong1288 (kenu1236)	Africa
Nubian	Nubian	*Nubian (Kunuz)*	kenu1243 (kenu1236)	Africa
Nuclear Macro-Je	Nuclear Macro-Je	*Ofayé*	ofay1240	South America

Family	Genus	Language	Glottocode	Macroarea
Nuclear Torricelli	Kombio-Arapesh-Urat	*Arapesh (Abu)*	abua1245	Papunesia
Nuclear Torricelli	Kombio-Arapesh-Urat	Arapesh (Mountain)	buki1249	Papunesia
Nuclear Torricelli	Marienberg	*Kamasau*	kama1367	Papunesia
Nuclear Torricelli	Urim	*Urim*	urim1252	Papunesia
Nuclear Torricelli	Wapei-Palei	*Au*	auuu1241	Papunesia
Nuclear Torricelli	Wapei-Palei	Olo	oloo1241	Papunesia
Nuclear Torricelli	Wapei-Palei	*Walman*	valm1241	Papunesia
Nuclear Torricelli	Wapei-Palei	*Yeri*	yapu1240	Papunesia
Nuclear Trans New Guinea	Asmat-Awyu-Ok	*Bimin*	bimi1240	Papunesia
Nuclear Trans New Guinea	Asmat-Awyu-Ok	*Mian*	mian1256	Papunesia
Nuclear Trans New Guinea	Asmat-Awyu-Ok	Oksapmin	oksa1245	Papunesia
Nuclear Trans New Guinea	Asmat-Awyu-Ok	Telefol	tele1256	Papunesia
Nuclear Trans New Guinea	Asmat-Awyu-Ok	Wambon	wamb1259	Papunesia
Nuclear Trans New Guinea	Asmat-Awyu-Ok	Yongkom	yong1280	Papunesia
Nuclear Trans New Guinea	Chimbu-Wahgi	Bo Ung (Ku Waru dialect)	boun1245	Papunesia
Nuclear Trans New Guinea	Chimbu-Wahgi	Dom	domm1246	Papunesia
Nuclear Trans New Guinea	Chimbu-Wahgi	*Golin*	goli1247	Papunesia
Nuclear Trans New Guinea	Chimbu-Wahgi	*Kuman*	kuma1280	Papunesia
Nuclear Trans New Guinea	Chimbu-Wahgi	Umbu-Ungu	umbu1258	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Nuclear Trans New Guinea	Dani	*Nggem*	ngge1241	Papunesia
Nuclear Trans New Guinea	Enga-Kewa-Huli	*Huli*	huli1244	Papunesia
Nuclear Trans New Guinea	Enga-Kewa-Huli	*Kewa (both East and West)*	east2516, west2599	Papunesia
Nuclear Trans New Guinea	Enga-Kewa-Huli	*Kyaka*	kyak1244	Papunesia
Nuclear Trans New Guinea	Finisterre-Huon	Awara	awar1248	Papunesia
Nuclear Trans New Guinea	Finisterre-Huon	*Iyo*	iyoo1238	Papunesia
Nuclear Trans New Guinea	Finisterre-Huon	Kovai	kova1243	Papunesia
Nuclear Trans New Guinea	Finisterre-Huon	Ma Manda	sauk1252	Papunesia
Nuclear Trans New Guinea	Finisterre-Huon	*Nabak*	naba1256	Papunesia
Nuclear Trans New Guinea	Finisterre-Huon	*Nankina*	nank1250	Papunesia
Nuclear Trans New Guinea	Finisterre-Huon	Nukna	nukn1238	Papunesia
Nuclear Trans New Guinea	Finisterre-Huon	Rawa (Karo)	rawa1267	Papunesia
Nuclear Trans New Guinea	Finisterre-Huon	*Selepet*	sele1250	Papunesia
Nuclear Trans New Guinea	Greater Binanderean	*Binandere*	bina1277	Papunesia
Nuclear Trans New Guinea	Greater Binanderean	*Guhu-Samane*	guhu1244	Papunesia
Nuclear Trans New Guinea	Greater Binanderean	*Korafe-Yegha*	kora1294	Papunesia
Nuclear Trans New Guinea	Greater Binanderean	*Orokaiva*	orok1269	Papunesia
Nuclear Trans New Guinea	Greater Binanderean	Suena	suen1241	Papunesia
Nuclear Trans New Guinea	Kainantu-Goroka	*Awa*	awap1236	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Nuclear Trans New Guinea	Kainantu-Goroka	*Fore*	fore1270	Papunesia
Nuclear Trans New Guinea	Kainantu-Goroka	*Hua*	huaa1250	Papunesia
Nuclear Trans New Guinea	Kainantu-Goroka	Kamano	kama1370	Papunesia
Nuclear Trans New Guinea	Kainantu-Goroka	Tairora	nort2920	Papunesia
Nuclear Trans New Guinea	Kainantu-Goroka	Usarufa	usar1243	Papunesia
Nuclear Trans New Guinea	Kainantu-Goroka	*Yagaria*	yaga1260 (huaa1250)	Papunesia
Nuclear Trans New Guinea	Madang	*Bargam*	barg1252	Papunesia
Nuclear Trans New Guinea	Madang	*Girawa*	gira1247	Papunesia
Nuclear Trans New Guinea	Madang	*Kobon*	kobo1249	Papunesia
Nuclear Trans New Guinea	Madang	Koromu	kesa1244	Papunesia
Nuclear Trans New Guinea	Madang	*Maia*	maia1254	Papunesia
Nuclear Trans New Guinea	Madang	*Mauwake*	mauw1238	Papunesia
Nuclear Trans New Guinea	Madang	Nend	nend1239	Papunesia
Nuclear Trans New Guinea	Madang	*Siroi*	siro1249	Papunesia
Nuclear Trans New Guinea	Madang	*Tauya*	tauy1241	Papunesia
Nuclear Trans New Guinea	Madang	*Usan*	usan1239	Papunesia
Nuclear Trans New Guinea	Madang	*Waskia*	wask1241	Papunesia
Nuclear Trans New Guinea	Mek	*Una*	unaa1239	Papunesia
Nuclear Trans New Guinea	Mek	*Yale (Kosarek)*	kosa1249	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Nyulyulan	Nyulyulan	Nyulnyul	nyul1247	Australia
Ongota	Ongota	Ongota	bira1253	Africa
Otomanguean	Eastern Otomanguean	Alacatlatzala Mixtec	alac1244	North America
Otomanguean	Eastern Otomanguean	*Amuzgo (Guerrero variety)*	guer1243	North America
Otomanguean	Eastern Otomanguean	*Chatino (Sierra Occidental)*	west2644	North America
Otomanguean	Eastern Otomanguean	Chatino (Yaitepec)	west2644	North America
Otomanguean	Eastern Otomanguean	Magdalena Peñasco Mixtec	magd1235	North America
Otomanguean	Eastern Otomanguean	Me'phaa (Malinaltepec variety)	mali1285	North America
Otomanguean	Eastern Otomanguean	Mixtec (Chalcatongo)	sanm1296	North America
Otomanguean	Eastern Otomanguean	*Mixtec (Jicaltepec)*	pino1237	North America
Otomanguean	Eastern Otomanguean	Mixtec (Ocotepec)	ocot1243	North America
Otomanguean	Eastern Otomanguean	*Mixtec (Yosondúa)*	yoso1239	North America
Otomanguean	Eastern Otomanguean	Mixtepec Zapotec (San Juan and San Pedro)	mixt1426	North America
Otomanguean	Eastern Otomanguean	Quiegolani Zapotec	sant1451	North America
Otomanguean	Eastern Otomanguean	Sierra de Juárez Zapotec	sier1250	North America
Otomanguean	Eastern Otomanguean	Tacuate de Santa María Zacatepec	sant1436	North America
Otomanguean	Eastern Otomanguean	Tataltepec Chatino	tata1258	North America
Otomanguean	Eastern Otomanguean	Teotepec Chatino	nopa1235	North America
Otomanguean	Eastern Otomanguean	Tezoatlán Mixtec	tezo1238	North America

Family	Genus	Language	Glottocode	Macroarea
Otomanguean	Eastern Otomanguean	Tlacolula Valley Zapotec (Western)	sanj1284	North America
Otomanguean	Eastern Otomanguean	*Trique (Copala)*	copa1237	North America
Otomanguean	Eastern Otomanguean	Yatzachi Zapotec	yatz1235	North America
Otomanguean	Eastern Otomanguean	Zacatepec Chatino	zaca1242	North America
Otomanguean	Eastern Otomanguean	*Zapotec (Isthmus)*	isth1244	North America
Otomanguean	Eastern Otomanguean	Zapotec (Mitla)	mitl1236	North America
Otomanguean	Eastern Otomanguean	*Zapotec (Zoogocho)*	zoog1238	North America
Otomanguean	Western Otomanguean	*Chinantec (Comaltepec)*	coma1246	North America
Otomanguean	Western Otomanguean	*Chinantec (Lealao)*	leal1235	North America
Otomanguean	Western Otomanguean	Chinantec (Palantla)	pala1351	North America
Otomanguean	Western Otomanguean	Ocuilteco	atzi1235	North America
Otomanguean	Western Otomanguean	*Pame (Southern)*	sout2999	North America
Otomanguean	Western Otomanguean	Quiotepec Chinantec	quio1240	North America
Otomanguean	Western Otomanguean	Sochiapam Chinantec	soch1239	North America
Otomanguean	Western Otomanguean	Tlapanec	acat1239	North America
Páez	Páez	*Páez*	paez1247	South America
Pama-Nyungan	Arandic-Thura-Yura	*Arrernte (Mparntwe)*	east2379	Australia
Pama-Nyungan	Desert Nyungic	*Ngaanyatjarra*	ngaa1240	Australia
Pama-Nyungan	Greater Maric	Warrongo	waru1264	Australia
Pama-Nyungan	Kalkatungic	Yalarnnga	yala1262	Australia

Family	Genus	Language	Glottocode	Macroarea
Pama-Nyungan	Paman	*Kugu Nganhcara*	kuku1280	Australia
Pama-Nyungan	Paman	Wik-Mungkan	wikm1247	Australia
Pama-Nyungan	South-West Pama-Nyungan	Biri	biri1256	Australia
Pama-Nyungan	Western Pama-Nyungan	Martuthunira	mart1255	Australia
Pama-Nyungan	Yimidhirr-Yalanji-Yidinic	Gugu-Yalanji	kuku1273	Australia
Pama-Nyungan	Yimidhirr-Yalanji-Yidinic	Guugu-Yimidjir	gugu1255	Australia
Pama-Nyungan	Yimidhirr-Yalanji-Yidinic	Yidiny	yidi1250	Australia
Pano-Tacanan	Panoan	Amahuaca	amah1246	South America
Pano-Tacanan	Panoan	Shipibo-Konibo	ship1254	South America
Pano-Tacanan	Panoan	Yaminawa	yami1256	South America
Pano-Tacanan	Tacanan	*Araona*	arao1248	South America
Pano-Tacanan	Tacanan	Ese Ejja	esee1248	South America
Pawaia	Pawaia	*Pawaia*	pawa1255	Papunesia
Peba-Yagua	Peba-Yagua	*Yagua*	yagu1244	South America
Quechuan	Quechuan	Ayacucho Quechua	ayac1239	South America
Quechuan	Quechuan	Cusco Quechua	cusc1236	South America
Quechuan	Quechuan	*Quechua (Huallaga)*	hual1241	South America
Quechuan	Quechuan	Quechua (Imbabura)	imba1240	South America
Quechuan	Quechuan	San Martín Quechua	sanm1289	South America
Sahaptian	Sahaptian	Nez Perce	nezp1238	North America

Family	Genus	Language	Glottocode	Macroarea
Saharan	Saharan	Beria	zagh1240	Africa
Saharan	Saharan	Dazaga	daza1242	Africa
Saharan	Saharan	*Kanuri*	cent2050	Africa
Salishan	Coast Salish	Halkomelem (Upriver)	halk1245	North America
Salishan	Coast Salish	*Musqueam*	halk1245, musq1240	North America
Salishan	Coast Salish	Squamish	squa1248	North America
Sandawe	Sandawe	*Sandawe*	sand1273	Africa
Savosavo	Savosavo	*Savosavo*	savo1255	Papunesia
Senagi	Senagi	Menggwa Dla	dera1245	Papunesia
Sentanic	Sentanic	Sentani	nucl1632	Papunesia
Sepik	Abau	*Abau*	abau1245	Papunesia
Sepik	Iwam-Wogamus	Iwam	iwam1256	Papunesia
Sepik	Nukuma	Kwoma	kwom1262	Papunesia
Sepik	Nukuma	Mende (Papua New Guinea)	mend1268	Papunesia
Sepik	Ram	*Awtuw*	awtu1239	Papunesia
Sepik	Sepik Hill	*Alamblak*	alam1246	Papunesia
Sepik	Sepik Hill	*Sare*	kapr1245	Papunesia
Sepik	Sepik Tama	Yessan-Mayo	yess1239	Papunesia
Sepik	Yellow River	*Namia*	nami1256	Papunesia
Shabo	Shabo	Shabo	shab1252	Africa
Sino-Tibetan	Bodic	(Dongwang) Tibetan	kham1282	Eurasia
Sino-Tibetan	Bodic	(Lamjung) Yo(h)lmo	lamj1247	Eurasia
Sino-Tibetan	Bodic	Balti	balt1258	Eurasia
Sino-Tibetan	Bodic	Chaudangsi	chau1259	Eurasia
Sino-Tibetan	Bodic	*Darma*	darm1243	Eurasia
Sino-Tibetan	Bodic	Dzongkha	dzon1239	Eurasia
Sino-Tibetan	Bodic	*Gurung*	guru1261	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Sino-Tibetan	Bodic	Johari	rang1266	Eurasia
Sino-Tibetan	Bodic	*Kham (Dege)*	kham1282	Eurasia
Sino-Tibetan	Bodic	*Kham (Nangchen)*	kham1282	Eurasia
Sino-Tibetan	Bodic	Kinnauri	kinn1249	Eurasia
Sino-Tibetan	Bodic	*Kurtöp*	kurt1248	Eurasia
Sino-Tibetan	Bodic	Kyirong	kyer1238	Eurasia
Sino-Tibetan	Bodic	Manange	mana1288	Eurasia
Sino-Tibetan	Bodic	Marchha	rong1264	Eurasia
Sino-Tibetan	Bodic	Pattani	patt1248	Eurasia
Sino-Tibetan	Bodic	*Purki*	puri1258	Eurasia
Sino-Tibetan	Bodic	*Tamang (Eastern)*	east2347	Eurasia
Sino-Tibetan	Bodic	Thakali	thak1245	Eurasia
Sino-Tibetan	Bodic	Tibetan (Amdo)	amdo1237	Eurasia
Sino-Tibetan	Bodic	*Tibetan (Drokpa)*	tibe1272	Eurasia
Sino-Tibetan	Bodic	*Tibetan (Modern Literary)*	tibe1272	Eurasia
Sino-Tibetan	Bodic	*Tibetan (Standard Spoken)*	tibe1272	Eurasia
Sino-Tibetan	Bodic	Tinani	tina1246	Eurasia
Sino-Tibetan	Bodic	*Tshangla*	tsha1245	Eurasia
Sino-Tibetan	Bodic	Western Tamang	west2415	Eurasia
Sino-Tibetan	Brahmaputran	(Haʔwa) Nocte	noct1238	Eurasia
Sino-Tibetan	Brahmaputran	Bodo	bodo1269	Eurasia
Sino-Tibetan	Brahmaputran	Cak	chak1270	Eurasia
Sino-Tibetan	Brahmaputran	Chang	chan1313	Eurasia
Sino-Tibetan	Brahmaputran	Dimasa	dima1251	Eurasia
Sino-Tibetan	Brahmaputran	*Garo*	garo1247	Eurasia
Sino-Tibetan	Brahmaputran	Jingpo	kach1280	Eurasia
Sino-Tibetan	Brahmaputran	Kachari	kach1279	Eurasia
Sino-Tibetan	Brahmaputran	Kadu (Asak)	kado1242	Eurasia
Sino-Tibetan	Brahmaputran	*Kokborok*	kokb1239	Eurasia
Sino-Tibetan	Brahmaputran	Konyak	kony1248	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Sino-Tibetan	Brahmaputran	Rabha	rabh1238	Eurasia
Sino-Tibetan	Dhimalish	Dhimal	dhim1246	Eurasia
Sino-Tibetan	Dhimalish	Toto	toto1302	Eurasia
Sino-Tibetan	Himalayish	*Athpare*	athp1241	Eurasia
Sino-Tibetan	Himalayish	Bantawa	bant1281	Eurasia
Sino-Tibetan	Himalayish	*Chepang*	chep1245	Eurasia
Sino-Tibetan	Himalayish	Chintang	chhi1245	Eurasia
Sino-Tibetan	Himalayish	Khaling	khal1275	Eurasia
Sino-Tibetan	Himalayish	*Kham*	west2420	Eurasia
Sino-Tibetan	Himalayish	Limbu	limb1266	Eurasia
Sino-Tibetan	Himalayish	*Newari (Kathmandu)*	newa1246	Eurasia
Sino-Tibetan	Himalayish	Thangmi	than1259	Eurasia
Sino-Tibetan	Himalayish	Thulung	thul1246	Eurasia
Sino-Tibetan	Himalayish	Wambule	wamb1257	Eurasia
Sino-Tibetan	Himalayish	Western Magar	west2418	Eurasia
Sino-Tibetan	Himalayish	Yakkha	yakk1236	Eurasia
Sino-Tibetan	Karenic	Geba Karen	geba1237	Eurasia
Sino-Tibetan	Karenic	*Karen (Sgaw)*	sgaw1245	Eurasia
Sino-Tibetan	Karenic	Kayan	kaya1315	Eurasia
Sino-Tibetan	Kho-Bwa	Duhumbi (Chugpa)	chug1252	Eurasia
Sino-Tibetan	Kho-Bwa	Sherdukpen	sher1257	Eurasia
Sino-Tibetan	Kman-Meyor	Mishmi	miju1243	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Akha Lai	haka1240	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	*Angami*	anga1288	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	*Ao*	aona1235	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	*Bawm*	bawm1236	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Chiru	chir1283	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Sino-Tibetan	Kuki-Chin-Naga	Daai Chin	daai1236	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Falam Chin	fala1243	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Gangte	gang1266	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Hakha Chin	haka1240	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Jejara	para1302	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Karbi (Amri variety)	amri1238	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Karbi (Hills variety)	karb1241	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Kharam	khar1288	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Khezha Naga	khez1235	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Kom (2)	komi1270	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Lamkang	lamk1238	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Liangmai	lian1251	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	*Lotha*	loth1237	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Mizo	lush1249	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	*Naga (Mao)*	maon1238	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	*Naga (Tangkhul)*	tang1336	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Paite	pait1244	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Rongmei	rong1266	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Sema	sumi1235	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Sino-Tibetan	Kuki-Chin-Naga	Sizang Chin	siyi1240	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Tiddim Chin	tedi1235	Eurasia
Sino-Tibetan	Kuki-Chin-Naga	Zeme	zeme1240	Eurasia
Sino-Tibetan	Lolo-Burmese	*Akha*	akha1245	Eurasia
Sino-Tibetan	Lolo-Burmese	*Bisu*	bisu1244	Eurasia
Sino-Tibetan	Lolo-Burmese	*Burmese*	nucl1310	Eurasia
Sino-Tibetan	Lolo-Burmese	*Hani*	hani1248	Eurasia
Sino-Tibetan	Lolo-Burmese	Khatso	kats1235	Eurasia
Sino-Tibetan	Lolo-Burmese	Lacid	lash1243	Eurasia
Sino-Tibetan	Lolo-Burmese	Lahu	lahu1253	Eurasia
Sino-Tibetan	Lolo-Burmese	*Lalo*	xish1235	Eurasia
Sino-Tibetan	Lolo-Burmese	*Lisu*	hual1240	Eurasia
Sino-Tibetan	Lolo-Burmese	*Lolo (Yao'An)*	lolo1259	Eurasia
Sino-Tibetan	Lolo-Burmese	Maru (Lhaovo)	maru1249	Eurasia
Sino-Tibetan	Lolo-Burmese	Mising	misi1242	Eurasia
Sino-Tibetan	Lolo-Burmese	Nyishi	nyis1236	Eurasia
Sino-Tibetan	Lolo-Burmese	Youle Jino	youl1235	Eurasia
Sino-Tibetan	Lolo-Burmese	Zaiwa	zaiw1241	Eurasia
Sino-Tibetan	Macro-Bai	*Bai*	cent2004, nort2724, sout2730	Eurasia
Sino-Tibetan	Macro-Tani	*Apatani*	apat1240	Eurasia
Sino-Tibetan	Macro-Tani	Milang	mila1245	Eurasia
Sino-Tibetan	Mixed Language	Wutun	wutu1241	Eurasia
Sino-Tibetan	Mruic	*Hkongso (Anu-Hkongso superset)*	anuu1241	Eurasia
Sino-Tibetan	Na-Qiangic	(Wobzi) Khroskyab	guan1266	Eurasia
Sino-Tibetan	Na-Qiangic	Eastern Geshiza	east2851	Eurasia
Sino-Tibetan	Na-Qiangic	Ersu	ersu1241	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Sino-Tibetan	Na-Qiangic	Guiqiong	guiq1238	Eurasia
Sino-Tibetan	Na-Qiangic	*Gyarong (Cogtse)*	situ1238	Eurasia
Sino-Tibetan	Na-Qiangic	Munya	west2417	Eurasia
Sino-Tibetan	Na-Qiangic	Namuyi Khato	namu1246	Eurasia
Sino-Tibetan	Na-Qiangic	Naxi	naxi1245	Eurasia
Sino-Tibetan	Na-Qiangic	nDrapa	zhab1238	Eurasia
Sino-Tibetan	Na-Qiangic	*Pumi*	pumi1242	Eurasia
Sino-Tibetan	Na-Qiangic	Puxi Qiang	sout2728	Eurasia
Sino-Tibetan	Na-Qiangic	*Qiang*	nort2722	Eurasia
Sino-Tibetan	Na-Qiangic	*rGyalrong (Caodeng)*	zbua1234	Eurasia
Sino-Tibetan	Na-Qiangic	rGyalrong (Japhug)	japh1234	Eurasia
Sino-Tibetan	Na-Qiangic	Shǐxīng	shix1238	Eurasia
Sino-Tibetan	Na-Qiangic	Tangut	tang1334	Eurasia
Sino-Tibetan	Nungish	*Anong*	nung1282	Eurasia
Sino-Tibetan	Nungish	Dulong	drun1238	Eurasia
Sino-Tibetan	Sinitic	*Cantonese*	cant1236	Eurasia
Sino-Tibetan	Sinitic	Chaozhou	chao1238	Eurasia
Sino-Tibetan	Sinitic	Hakka Chinese	hakk1236	Eurasia
Sino-Tibetan	Sinitic	*Mandarin*	mand1415	Eurasia
Sino-Tibetan	Sinitic	Wu Chinese	wuch1236	Eurasia
Sino-Tibetan	Tujia	*Tujia (Northern)*	nort2732	Eurasia
Siouan	Siouan	*Assiniboine*	assi1247	North America
Siouan	Siouan	Crow	crow1244	North America
Siouan	Siouan	Hidatsa	hida1246	North America
Siouan	Siouan	*Mandan*	mand1446	North America
Siouan	Siouan	Omaha-Ponca	omah1247	North America
Sko	Sko	*I'saka*	kris1246	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Sko	Sko	*Poko-Rawo*	rawo1244	Papunesia
Songhay	Songhay	Humburi Senni	humb1243	Africa
Songhay	Songhay	*Koyra Chiini*	koyr1240	Africa
Songhay	Songhay	*Koyraboro Senni*	koyr1242	Africa
Songhay	Songhay	Tondi Sonway Kiini	tond1249	Africa
Songhay	Songhay	*Zarma*	zarm1239	Africa
South Bird's Head	Inanwatan	*Inanwatan*	suab1238	Papunesia
South Bougainville	South Bougainville	*Motuna*	siwa1245	Papunesia
Southern Daly	Southern Daly	Murriny Patha	murr1258	Australia
Sulka	Sulka	*Sulka*	sulk1246	Papunesia
Surmic	South Surmic	Murle	murl1244	Africa
Surmic	Surmic	Majang	maja1242	Africa
Taiap	Taiap	Tayap (or Taiap)	taia1239	Papunesia
Tai-Kadai	Hlaic	Hlai (Baoding)	hlai1239	Eurasia
Tai-Kadai	Kadaic	Gelao (Central Qau variety)	qaua1234	Eurasia
Tai-Kadai	Kadaic	Lachi	lach1248	Eurasia
Tai-Kadai	Kadaic	Zoulei	aoua1234	Eurasia
Tai-Kadai	Kam-Tai	Ahom	ahom1240	Eurasia
Tai-Kadai	Kam-Tai	Aiton	aito1238	Eurasia
Tai-Kadai	Kam-Tai	*Dong (Southern)*	sout2741	Eurasia
Tai-Kadai	Kam-Tai	Lao	laoo1244	Eurasia
Tai-Kadai	Kam-Tai	Mak (China)	makc1235	Eurasia
Tai-Kadai	Kam-Tai	Maonan	maon1241	Eurasia
Tai-Kadai	Kam-Tai	*Nung (in Vietnam)*	nung1283	Eurasia
Tai-Kadai	Kam-Tai	Phake	phak1238	Eurasia
Tai-Kadai	Kam-Tai	*Sanjiang Kam*	caom1238	Eurasia
Tai-Kadai	Kam-Tai	*Thai*	thai1261	Eurasia
Tai-Kadai	Kam-Tai	*Yay*	bouy1240	Eurasia
Tamaic	Tamaic	Tama	tama1331	Africa

Family	Genus	Language	Glottocode	Macroarea
Ta-Ne-Omotic	Ta-Ne-Omotic	*Gamo*	gamo1243	Africa
Ta-Ne-Omotic	Ta-Ne-Omotic	Koorete	koor1239	Africa
Ta-Ne-Omotic	Ta-Ne-Omotic	Wolaytta	wola1242	Africa
Ta-Ne-Omotic	Ta-Ne-Omotic	*Yemsa*	yems1235	Africa
Ta-Ne-Omotic	Ta-Ne-Omotic	Zayse	zays1235	Africa
Tangkic	Tangkic	Kayardild	kaya1319	Australia
Tarascan	Tarascan	Purépecha	pure1242	North America
Taushiro	Taushiro	Taushiro	taus1253	South America
Teberan	Teberan	Dadibi	dadi1250	Papunesia
Timor-Alor-Pantar	Alor-Pantar	*Abui*	abui1241	Papunesia
Timor-Alor-Pantar	Alor-Pantar	*Adang*	adan1251	Papunesia
Timor-Alor-Pantar	Alor-Pantar	*Klon*	kelo1247	Papunesia
Timor-Alor-Pantar	Bunak	Bunaq	buna1278	Papunesia
Timor-Alor-Pantar	East Timor	*Makasae*	maka1316	Papunesia
Tiwi	Tiwi	*Tiwi*	tiwi1244	Australia
Tonkawa	Tonkawa	Tonkawa	tonk1249	North America
Tsimshian	Tsimshian	Nisgha	nisg1240	North America
Tsimshian	Tsimshian	*Tsimshian (Coast)*	coas1300	North America
Tucanoan	Tucanoan	*Kotiria*	guan1269	South America
Tucanoan	Tucanoan	*Retuarã*	tani1257	South America
Tucanoan	Tucanoan	*Siona-Tetete*	sion1247	South America
Tungusic	Tungusic	Even	even1260	Eurasia
Tungusic	Tungusic	*Evenki*	even1259	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Tungusic	Tungusic	Kilen	kile1243	Eurasia
Tungusic	Tungusic	*Udihe*	udih1248	Eurasia
Tupian	Arikemic	Karitiâna	kari1311	South America
Tupian	Maweti-Guarani	Cocama	coca1259	South America
Tupian	Maweti-Guarani	Émérillon	emer1243	South America
Tupian	Maweti-Guarani	Guajajara	guaj1255	South America
Tupian	Maweti-Guarani	*Guaraní*	para1311	South America
Tupian	Maweti-Guarani	Kamaiurá	kama1373	South America
Tupian	Maweti-Guarani	Mbyá Guaraní	mbya1239	South America
Tupian	Maweti-Guarani	Urubú-Kaapor	urub1250	South America
Tupian	Tupi-Guarani	*Yuqui*	yuqu1240	South America
Turkic	Bolgar	Chuvash	chuv1255	Eurasia
Turkic	Common Turkic	*Bashkir*	bash1264	Eurasia
Turkic	Common Turkic	Karachay-Balkar	kara1465	Eurasia
Turkic	Common Turkic	Karakalpak	kara1467	Eurasia
Turkic	Common Turkic	Sakha	yaku1254	Eurasia
Turkic	Common Turkic	*Tatar*	tata1255	Eurasia
Turkic	Common Turkic	*Turkish*	nucl1301	Eurasia
Turkic	Common Turkic	Turkmen	turk1304	Eurasia
Turkic	Common Turkic	*Uzbek*	nort2690, sout2699	Eurasia

Family	Genus	Language	Glottocode	Macroarea
Uralic	Finnic	*Estonian*	esto1258	Eurasia
Uralic	Finnic	*Finnish*	finn1318	Eurasia
Uralic	Khantyic	*Khanty*	khan1273	Eurasia
Uralic	Khantyic	Surgut Khanty	surg1248	Eurasia
Uralic	Mansic	Mansi	mans1258	Eurasia
Uralic	Mordvin	Erzya	erzy1239	Eurasia
Uralic	Permian	*Udmurt*	udmu1245	Eurasia
Uralic	Saami	North Saami	nort2671	Eurasia
Uralic	Saami	Pite Saami	pite1240	Eurasia
Uralic	Saami	Skolt Saami	skol1241	Eurasia
Uralic	Samoyedic	Kamas	kama1378, kama1351	Eurasia
Uralic	Samoyedic	*Nenets*	nene1249	Eurasia
Uralic	Samoyedic	Selkup	selk1253	Eurasia
Urarina	Urarina	*Urarina*	urar1246	South America
Uru-Chipaya	Uru-Chipaya	*Chipaya*	chip1262	South America
Uto-Aztecan	Northern Uto-Aztecan	Comanche	coma1245	North America
Uto-Aztecan	Northern Uto-Aztecan	Cupeño	cupe1243	North America
Uto-Aztecan	Northern Uto-Aztecan	Hopi	hopi1249	North America
Uto-Aztecan	Northern Uto-Aztecan	Paiute (Northern)	nort2954	North America
Uto-Aztecan	Northern Uto-Aztecan	Shoshoni	shos1248	North America
Uto-Aztecan	Northern Uto-Aztecan	Ute-Southern Paiute	utes1238	North America
Uto-Aztecan	Southern Uto-Aztecan	Cora	elna1235	North America
Uto-Aztecan	Southern Uto-Aztecan	*Nahuatl (Huasteca)*	west2624	North America

Family	Genus	Language	Glottocode	Macroarea
Uto-Aztecan	Southern Uto-Aztecan	Nahuatl (Michoacán)	mich1245	North America
Uto-Aztecan	Southern Uto-Aztecan	Nahuatl (Tetelcingo)	tete1251	North America
Uto-Aztecan	Southern Uto-Aztecan	Nevome	pima1248	North America
Uto-Aztecan	Southern Uto-Aztecan	O'odham	toho1245	North America
Uto-Aztecan	Southern Uto-Aztecan	Tarahumara (Western)	lowl1265	North America
Uto-Aztecan	Southern Uto-Aztecan	*Yaqui*	yaqu1251	North America
Warao	Warao	Warao	wara1303	South America
Washo	Washo	Washo	wash1253	North America
West Bomberai	West Bomberai	Kalamang	kara1499	Papunesia
Western Daly	Western Daly	*Maranungku*	mara1386	Australia
Xincan	Xincan	Chiquimulilla	xinc1242	North America
Xincan	Xincan	*Guazacapán*	xinc1246	North America
Xincan	Xincan	Jumaytepeque	xink1235	North America
Xincan	Xincan	Yupiltepeque	xinc1243	North America
Yam	Morehead-Maro	Arammba	aram1253	Papunesia
Yam	Morehead-Maro	Komnzo	wara1294	Papunesia
Yangmanic	Yangmanic	Wardaman	ward1246	Australia
Yanomamic	Yanomamic	Shiriana	nina1238	South America
Yanomamic	Yanomamic	Yanomamö	yano1261	South America
Yele	West New Britain	Yélî Dnye	yele1255	Papunesia

Family	Genus	Language	Glottocode	Macroarea
Yeniseian	Yeniseian	*Ket*	kett1243	Eurasia
Yokutsan	Yokutsan	Yokuts (Yaudanchi)	yoku1256	North America
Yuchi	Yuchi	*Yuchi*	yuch1247	North America
Yukaghir	Yukaghir	Yukaghir (Kolyma)	sout2750	Eurasia
Yuki-Wappo	Yuki-Wappo	*Wappo*	wapp1239	North America
Zamucoan	Zamucoan	Chamacoco	cham1315	South America
Zuni	Zuni	Zuni	zuni1245	North America

A.2 N-A-Num-Dem

Language	**Type**	**WALS**	**V-I**	**V-M**	**V-F**	**Glottocode**
Uduk	0		Yes	Yes	Yes	uduk1239
Lamang	1	VSO	Yes	No	No	lama1288
Maori	1	VSO	Yes	No	No	maor1246
Nias	1	VOS				nias1242
Tukang Besi [197]	1	VOS	Yes	Yes	No	tuka1247, tuka1248
Turkana	1	VSO	Yes	No	No	turk1308
Yakan	1		Yes	No	No	yaka1277
Balinese	1.5		Yes	Yes	No	bali1278
Begak-Ida'an	1.5		Yes	Yes	No	idaa1241
Buang (Mapos)	1.5		Yes	Yes	No	mapo1242
Lawa (Eastern)	1.5		Yes	Yes	No	east2330
Abun	2	SVO	No	Yes	No	abun1252
Adzera	2	SVO	No	Yes	No	adze1240
Aiton	2		No	Yes	No	aito1238
Akan	2	SVO			No	akan1250

[197] Grambank argues for the presence of an unmarked SVO orders due to a misinterpretation of Donohue (2011: 60). Donohue (2011: 51 ff.) clarifies that the "s-V-o" order reported there actually indicates the indexing on the verbal affixes, whereas overt arguments are postverbal when not dropped or focalized.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Alorese	2		No	Yes	No	alor1247
Amara	2	SVO	No	Yes	No	amar1272
Ambae (Lolovoli Northeast)	2	SVO	No	Yes	No	east2443
Ambai	2	SVO			No	amba1265
Ambonese Malay	2	SVO (APiCS)	No	Yes	No	ambo1250
Arop-Lokep	2	SVO	No	Yes	No	arop1243
Arop-Sissano	2		No	Yes	No	arop1242
Au	2	SVO	No	Yes	No	auuu1241
Avatime	2		No	Yes	No	avat1244
Avava	2		No	Yes	No	katb1237
Bafia	2		No	Yes	No	bafi1243
Bafut	2	SVO	No	Yes	No	bafu1246
Bali-Vitu	2	SVO	No	Yes	No	unea1237
Bariai	2		No	Yes	No	bari1286
Biak	2	SVO	No	Yes	No	biak1248
Big Nambas	2		No	Yes	No	bign1238
Birom	2	SVO				bero1242
Blang	2		No	Yes	No	blan1242
Bola	2		No	Yes	No	bola1250
Bongo	2	SVO	No	Yes	No	bong1285
Brao	2	SVO	No	Yes	No	lave1249
Buduma	2	SVO	No	Yes	No	budu1265
Bugan	2		No	Yes	No	buga1247
Bugawac	2		No	Yes	No	buga1250
Buru	2	SVO	No	Yes	No	buru1303
Buwal	2		No	Yes	No	buwa1243
Chakali	2		No	Yes	No	chak1271
Chichewa	2	SVO				nyan1308
Chong	2		No	Yes	No	cent2314
Dagbani	2	SVO	No	Yes	No	dagb1246
Dakaka	2		No	Yes	No	daka1243

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Day	2	SVO			No	dayy1236
Degema	2	SVO	No	Yes	No	dege1246
Dhao	2		No	Yes	No	dhao1237
Doyayo	2	SVO			No	doya1240
Emai	2	SVO	No	Yes	No	emai1242
Engenni	2	SVO			No	enge1239
Ewe	2	SVO	No	Yes	No	ewee1241
Fali (South)	2		No	Yes	No	sout2782
Fongbe	2	SVO	No	Yes	No	fonn1241
Geba Karen	2		No	Yes	No	geba1237
Gola	2		No	Yes	No	gola1255
Gwama	2		No	Yes	No	kwam1249
Hatam	2	SVO	No	Yes	No	hata1243
Hote	2		No	Yes	No	hote1245
Ifira-Mele	2	SVO	No	Yes	No	mele1250
Igbo	2	SVO	No	Yes	No	nucl1417
Igede	2	SVO			No	iged1239
Irarutu	2	SVO	No	Yes	No	irar1238
Jabêm	2	SVO	No	Yes	No	yabe1254
Jambi Malay	2		No	Yes	No	jamb1236
Jukun	2		No	Yes	No	juku1254
Kam	2		No	Yes	No	kamm1249
Kara (in Papua New Guinea)	2		No	Yes	No	kara1486
Karen (Sgaw)	2	SVO	No	Yes	No	sgaw1245
Karnai	2		No	Yes	No	karn1252
Kasong	2	SVO	No	Yes	No	chon1285
Kenga	2	SVO	No	Yes	No	keng1240
Kéo	2		No	Yes	No	keoo1238
Khmer	2	SVO	No	Yes	No	cent1989
Khmu	2	SVO	No	Yes	No	khmu1256
Kolbila	2		No	Yes	No	kolb1240

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Koromfe	2	SVO	No	Yes	No	koro1298
Kosraean	2	SVO	No	Yes	No	kosr1238
Koyra Chiini	2	SVO	No	Yes	No	koyr1240
Kulango	2		No	Yes	No	boun1243
Kuy	2		No	Yes	No	kuyy1240
Laal	2	SVO	No	Yes	No	laal1242
Labu	2	SVO	No	Yes	No	labu1248
Lao	2	SVO	No	Yes	No	laoo1244
Laven	2		No	Yes	No	lave1248
Leggbó	2	SVO			No	legb1242
Lelepa	2		No	Yes	No	lele1267
Lemerig	2		No	Yes	No	leme1238
Lenakel	2	SVO	No	Yes	No	lena1238
Leti	2	SVO	No	Yes	No	leti1246
Lewo	2	SVO	No	Yes	No	lewo1242
Loniu	2	SVO	No	Yes	No	loni1238
Luang	2		No	Yes	No	luan1263
Maleu	2	SVO	No	Yes	No	male1289
Malgwa	2	SVO	No	Yes	No	malg1251
Malua Bay	2		No	Yes	No	malu1245
Mambai	2		No	Yes	No	mamb1294
Mangseng	2		No	Yes	No	mang1403
Matbat	2		No	Yes	No	matb1237
Mato	2		No	Yes	No	mato1252
Maybrat	2	SVO	No	Yes	No	maib1239
Mbay	2	SVO	No	Yes	No	mbay1241
Mbum	2	SVO			No	mbum1254
Meramera	2		No	Yes	No	mera1242
Meyah	2	SVO	No	Yes	No	meya1236
Mlabri	2	SVO	No	Yes	No	mlab1235
Mnong (Central)	2		No	Yes	No	cent1992
Moken	2		No	Yes	No	moke1242

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Mokilese	2	SVO	No	Yes	No	moki1238
Mon	2	SVO	No	Yes	No	monn1252
Mpade	2		No	Yes	No	mpad1242
Mpur	2	SVO	No	Yes	No	mpur1239
Mundang	2	SVO	No	Yes	No	mund1325
Mungbam	2		No	Yes	No	abar1238
Mwotlap	2	SVO	No	Yes	No	motl1237
Nalik	2	SVO	No	Yes	No	nali1244
Navahaq	2		No	Yes	No	sout2857
Nawuri	2		No	Yes	No	nawu1242
Ndumu	2	SVO				ndum1239
Ndut	2	SVO			No	ndut1239
Nese	2		No	Yes	No	nese1235
Neverver	2		No	Yes	No	ling1265
Ngizim	2	SVO			No	ngiz1242
Ninde	2		No	Yes	No	labo1244
Nkonya	2	SVO	No	Yes	No	nkon1248
Nsenga	2	SVO	No	Yes	No	nsen1242
Nuaulu	2	SVO	No	Yes	No	sout2895
Nukna	2		No	Yes	No	nukn1238
Oko	2		No	Yes	No	okoe1238
Olo	2	SVO	No	Yes	No	oloo1241
Palaung	2	SVO	No	Yes	No	ruch1235, ruma1248, shwe1236
Papuan Malay	2		No	Yes	No	papu1250
Phake	2		No	Yes	No	phak1238
Pingelapese	2		No	Yes	No	ping1243
Rotuman	2	SVO	No	Yes	No	rotu1241
Samba Leko	2	SVO	No	Yes	No	samb1305
Sapuan	2	SVO	No	Yes	No	sapu1248
Sekpele	2		No	Yes	No	sekp1241

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Selaru	2		No	Yes	No	sela1259
Selee	2		No	Yes	No	sele1249
Sherbro	2		No	Yes	No	sher1258
Shilluk	2	SVO	No	Yes	No	shil1265
Sio	2	SVO	No	Yes	No	sioo1240
Siwu	2		No	Yes	No	siwu1238
Sougb	2	SVO	No	Yes	No	mani1235
South Efate [198]	2	SVO	No	No	Yes	sout2856
Southwest Tanna	2		No	Yes	No	sout2869
Stieng	2	SVO	No	Yes	No	bulo1242, bude1236
Sudest	2	SVO	No	Yes	No	sude1239
Taba	2	SVO	No	Yes	No	east2440
Tafi	2		No	Yes	No	tafi1243
Tami	2		No	Yes	No	tami1290
Ternate	2		No	Yes	No	tern1247
Thai	2	SVO	No	Yes	No	thai1261
Thavung	2		No	Yes	No	aheu1239
Tidore	2	SVO	No	Yes	No	tido1248
Tirax	2		No	Yes	No	maee1241
Tugun	2	SVO	No	Yes	No	tugu1245
Tuwuli	2		No	Yes	No	tuwu1238
Unua	2		No	Yes	No	unua1237
Urak Lawoi’	2	SVO	No	Yes	No	urak1238
Urim	2		No	Yes	No	urim1252
Ut-Ma’in	2		No	Yes	No	kagf1238
Vera’a	2		No	Yes	No	vera1241
Vinmavis	2	SVO	No	Yes	No	vinm1237
Walman	2	SVO			No	valm1241
Waropen	2		No	Yes	No	waro1242
Western Cham	2	SVO	No	Yes	No	west2650

[198] Thieberger (2006: 267) supports the classification that WALS gives for South Efate.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Whitesands	2		No	Yes	No	whit1269
Yaaku	2		No	Yes	No	yaak1241
Yeri	2		No	Yes	No	yapu1240
Yoruba	2	SVO			No	yoru1245
Ahom	2.5		No	Yes	Yes	ahom1240
Bunaq	2.5		No	Yes	Yes	buna1278
Gwari	2.5	SOV or SVO				gbag1258
Humburi Senni	2.5		No	Yes	Yes	humb1243
Koyraboro Senni [199]	2.5	SOV	No	Yes	Yes	koyr1242
Logoti [200]	2.5	SOV or SVO	No	Yes	No	logo1259
Maku'a	2.5		No	Yes	Yes	maku1277
Mangga Buang	2.5		No	Yes	Yes	mang1404
Poko-Rawo	2.5	SOV or SVO				rawo1244
Tonkawa [201]	2.5	SOV or SVO	No	No	Yes	tonk1249
Western Tamang	2.5		No	Yes	Yes	west2415
Wik-Mungkan	2.5		No	Yes	Yes	wikm1247
Yimas	2.5		No	Yes	Yes	yima1243
Zarma [202]	2.5	No dom	No	Yes	Yes	zarm1239
Abau	3	SOV	No	No	Yes	abau1245
Adang	3	SOV	No	No	Yes	adan1251
Ama (Papa New Guinea)	3		No	No	Yes	amap1240
Angami	3	SOV	No	No	Yes	anga1288
Ao	3	SOV	No	No	Yes	aona1235
Arrernte (Mparntwe)	3	SOV	No	No	Yes	east2379
Assiniboine	3		No	No	Yes	assi1247
Awad Bing	3		No	No	Yes	awad1244

199 There is a small subset of verbs whose mandatory order is SVO, and clauses without aspectual markers are SVO. See the data provided at https://songhaylanguage.wordpress.com/

200 Though see the analysis in Wright (1995: Sec. 4.2, 4.4, 7), which proposes an umarked SVO word order with OV-nominalizations.

201 What is reported by WALS finds further support in Wier (2016: 26-27).

202 Hanafiou (2007: 57) clarifies that indeterminacy on WALS is due to unmarked SVO and SOV word orders.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Barai	3	SOV	No	No	Yes	nucl1630
Bargam	3		No	No	Yes	barg1252
Bawm	3	SOV	No	No	Yes	bawm1236
Ben Tey	3		No	No	Yes	bent1238
Biangai	3		No	No	Yes	bian1252
Bimin	3		No	No	Yes	bimi1240
Binandere	3	SOV	No	No	Yes	bina1277
Bribri	3	SOV	No	No	Yes	brib1243
Buhutu	3		No	No	Yes	buhu1237
Busa	3	SOV	No	No	Yes	busa1253
Cabécar	3		No	No	Yes	cabe1245
Chickasaw	3		No	No	Yes	chic1270
Choctaw	3	SOV	No	No	Yes	choc1276
Daga	3	SOV				daga1275
Dazaga	3		No	No	Yes	daza1242
Dizin	3		No	No	Yes	dizi1235
Dom	3		No	No	Yes	domm1246
Duna	3	SOV	No	No	Yes	duna1248
Edolo	3	SOV				edol1239
Gangte	3		No	No	Yes	gang1266
Golin	3	SOV	No	No	Yes	goli1247
Guhu-Samane	3	SOV	No	No	Yes	guhu1244
Gumawana	3	SOV	No	No	Yes	guma1254
Hidatsa	3	SOV	No	No	Yes	hida1246
Huli	3		No	No	Yes	huli1244
I'saka	3	SOV	No	No	Yes	kris1246
Iyo	3		No	No	Yes	iyoo1238
Jalkunan	3		No	No	Yes	jalk1242
Kairiru	3	SOV	No	No	Yes	kair1263
Kaluli	3		No	No	Yes	kalu1248
Kamasau	3	SOV	No	No	Yes	kama1367

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Kamula[203]	3		No	Maybe	Yes	kamu1260
Kanuri	3	SOV	No	No	Yes	cent2050
Kham (Dege)	3	SOV	No	No	Yes	kham1282
Kham (Nangchen)	3	SOV	No	No	Yes	kham1282
Khezha Naga	3		No	No	Yes	khez1235
Klon	3		No	No	Yes	kelo1247
Koiali (Mountain)	3	SOV	No	No	Yes	moun1252
Koita	3	SOV	No	No	Yes	koit1244
Konai	3		No	No	Yes	kona1242
Korafe-Yegha	3		No	No	Yes	kora1294
Kovai	3		No	No	Yes	kova1243
Kugu Nganhcara	3		No	No	Yes	kuku1280
Kunama	3	SOV	No	No	Yes	kuna1268
Kyaka	3	SOV				kyak1244
Lotha	3	SOV	No	No	Yes	loth1237
Ma Manda	3		No	No	Yes	sauk1252
Maia	3		No	No	Yes	maia1254
Mairasi	3	SOV	No	No	Yes	nucl1594
Manam	3	SOV	No	No	Yes	mana1295
Mandan	3	SOV				mand1446
Mauwake	3		No	No	Yes	mauw1238
Mende (Papua New Guinea)	3		No	No	Yes	mend1268
Menya	3	SOV	No	No	Yes	meny1245
Mian	3	SOV	No	No	Yes	mian1256
Momu-Fas	3	SOV	No	No	Yes	fass1245
Murriny Patha	3	SOV	No	No	Yes	murr1258
Nabak	3	SOV	No	No	Yes	naba1256
Naga (Tangkhul)	3	SOV	No	No	Yes	tang1336
Namia	3	SOV	No	No	Yes	nami1256
Nend	3	SOV	No	No	Yes	nend1239

[203] The data discussed by Routamaa (1997) show that the ”medial verbs” on which Grambank shows some reservations are either coordinated or serialized verbs. Even in such cases, OV structures are present.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Ngaanyatjarra	3	SOV	No	No	Yes	ngaa1240
Ngäbere	3	SOV	No	No	Yes	ngab1239
Nggem	3		No	No	Yes	ngge1241
Omaha-Ponca	3	SOV	No	No	Yes	omah1247
Orokaiva	3	SOV	No	No	Yes	orok1269
Oromo (Harar)	3	SOV				east2652
Owa	3		No	No	Yes	owaa1237
Rawa (Karo)	3	SOV	No	No	Yes	rawa1267
Selepet	3		No	No	Yes	sele1250
Sheko	3		No	No	Yes	shek1245
Siroi	3	SOV	No	No	Yes	siro1249
Sizang Chin	3		No	No	Yes	siyi1240
Suena	3	SOV	No	No	Yes	suen1241
Telefol	3	SOV	No	No	Yes	tele1256
Tibetan (Drokpa)	3	SOV	No	No	Yes	tibe1272
Tobelo	3		No	No	Yes	tobe1252
Tommo So	3		No	No	Yes	tomm1242
Tondi Sonway Kiini	3		No	No	Yes	tond1249
Ulwa	3		No	No	Yes	ulwa1239
Usan	3	SOV	No	No	Yes	usan1239
Vai	3	SOV				vaii1241
Wan	3		No	No	Yes	wann1242
Waskia	3	SOV	No	No	Yes	wask1241
Weri	3		No	No	Yes	weri1253
Yale (Kosarek)	3	SOV	No	No	Yes	kosa1249
Yongkom	3		No	No	Yes	yong1280
Zaiwa	3		No	No	Yes	zaiw1241

A.3 Dem-Num-A-N

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Chukchi [204]	0	No dom	No	No	Yes	chuk1273
Jaqaru	0		Yes	Yes	Yes	jaqa1244
Kayardild	0	No dom	Yes	Yes	Yes	kaya1319
Nez Perce	0	No dom	Yes	Yes	Yes	nezp1238
O'odham	0	No dom	Yes	Yes	Yes	toho1245
Yokuts (Yaudanchi)	0	No dom	Yes	Yes	Yes	yoku1256
Amis	1	VSO or VOS	Yes	No	No	amis1246
Ashéninka Perené	1		Yes	No	No	ashe1272
Cakchiquel	1	VOS	Yes	No	No	kaqc1270
Cayuvava	1	VOS	Yes	No	No	cayu1262
Cebuano	1	VSO	Yes	No	No	cebu1242
Chamorro	1	VSO	Yes	No	No	cham1312
Halkomelem (Upriver)	1	VSO	Yes	No	No	halk1245
Maranao	1		Yes	No	No	mara1404
Masbatenyo	1		Yes	No	No	masb1238
Musqueam	1	VSO				halk1245, musq1240
Nicobarese (Car)	1	VOS	Yes	No	No	carn1240
Quileute	1	VSO	Yes	No	No	quil1240
Rukai	1	VSO or VOS	Yes	No	No	ruka1240
Sipakapense	1	VSO				sipa1247
Squamish	1	VSO				squa1248
Talaandig-Binukid	1		Yes	No	No	binu1244
Greek (Modern) [205]	1.5	SVO or VSO	No	Yes	No	mode1248
Maung [206]	1.5	SVO	Yes	Yes	No	maun1240
Nyulnyul	1.5		Yes	Yes	No	nyul1247
Tiwi [207]	1.5	SVO	Yes	Yes	No	tiwi1244

[204] See Kantarovich (2020: 316), whose discussion supports M. Dunn's comment on Grambank defining Chukchi as a "pragmatic word order language".
[205] See Alexiadou (2022: 46-47).
[206] See discussion in Singer (2006: 137 ff.) favoring Grambank's classification.
[207] See A. Wilson (2013: 60) in support of Grambank's classification.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Arapesh (Mountain)	2	SVO	No	Yes	No	buki1249
Arigidi	2		No	Yes	No	arig1246
Bosnian-Serbian-Croatian	2	SVO	No	Yes	No	sout1528
Bulgarian	2	SVO				bulg1262
Cantonese	2	SVO				cant1236
Chaozhou	2		No	Yes	No	chao1238
Czech	2	SVO	No	Yes	No	czec1258
English	2	SVO	No	Yes	No	stan1293
Erzya	2	SVO	No	Yes	No	erzy1239
Finnish	2	SVO	No	Yes	No	finn1318
Gunbalang	2	SVO	No	Yes	No	kunb1251
Hakka Chinese	2	SVO	No	Yes	No	hakk1236
Huastec	2	SVO	No	Yes	No	huas1242
Iwaidja	2	SVO				iwai1244
Latvian	2	SVO	No	Yes	No	latv1249
Macedonian	2	SVO	No	Yes	No	mace1250
Mandarin	2	SVO	No	Yes	No	mand1415
Mapudungun	2	SVO	No	Yes	No	mapu1245
North Saami	2	SVO	No	Yes	No	nort2671
Ocuilteco	2	SVO				atzi1235
Pite Saami	2		No	Yes	No	pite1240
Polish	2	SVO	No	Yes	No	poli1260
Saisiyat	2		No	Yes	No	sais1237
She	2		No	Yes	No	shee1238
Slovenian	2	SVO	No	Yes	No	slov1268
Ukrainian	2	SVO	No	Yes	No	ukra1253
Wu Chinese	2		No	Yes	No	wuch1236
Wuvulu-Aua	2		No	Yes	No	wuvu1239
Yurok [208]	2	No dom	No	Yes	No	yuro1248
Zoque (Chimalapa)	2	SVO				chim1300

[208] See Robins (1958: 19) and Blevins (2008: 2-4) for discussion supporting Grambank's classification.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Armenian (Eastern)	2.5	SOV or SVO	No	Yes	Yes	nucl1235
Leco	2.5		No	Yes	Yes	leco1242
Lendu	2.5	SOV or SVO	No	Yes	Yes	lend1245
Danish [209]	2-V2	SVO	No	Yes	No	dani1285
Estonian [210]	2-V2	SVO	No	Yes	No	esto1258
Faroese [211]	2-V2		No	Yes	No	faro1244
Icelandic [212]	2-V2	SVO	No	Yes	No	icel1247
Norwegian [213]	2-V2	SVO				norw1258
Swedish [214]	2-V2	SVO	No	Yes	No	swed1254
Ainu	3	SOV	No	No	Yes	ainu1240
Alaba-K'abeena	3		No	No	Yes	alab1254
Alamblak	3	SOV	No	Yes	Yes	alam1246
Amharic	3	SOV	No	No	Yes	amha1245
Arammba	3		No	No	Yes	aram1253
Archi	3	SOV	No	No	Yes	arch1244
Argobba	3		No	No	Yes	argo1244
Armenian (Western)	3	SOV				homs1234
Athpare	3	SOV	No	No	Yes	athp1241
Avar	3	SOV				avar1256
Awa	3	SOV				awap1236
Awa Pit	3	SOV	No	No	Yes	awac1239
Ayacucho Quechua	3		No	No	Yes	ayac1239
Aymara (Southern)	3		No	No	Yes	sout2996
Bagvalal	3		No	No	Yes	bagv1239
Bantawa	3		No	No	Yes	bant1281
Bashkir	3		No	No	Yes	bash1264
Bezhta	3		No	No	Yes	bezh1248
Bhojpuri	3	SOV	No	No	Yes	bhoj1244

209 Classification established following Holmberg (2015).
210 Classification established following Holmberg (2015).
211 Classification established following Holmberg (2015).
212 Classification established following Holmberg (2015).
213 Classification established following Holmberg (2015).
214 Classification established following Holmberg (2015).

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Brokskat	3	SOV				brok1247
Bundeli	3		No	No	Yes	bund1253
Buriat	3	SOV				mong1330
Burushaski	3	SOV	No	No	Yes	buru1296
Carib (Galibi)	3	SOV	No	No	Yes	gali1262
Cayapa	3	SOV	No	No	Yes	chac1249
Chaha	3	SOV				seba1251, chah1248
Chaudangsi	3	SOV				chau1259
Chepang	3	SOV	No	No	Yes	chep1245
Chintang	3		No	No	Yes	chhi1245
Chipaya	3	SOV				chip1262
Cholón	3	SOV	No	No	Yes	chol1284
Chuvash	3	SOV	No	No	Yes	chuv1255
Comanche	3	SOV	No	No	Yes	coma1245
Cusco Quechua	3		No	No	Yes	cusc1236
Dagur	3	SOV	No	No	Yes	daur1238
Dandami Maria	3		No	No	Yes	dand1238
Darai	3		No	No	Yes	dara1250
Darma	3	SOV	No	No	Yes	darm1243
Defaka	3		No	No	Yes	defa1248
Dhimal	3	SOV	No	No	Yes	dhim1246
Even	3	SOV	No	No	Yes	even1260
Evenki	3	SOV	No	No	Yes	even1259
Gamo	3	SOV	No	No	Yes	gamo1243
Godoberi	3		No	No	Yes	ghod1238
Gondi	3	SOV				nort2702
Gujarati	3	SOV				guja1252
Hadiyya	3		No	No	Yes	hadi1240
Hindi	3	SOV	No	No	Yes	hind1269
Hinuq	3		No	No	Yes	hinu1240
Hopi	3	SOV				hopi1249

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Hua	3	SOV	No	No	Yes	huaa1250
Huitoto (Minica)	3	SOV	No	No	Yes	mini1256
Hunzib	3	SOV	No	No	Yes	hunz1247
Ijo (Kolokuma)	3	SOV				izon1238
Japanese	3	SOV	No	No	Yes	nucl1643
Johari	3	SOV				rang1266
Kamas	3	SOV	No	No	Yes	kama1378, kama1351
Kannada	3	SOV	No	No	Yes	nucl1305
Karachay-Balkar	3	SOV				kara1465
Karakalpak	3	SOV				kara1467
Ket	3	SOV	No	No	Yes	kett1243
Khaling	3	SOV	No	No	Yes	khal1275
Khalkha	3	SOV	No	No	Yes	halh1238
Kham	3	SOV				west2420
Khanty	3	SOV				khan1273
Khwarshi-Inhoqwari	3		No	No	Yes	khva1239
Kilen	3		No	No	Yes	kile1243
Kinnauri	3	SOV	No	No	Yes	kinn1249
Kolami	3	SOV	No	No	Yes	nort2699
Komnzo	3		No	No	Yes	wara1294
Konkani	3		No	No	Yes	konk1267
Korku	3	SOV	No	No	Yes	kork1243
Koya	3	SOV	No	No	Yes	koya1251
Kryz	3	SOV	No	No	Yes	kryt1240
Kumauni	3	SOV				kuma1273
Kusunda	3	SOV	No	No	Yes	kusu1250
Kuvi	3	SOV	No	No	Yes	kuvi1243
Kwoma	3	SOV	No	No	Yes	kwom1262
Lak	3		No	No	Yes	lakk1252
Lamani	3	SOV	No	No	Yes	lamb1269
Laz	3		No	No	Yes	lazz1240

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Lezgian	3	SOV	No	No	Yes	lezg1247
Limbu	3		No	No	Yes	limb1266
Magahi	3	SOV	No	No	Yes	maga1260
Maithili	3		No	No	Yes	mait1250
Malayalam	3	SOV	No	No	Yes	mala1464
Malto	3	SOV				kuma1274
Mangghuer	3	SOV				huzh1238
Mansi	3	SOV				mans1258
Marathi	3	SOV	No	No	Yes	mara1378
Marchha	3	SOV	No	No	Yes	rong1264
Mundari	3	SOV	No	No	Yes	mund1320
Nama	3	SOV	No	No	Yes	nama1264
Nenets	3	SOV	No	No	Yes	nene1249
Nepali	3	SOV	No	No	Yes	nepa1254
Nevome	3	SOV	No	No	Yes	pima1248
Nkoroo	3		No	No	Yes	nkor1239
Oirad-Kalmyk-Darkhat	3		No	No	Yes	kalm1243
Paiute (Northern)	3	SOV	No	No	Yes	nort2954
Panjabi	3	SOV	No	No	Yes	panj1256
Pashto	3	SOV				cent1973
Pattani	3	SOV	No	No	Yes	patt1248
Prasuni	3	SOV	No	No	Yes	pras1239
Qafar	3	SOV	No	No	Yes	afar1241
Quechua (Huallaga)	3	SOV	No	No	Yes	hual1241
Quechua (Imbabura)	3	SOV	No	No	Yes	imba1240
Resígaro	3	SOV	No	No	Yes	resi1247
Rutul	3	SOV	No	No	Yes	rutu1240
Sakha	3		No	No	Yes	yaku1254
San Martín Quechua	3		No	No	Yes	sanm1289
Santali	3	SOV	No	No	Yes	sant1410
Sare	3	SOV				kapr1245

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Savosavo	3	SOV	No	No	Yes	savo1255
Selkup	3	SOV				selk1253
Shabo	3	SOV	No	No	Yes	shab1252
Shina	3	SOV				shin1264
Shoshoni	3	SOV	No	No	Yes	shos1248
Sidaama	3	SOV	No	No	Yes	sida1246
Siona-Tetete	3		No	No	Yes	sion1247
Skolt Saami	3		No	No	Yes	skol1241
Surgut Khanty	3		No	No	Yes	surg1248
Tamil	3	SOV	No	No	Yes	tami1289
Tarahumara (Western)	3	SOV	No	No	Yes	lowl1265
Tatar	3	SOV	No	No	Yes	tata1255
Telugu	3	SOV	No	No	Yes	telu1262
Thangmi	3	SOV	No	No	Yes	than1259
Thulung	3	SOV	No	No	Yes	thul1246
Tinani	3	SOV				tina1246
Torwali	3	SOV				torw1241
Toto	3		No	No	Yes	toto1302
Tsafiki	3	SOV	No	No	Yes	colo1256
Tsez	3	SOV	No	No	Yes	dido1241
Tsova-Tush	3	SOV	No	No	Yes	bats1242
Tulu	3	SOV	No	No	Yes	tulu1258
Turkish	3	SOV	No	No	Yes	nucl1301
Turkmen	3	SOV	No	No	Yes	turk1304
Udi	3	SOV	No	No	Yes	udii1243
Udihe	3	SOV	No	No	Yes	udih1248
Udmurt	3	SOV				udmu1245
Urali Irula	3		No	No	Yes	irul1244
Urdu	3	SOV				urdu1245
Usarufa	3	SOV				usar1243
Ute-Southern Paiute	3		No	No	Yes	utes1238

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Uzbek	3	SOV				nort2690, sout2699
Wakhi	3	SOV				wakh1245
Wambule	3	SOV	No	No	Yes	wamb1257
Washo	3	SOV	No	No	Yes	wash1253
Western Magar	3		No	No	Yes	west2418
Wolane	3		No	No	Yes	wola1253
Wolaytta	3	SOV	No	No	Yes	wola1242
Xamtanga	3		No	No	Yes	xamt1239
Yakkha	3		No	No	Yes	yakk1236
Yaqui	3	SOV	No	No	Yes	yaqu1251
Yemsa	3		No	No	Yes	yems1235
Yukaghir (Kolyma)	3	SOV				sout2750
Zayse	3	SOV	No	No	Yes	zays1235
Dutch [215]	3-V2	SOV or SVO	No	Yes	No	dutc1256
German [216]	3-V2	SOV or SVO				stan1295
Kashmiri [217]	3-V2	SVO	No	Yes	No	kash1277

A.4 Num-N-A-Dem

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Selayar [218]	0		Yes	Yes	Yes	sela1260
Alacatlatzala Mixtec	1		Yes	No	No	alac1244
Amuzgo (Guerrero variety)	1		Yes	No	No	guer1243
Ancient Hebrew	1		Yes	No	No	anci1244
Bajau (West Coast)	1	SVO or VSO	Yes	No	No	west2560
Batak (Toba)	1	VOS	Yes	No	No	bata1289

215 Classification established following S. Müller (2023a: 26-27).
216 Classification established following S. Müller (2023a: 26-27).
217 Classification established following Holmberg (2015) and Verbeke (2017).
218 See however Finer (1997: 679) and Basri (1999: 240 ff.) for the claim that any order that is not V-initial is pragmatically marked.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Chatino (Sierra Occidental)	1	VSO	Yes	No	No	west2644
Chatino (Yaitepec)	1	VSO				west2644
Chinantec (Comaltepec)	1	VSO				coma1246
Chinantec (Lealao)	1	VOS				leal1235
Chinantec (Palantla)	1	VSO				pala1351
Fijian [219]	1	VSO or VOS	No	Yes	No	fiji1243
Gaelic (Scots)	1	VSO				scot1245
Gela	1	VOS	Yes	No	No	gela1263
Goajiro	1	VSO	Yes	No	No	wayu1243
Hawu (Sawu)	1		Yes	No	No	sabu1255
Hoava	1	VSO	Yes	No	No	hoav1238
Irish [220]	1	VSO	Yes	Yes	No	iris1253
Kimarang	1		Yes	No	No	kima1244
Kiribati	1	VOS	No	Yes	No	gilb1244
Kokota	1	VSO	Yes	No	No	koko1269
Kubokota (Ghanongga superset)	1		Yes	No	No	ghan1242
Lengo	1		Yes	No	No	leng1259
Longgu	1	VOS	Yes	No	No	long1395
Magdalena Peñasco Mixtec	1		Yes	No	No	magd1235
Marquesan (northern variety)	1	VSO	Yes	No	No	nort2845
Mixtec (Chalcatongo)	1	VSO				sanm1296
Mixtec (Jicaltepec)	1	VSO				pino1237
Mixtec (Ocotepec)	1	VSO				ocot1243
Mixtec (Yosondúa)	1	VSO				yoso1239

[219] Rhiannon Schembri argues that Fijian is coded as a V-medial language on Grambank based on Dixon (1988: 33-34, 63). However, the latter source describes the language as V-initial. See also Dixon (1988: 47, 242,245).

[220] See discussion in Doyle (2001: 67 ff.).

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Mixtepec Zapotec (San Juan and San Pedro)	1		Yes	No	No	mixt1426
Niuean	1	VSO	Yes	No	No	niue1239
Popti’	1		Yes	No	No	popt1235
Quiegolani Zapotec	1		Yes	No	No	sant1451
Rapanui	1	VSO	Yes	No	No	rapa1244
Roviana	1	VSO	Yes	No	No	rovi1238
Sama Balangingi	1		Yes	No	No	bala1311
Sierra de Juárez Zapotec	1		Yes	No	No	sier1250
Sochiapam Chinantec	1		Yes	No	No	soch1239
Tacuate de Santa María Zacatepec	1		Yes	No	No	sant1436
Tamashek	1	VSO	Yes	No	No	tama1365
Tataltepec Chatino	1		Yes	No	No	tata1258
Tatana	1		Yes	No	No	tata1257
Tboli	1	VSO	Yes	No	No	tbol1240
Teotepec Chatino	1		Yes	No	No	nopa1235
Tezoatlán Mixtec	1		Yes	No	No	tezo1238
Timugon Murut	1	VSO	Yes	No	No	timu1262
Tlacolula Valley Zapotec (Western)	1		Yes	No	No	sanj1284
Tlapanec	1	VSO				acat1239
Tokelau	1		Yes	No	No	toke1240
Trique (Copala)	1	VSO	Yes	No	No	copa1237
Ughele	1		Yes	No	No	ughe1237
Welsh	1	VSO	Yes	No	No	wels1247
Wolio	1	VSO	Yes	No	No	woli1241
Yatzachi Zapotec	1		Yes	No	No	yatz1235
Zacatepec Chatino	1		Yes	No	No	zaca1242
Zapotec (Isthmus)	1	VSO				isth1244
Zapotec (Mitla)	1	VSO				mitl1236

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Zapotec (Zoogocho)	1	VSO	Yes	No	No	zoog1238
Batak (Karo)	1.5	SVO or VSO	Yes	Yes	No	bata1293
Bugis (Buginese variety) [221]	1.5	SVO	Yes	Yes	No	bugi1244
Gayo	1.5		Yes	Yes	No	gayo1244
Manggarai [222]	1.5	SVO	Yes	Yes	No	mang1405
Me'phaa (Malinaltepec variety)	1.5		Yes	Yes	No	mali1285
Pendau	1.5		Yes	Yes	No	pend1242
Ratahan	1.5		Yes	Yes	No	rata1244
Semaq Beri	1.5		Yes	Yes	No	sema1265
Semelai	1.5	VSO or VOS	Yes	Yes	No	seme1247
Tajio	1.5		Yes	Yes	No	taji1246
Tuvaluan [223]	1.5	OVS	Yes	No	No	tuva1244
Breton [224]	1-V2	SVO	Yes	No	No	bret1244
'Are'are	2		No	Yes	No	area1240
Achagua	2	SVO	No	Yes	No	acha1250
Arabic (Egyptian) [225]	2	SVO	Yes	No	No	egyp1253
Arosi	2	SVO	No	Yes	No	aros1241
Baba	2		No	Yes	No	baba1264
Balantak	2		No	Yes	No	bala1315
Barang-Barang (Laiyolo superset)	2		No	Yes	No	laiy1246
Barok	2		No	Yes	No	baro1253
Berber (Rif)	2	SVO				tari1263
Bonggi	2		No	Yes	No	bong1289
Brôu (Eastern variety)	2		No	Yes	No	east2332

221 See discussion in Pedrós Caballero (2014: 36).
222 See discussion in Salahuddin (2023: 189-190).
223 See discussion in Besnier (2002: 131-134).
224 Classification based on Holmberg (2015: Sec. 2.4).
225 Woidich (2006: Sec. 3.3.2) and Tvica (2017: 208) agree in arguing that the basic order is SVO. See Woidich (2006: Sec. 4.2.1.2.1, 4.3.2.5) for discussion of the marked VSO alternatives.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Brôu (Western variety)	2		No	Yes	No	west2397
Bunu (Younuo)	2	SVO	No	Yes	No	youn1235
Cheke Holo	2		No	Yes	No	chek1238
Chrau	2	SVO	No	Yes	No	chra1242
Cornish	2	SVO				corn1251
Cuoi (Hung variety)	2		No	Yes	No	hung1275
Dong (Southern)	2	SVO	No	Yes	No	sout2741
Eastern Cham	2	SVO	No	Yes	No	east2563
Gelao (Central Qau variety)	2	SVO	No	Yes	No	qaua1234
Haitian	2	SVO (APiCS)	No	Yes	No	hait1244
Halia	2	SVO	No	Yes	No	hali1244
Hebrew (Modern)	2	SVO	No	Yes	No	hebr1245
Hlai (Baoding)	2	SVO	No	Yes	No	hlai1239
Hmong Njua	2	SVO				hmon1264
Iban	2	SVO	No	Yes	No	iban1264
Indonesian	2	SVO	No	Yes	No	indo1316
Jahai	2		No	Yes	No	jeha1242
Jah-Hut	2		No	Yes	No	jahh1242
Jarai	2		No	Yes	No	jara1266
Jedek	2		No	Yes	No	jede1239
Jeh	2		No	Yes	No	jehh1245
Karitiâna	2		No	Yes	No	kari1311
Kensiu	2		No	Yes	No	kens1248
Kwaio	2	SVO	No	Yes	No	kwai1243
Kwaraqae	2		No	Yes	No	kwar1239
Lachi	2		No	Yes	No	lach1248
Lampung	2	SVO				lamp1243
Lun Bawang (Lundayeh superset)	2		No	Yes	No	lund1271
Maanyan	2		No	Yes	No	maan1238

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Madurese [226]	2		No	Yes	Yes	nucl1460
Mak (China)	2		No	Yes	No	makc1235
Maniq	2		No	Yes	No	tong1308
Maonan	2		No	Yes	No	maon1241
Marshallese	2		No	Yes	No	mars1254
Mateq	2		No	Yes	No	kemb1249
Mono-Alu [227]	2	No dom	No	Yes	Yes	mono1273
Muna	2	SVO	No	Yes	No	muna1247
Mussau	2	SVO	No	Yes	No	muss1246
Nehan	2	SVO	No	Yes	No	neha1247
Northern Roglai	2		No	Yes	No	nort2994
Nung (in Vietnam)	2	SVO				nung1283
Pacoh	2		No	Yes	No	paco1243
Phan Rang Chamb	2		No	Yes	No	east2563
Rejang Kayan	2		No	Yes	No	reja1241
Rengao	2		No	Yes	No	reng1252
Sa'a	2		No	Yes	No	saaa1240
Sanjiang Kam	2		No	Yes	No	caom1238
Sasak	2		No	Yes	No	sasa1249
Satawalese	2		No	Yes	No	sata1237
Sedang	2	SVO	No	Yes	No	seda1262
Seimat	2		No	Yes	No	seim1238
Siar (Siar-Lak)	2	SVO	No	Yes	No	siar1238
Sikaiana	2		No	Yes	No	sika1261
Sisiqa	2	SVO	No	Yes	No	sisi1250
Siwi	2		No	Yes	No	siwi1239
Sre (Koho superset)	2	SVO	No	Yes	No	koho1244
Sundanese	2	SVO				sund1252
Sursurunga	2		No	Yes	No	surs1246

[226] According to Davies (1999, 2005: 645-647), the basic word order is SVO. The V-final counterparts only occur with bare-stem verbs.

[227] According to Meier (2020: 14), preverbal objects occupy a focalized position.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Taiof (Saposa superset)	2	SVO	No	Yes	No	sapo1253
Temiar	2	SVO	No	Yes	No	temi1246
Tigak	2	SVO	No	Yes	No	tiga1245
Toqabaita	2		No	Yes	No	toab1237
Tungak	2	SVO	No	Yes	No	tung1290
Ulithian	2	SVO	No	Yes	No	ulit1238
Vietnamese	2	SVO	No	Yes	No	viet1252
Wala	2		No	Yes	No	wala1266
Wichí	2	SVO				wich1264
Woleaian	2	SVO	No	Yes	No	wole1240
Wolof	2	SVO	No	Yes	No	nucl1347
Wotu	2		No	Yes	No	wotu1240
Yay	2	SVO				bouy1240
Zoulei	2		No	Yes	No	aoua1234
Basque	3	SOV	No	No	Yes	basq1248

A.5 Dem-N-A-Num

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Tahitian	1	VSO	Yes	No	No	tahi1242
Miya [228]	1.5	SVO or VOS	Yes	No	No	miya1266
Akoose	2		No	Yes	No	akoo1248
Anufo	2	SVO	No	Yes	No	anuf1239
Bari	2	SVO				bari1284
Ga (Gã)	2	SVO	No	Yes	No	gaaa1244
Hkongso (Anu-Hkongso superset)	2		No	Yes	No	anuu1241
Kaket (Qaqet)	2		No	Yes	No	qaqe1238
Karbi (Amri variety)	2		No	Yes	No	amri1238
Lwena	2		No	Yes	No	luva1239

[228] See discussion in Dryer (2013f: Sec. 1) and references therein.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Mandara	2		No	Yes	No	mand1440
nDrapa	2		No	Yes	No	zhab1238
Nochi (Notsi)	2		No	Yes	No	nots1237
Pular	2		No	Yes	No	pula1262
Tiriyo	2	OVS	No	Yes	No	trio1238
Warao	2		No	Yes	No	wara1303
Amahuaca [229]	2.5	No dom	No	No	Yes	amah1246
(Dongwang) Tibetan	3		No	No	Yes	kham1282
(Haʔwa) Nocte	3	SOV	No	No	Yes	noct1238
(Lamjung) Yo(h)lmo	3		No	No	Yes	lamj1247
(Wobzi) Khroskyab	3		No	No	Yes	guan1266
Abui	3	SOV	No	No	Yes	abui1241
Akha Lai	3		No	No	Yes	haka1240
Angaataha	3		No	No	Yes	anga1290
Apatani	3	SOV	No	No	Yes	apat1240
Araona	3		No	No	Yes	arao1248
Awar	3		No	No	Yes	awar1249
Awtuw	3	SOV	No	No	Yes	awtu1239
Bambara	3	SOV	No	No	Yes	bamb1269
Bo Ung (Ku Waru dialect)	3		No	No	Yes	boun1245
Burmese	3	SOV	No	No	Yes	nucl1310
Cak	3		No	No	Yes	chak1270
Creek	3		No	No	Yes	cree1270
Crow	3	SOV	No	No	Yes	crow1244
Cuna (Border)	3	SOV	No	No	Yes	bord1248
Cuna (San Blas varieties)	3	SOV	No	No	Yes	sanb1242
Daai Chin	3		No	No	Yes	daai1236
Dimasa	3		No	No	Yes	dima1251

[229] The fact that Sparing-Chávez (2012) discusses V-medial sentences with "marked aspect on the verb" does not imply that they are pragmatically marked. See in particular Sparing-Chávez (2012: 5).

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Doromu-Koki	3		No	No	Yes	doro1266
Duhumbi (Chugpa)	3		No	No	Yes	chug1252
Dzongkha	3		No	No	Yes	dzon1239
Eastern Geshiza	3		No	No	Yes	east2851
Emberá (Northern)	3		No	No	Yes	nort2972
Ersu	3		No	No	Yes	ersu1241
Falam Chin	3		No	No	Yes	fala1243
Fulniô	3		No	No	Yes	fuln1247
Fuyug	3		No	No	Yes	fuyu1242
Garo	3	SOV	No	No	Yes	garo1247
Gugu-Yalanji	3		No	No	Yes	kuku1273
Guiqiong	3		No	No	Yes	guiq1238
Guugu-Yimidjir	3	SOV	No	No	Yes	gugu1255
Gyarong (Cogtse)	3		No	No	Yes	situ1238
Haida	3	SOV	No	No	Yes	nort2938, sout2956
Hruso	3		No	No	Yes	hrus1242
Hualapai	3	SOV	No	No	Yes	hava1248
Iduna	3	SOV	No	No	Yes	idun1242
Iwam	3	SOV	No	No	Yes	iwam1256
Kabardian	3	SOV	No	No	Yes	kaba1278
Kadu (Asak)	3		No	No	Yes	kado1242
Kaki Ae	3	SOV	No	No	Yes	kaki1249
Karbi (Hills variety)	3		No	No	Yes	karb1241
Kharam	3		No	No	Yes	khar1288
Kokborok	3	SOV				kokb1239
Konyak	3		No	No	Yes	kony1248
Koromu	3		No	No	Yes	kesa1244
Kulina	3		No	No	Yes	culi1244
Kuman	3	SOV	No	No	Yes	kuma1280
Kurtöp	3		No	No	Yes	kurt1248
Kyirong	3		No	No	Yes	kyer1238

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Lacid	3		No	No	Yes	lash1243
Manange	3		No	No	Yes	mana1288
Mandinka (Gambian)	3	SOV	No	No	Yes	mand1436
Maru (Lhaovo)	3	SOV	No	No	Yes	maru1249
Mekeo	3		No	No	Yes	meke1243
Mishmi	3		No	No	Yes	miju1243
Mizo	3	SOV	No	No	Yes	lush1249
Mona	3		No	No	Yes	mwan1250
Motu	3	SOV	No	No	Yes	motu1246
Motuna	3		No	No	Yes	siwa1245
Muisca	3	SOV	No	No	Yes	chib1270
Munya	3		No	No	Yes	west2417
Nankina	3	SOV				nank1250
Nimboran	3		No	No	Yes	nucl1633
Nubian (Dongolese)	3	SOV	No	No	Yes	dong1288 (kenu1236)
Nubian (Kunuz)	3	SOV	No	No	Yes	kenu1243 (kenu1236)
Ömie	3		No	No	Yes	omie1241
Orokolo	3		No	No	Yes	orok1267
Paite	3		No	No	Yes	pait1244
Pech (Pesh)	3		No	No	Yes	pech1241
Pumi	3	SOV				pumi1242
rGyalrong (Caodeng)	3	SOV				zbua1234
rGyalrong (Japhug)	3		No	No	Yes	japh1234
Rongmei	3		No	No	Yes	rong1266
Sandawe	3	SOV	No	No	Yes	sand1273
Sentani	3	SOV	No	No	Yes	nucl1632
Sherdukpen	3		No	No	Yes	sher1257
Shǐxīng	3		No	No	Yes	shix1238
Sinaugoro	3	SOV	No	No	Yes	sina1266

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Sucite Senoufo	3		No	No	Yes	sici1249
Supyire Senoufo	3	SOV	No	No	Yes	supy1237
Tairuma (Toaripi)	3		No	No	Yes	toar1246
Tama	3	SOV	No	No	Yes	tama1331
Tangut	3		No	No	Yes	tang1334
Tauya	3	SOV	No	No	Yes	tauy1241
Tawala	3	SOV	No	No	Yes	tawa1275
Tayap (or Taiap)	3		No	No	Yes	taia1239
Thakali	3	SOV	No	No	Yes	thak1245
Tiddim Chin	3		No	No	Yes	tedi1235
Tujia (Northern)	3		No	No	Yes	nort2732
Tunen	3	SOV	No	No	Yes	tune1261
Ubir	3		No	No	Yes	ubir1237
Umbu-Ungu	3		No	No	Yes	umbu1258
Wambon	3	SOV	No	No	Yes	wamb1259
Xasonga	3	SOV				xaas1235
Yaminawa	3	SOV	No	No	Yes	yami1256
Yélî Dnye [230]	3	SOV	Yes	Yes	Yes	yele1255
Yessan-Mayo	3		No	No	Yes	yess1239
Youle Jino	3		No	No	Yes	youl1235
Yuchi	3	SOV	No	No	Yes	yuch1247
Yuqui	3		No	No	Yes	yuqu1240
Zeme	3		No	No	Yes	zeme1240

[230] See Levinson (2022: 7, 22-23, 253-254).

A.6 Dem-Num-N-A

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Karok	0	No dom	Yes	Yes	Yes	karo1304
Mbyá Guaraní [231]	0		Maybe	Yes	Yes	mbya1239
Baniwa	1		Yes	No	No	bani1255
Baure	1	VOS	Yes	No	No	baur1253
Berber (Figuig)	1	VSO				tazn1238
Cora	1	VSO	Yes	No	No	elna1235
Curripaco	1		Yes	No	No	curr1243
Guajajara	1	VSO				guaj1255
Guató	1	VSO	Yes	No	No	guat1253
Hawaiian	1	VSO	Yes	No	No	hawa1245
Kipea	1	VOS				kari1255
Kuot	1	VSO	Yes	No	No	kuot1243
Mam	1	VSO	Yes	No	No	mamm1241
Nahuatl (Huasteca)	1	VSO	Yes	No	No	west2624
Ottawa	1		Yes	No	No	otta1242
Pnar	1		Yes	No	No	pnar1238
Pukapuka	1		Yes	No	No	puka1242
Quiotepec Chinantec	1	VSO	Yes	No	No	quio1240
Yagua [232]	1	No dom	Yes	No	No	yagu1244
Yalarnnga	1		Yes	No	No	yala1262
Arabic (Modern Standard) [233]	1.5	VSO	Yes	No	No	stan1318
Asháninka	1.5		Yes	Yes	No	asha1243
Central Alaskan Yupik	1.5		Yes	Yes	No	cent2127
Gulf Arabic [234]	1.5	SVO	Yes	Yes	No	gulf1241

[231] See also Martins (2003: 118-119).
[232] According to Payne & Payne (1990: 259) an argument can appear preverbally only when focalized or topicalized.
[233] See also Alkuwaihes (2017: 28-29).
[234] See also Alkuwaihes (2017: 29).

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Albanian	2	SVO	No	Yes	No	tosk1239, gheg1238
Baré	2	SVO	No	Yes	No	bare1276
Biri [235]	2	SVO	No	No	Yes	biri1256
Catalan	2	SVO	No	Yes	No	stan1289
Cocama	2	SVO	No	Yes	No	coca1259
French	2	SVO	No	Yes	No	stan1290
Guaraní	2	SVO				para1311
Italian	2	SVO	No	Yes	No	ital1282
Khasi	2	SVO	No	Yes	No	khas1269
Madak	2		No	Yes	No	mada1285
Maltese	2		No	Yes	No	malt1254
Manadonese	2	SVO	No	Yes	No	mala1481
Martuthunira	2	SVO	No	Yes	No	mart1255
Mien	2	SVO	No	Yes	No	iumi1238
Nahuatl (Michoacán) [236]	2	SVO	Yes	Yes	No	mich1245
Nahuatl (Tetelcingo)	2	SVO	No	Yes	No	tete1251
Palikur	2	SVO	No	Yes	No	pali1279
Pame (Southern)	2	SVO				sout2999
Portuguese	2	SVO	No	Yes	No	port1283
Romansch (Sursilvan)	2	SVO	No	Yes	No	roma1326
Spanish	2	SVO				stan1288
Tok Pisin	2	SVO (APiCS)	No	Yes	No	tokp1240
Urarina	2	OVS	No	Yes	No	urar1246
Wardaman [237]	2	No dom	No	Yes	No	ward1246
Warekena	2	SVO	No	Yes	No	ware1255

[235] Terrill (1998: 50-51) provides mono- and di-transitive clauses that are neither V-final nor V-initial.

[236] Grambank classifies this language following Sischo & Hollenbach (2005: 14-15), albeit no example where subject and object are both non-pronominal is provided. WALS refers instead to Sischo (1979: 313), where such an example is provided together with a general template.

[237] Merlan (1994: 248-250) shows that, whenever both arguments of a mono-transitive clause are overtly realized, either one of the two precedes the verb. The classification given on WALS (Dryer 2013f) follows from the frequencies discussed in Merlan (1994: 248), where the verb commonly appears as the first item (with a prefix co-indexed with the thematic agent) due to the fact that at least one of the two arguments is not overtly realized.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Purépecha [238]	2.5	SVO	No	No	Yes	pure1242
Boruca	3		No	No	Yes	boru1252
Carib (De'kwana)	3	SOV	No	No	Yes	maqu1238
Cupeño	3	SOV	No	No	Yes	cupe1243
Dâw	3	SOV				daww1239
Émérillon	3	SOV				emer1243
Ese Ejja	3	SOV	No	No	Yes	esee1248
Hup	3	SOV	No	No	Yes	hupd1244
Kamaiurá	3	SOV				kama1373
Kanoê	3	SOV				kano1245
Kiowa	3	SOV	No	No	Yes	kiow1266
Kotiria	3		No	No	Yes	guan1269
Kumzari	3		No	No	Yes	kumz1235
Latin	3		No	No	Yes	lati1261
Miskito	3	SOV	No	No	Yes	misk1235
Navajo	3	SOV	No	No	Yes	nava1243
Neo-Aramaic (Arbel Jewish)	3	SOV	No	No	Yes	lish1245
Ofayé	3		No	No	Yes	ofay1240
Páez	3	SOV	No	No	Yes	paez1247
Persian	3	SOV	No	No	Yes	west2369
Sarcee	3	SOV	No	No	Yes	sars1236
Shiriana	3	SOV	No	No	Yes	nina1238
Tajik	3	SOV				taji1245
Ubykh	3	SOV	No	No	Yes	ubyk1235
Urubú-Kaapor	3	SOV				urub1250
Yanomamö	3		No	No	Yes	yano1261
Zazaki	3	SOV				diml1238
Zuni	3	SOV	No	No	Yes	zuni1245

[238] See Maldonado (2012: 11-12).

A.7 N-A-Dem-Num

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Anejom [239]	1	VOS	Yes	Yes	No	anei1239
East Uvean	1		Yes	No	No	wall1257
Natqgu	1		Yes	No	No	natu1246
Ngas	1		Yes	No	No	ngas1240
Parauk	1.5		Yes	Yes	No	para1301
Aghem	2	SVO	No	Yes	No	aghe1239
Araki	2		No	Yes	No	arak1252
Babungo	2	SVO				veng1238
Berta	2	SVO	No	Yes	No	bert1248
Bierebo	2		No	Yes	No	bier1244
Dusner	2		No	Yes	No	dusn1237
Ega	2	SVO	No	Yes	No	egaa1242
Farefare (Gurune)	2		No	Yes	No	fare1241
Gemzek (Zulgo-Gemzek superset)	2		No	Yes	No	zulg1242
Kaulong	2	SVO	No	Yes	No	kaul1240
Kayan	2		No	Yes	No	kaya1315
Kom	2		No	Yes	No	komc1235
Lote	2		No	Yes	No	lote1237
Lusi	2	SVO	No	Yes	No	lusi1240
Maskelynes	2		No	Yes	No	mask1242
Mavea	2		No	Yes	No	mafe1237
Mmen	2		No	Yes	No	mmen1238
Mundabli	2		No	Yes	No	mund1328
Mwani	2		No	Yes	No	mwan1247
Neve'ei	2		No	Yes	No	vinm1237
Nubi	2	SVO (APiCS)	No	Yes	No	nubi1253
Nupe	2	SVO	No	Yes	No	nupe1254

[239] Lynch (2000: 114-118) never mentions SVO as a possible word order in mono-transitive clauses, *contra* what is argued on Grambank based on the same source.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Port Sandwich	2	SVO	No	Yes	No	port1285
Sakao	2	SVO	No	Yes	No	saka1289
Sisaala	2	SVO	No	Yes	No	tumu1242
Tama(m)bo	2	SVO	No	Yes	No	malo1243
Tibetan (Amdo)	2		No	Yes	No	amdo1237
Uldeme	2	SVO				wuzl1236
Wamesa	2		No	Yes	No	wand1267
Warembori	2	SVO	No	Yes	No	ware1253
West Makian	2	SVO	No	Yes	No	west2600
Wooi	2		No	Yes	No	woii1237
Akha	3	SOV	No	No	Yes	akha1245
Amam	3		No	No	Yes	amam1246
Anong	3	SOV	No	No	Yes	nung1282
Awara	3		No	No	Yes	awar1248
Beria	3	SOV	No	No	Yes	zagh1240
Bisu	3	SOV	No	No	Yes	bisu1244
Hani	3		No	No	Yes	hani1248
Jamsay	3	SOV	No	No	Yes	jams1239
Khatso	3		No	No	Yes	kats1235
Kobon	3	SOV	No	No	Yes	kobo1249
Koiari	3		No	No	Yes	gras1249
Lalo	3	SOV	No	No	Yes	xish1235
Lisu	3	SOV	No	No	Yes	hual1240
Lolo (Yao'An)	3		No	No	Yes	lolo1259
Maisin	3	SOV	No	No	Yes	mais1250
Maranungku	3	SOV	No	No	Yes	mara1386
Naga (Mao)	3	SOV				maon1238
Namuyi Khato	3		No	No	Yes	namu1246
Naxi	3		No	No	Yes	naxi1245
Panama (Naso)	3		No	No	Yes	teri1250
Qiang	3		No	No	Yes	nort2722

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Roro (Waima superset)	3		No	No	Yes	waim1251
Sema	3		No	No	Yes	sumi1235
Soninke	3	SOV	No	No	Yes	soni1259
Takia	3	SOV	No	No	Yes	taki1248
Tibetan (Modern Literary)	3	SOV	No	No	Yes	tibe1272
Tibetan (Standard Spoken)	3	SOV	No	No	Yes	tibe1272

A.8 N-Dem-Num-A

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Warrongo	1		Yes	No	No	waru1264
Deiga	1		Yes	No	No	keig1242
Ik	1	VSO	Yes	No	No	ikkk1242
So	1	VSO	Yes	No	No	sooo1256
Murle	1	VSO	Yes	No	No	murl1244
Kipsigis	1		Yes	No	No	kips1239
Majang	1	VSO	Yes	No	No	maja1242
Kiuru	2	SVO				moch1256
Nyakyusa	2		No	Yes	No	nyak1260
Ikihehe	2	SVO	No	Yes	No	hehe1240
Lingala	2	SVO	No	Yes	No	ling1263
Nchane	2		No	Yes	No	ncan1245
Chimakonde	2	SVO	No	Yes	No	mako1251
Cuwabo	2		No	Yes	No	chuw1238
Ekegusii	2		No	Yes	No	gusi1247
Gîtigania (Meru superset)	2		No	Yes	No	tiga1244, meru1245
Gyeli	2		No	Yes	No	gyel1242
Igembe (Meru superset)	2		No	Yes	No	igem1238, meru1245

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Ikalanga (Kalanga superset)	2		No	Yes	No	ikal1242, kala1384
Kagulu	2		No	Yes	No	kagu1239
Kemezung	2		No	Yes	No	keme1240
Kîîtharaka	2		No	Yes	No	thar1283
Kikamba	2		No	Yes	No	kamb1297
Kikuyu	2	SVO	No	Yes	No	kiku1240
Kiyaka	2		No	Yes	No	yaka1269
Kivunjo	2		No	Yes	No	vunj1238
Lubukusu	2		No	Yes	No	buku1249
Ndengeleko	2		No	Yes	No	nden1248
Nomaande	2		No	Yes	No	noma1260
Setswana	2		No	Yes	No	tswa1253
Shona	2	SVO	No	Yes	No	shon1251
Yasa	2		No	Yes	No	yasa1242
Xitsonga	2		No	Yes	No	tson1249
Chimiini	2		No	Yes	No	chim1312
Zulu	2	SVO	No	Yes	No	zulu1248
Moro	2	SVO	No	Yes	No	moro1285
Arapesh (Abu)	2	SVO	No	Yes	No	abua1245
Asheron	2		No	Yes	No	ache1245
Namakir	2		No	Yes	No	nama1268
Paamese	2	SVO	No	Yes	No	paam1238
Arbore [240]	3	SOV	No	Yes	No	arbo1245
Elmolo	3		No	No	Yes	elmo1238
Rendille	3		No	No	Yes	rend1243
Somali	3	SOV	No	No	Yes	soma1255
Iraqw	3	SOV	No	No	Yes	iraq1241
Gorowa	3		No	No	Yes	goro1270
Laragia	3		No	No	Yes	lara1258

[240] See Hayward (1984: 110, 240) for discussion supporting the classification that WALS (Dryer 2013f) applies for Arbore.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Wutun	3		No	No	Yes	wutu1241
Ongota	3		No	No	Yes	bira1253

A.9 Dem-A-N-Num

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Laguna Keres (Western Keres superset)	0		Yes	Yes	Yes	lagu1251 (west2632)
Cotabato Manobo	1		Yes	No	No	cota1241
Arapaho	1.5		Yes	Yes	No	arap1274
Hausa	2	SVO	No	Yes	No	haus1257
Zande	2	SVO	No	Yes	No	zand1248
Ambulas [241]	3	SOV	No	Yes	Yes	ambu1247
Balti	3	SOV				balt1258
Bodo	3		No	No	Yes	bodo1269
Chang	3	SOV				chan1313
Chiru	3		No	No	Yes	chir1283
Dadibi	3		No	No	Yes	dadi1250
Dulong	3		No	No	Yes	drun1238
Fore	3	SOV	No	No	Yes	fore1270
Gurung	3	SOV	No	No	Yes	guru1261
Hanga Hundi	3	SOV	No	No	Yes	hang1263
Hoozo	3		No	No	Yes	hozo1236
Iatmul	3	SOV	No	No	Yes	iatm1242
Jingpo	3	SOV	No	No	Yes	kach1280
Kachari	3	SOV				kach1279
Kamano	3		No	No	Yes	kama1370
Kewa (both East and West)	3	SOV	No	No	Yes	east2516, west2599
Korean	3	SOV	No	No	Yes	kore1280

[241] See P. R. Wilson (1980: 186) for discussion favoring the classification provided by WALS (Dryer 2013f).

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Kwomtari	3	SOV	No	No	Yes	nucl1593
Liangmai	3		No	No	Yes	lian1251
Manambu	3		No	No	Yes	mana1298
Mising	3	SOV				misi1242
Newari (Kathmandu)	3	SOV	No	No	Yes	newa1246
Nivkh	3	SOV	No	No	Yes	gily1242
Nyishi	3		No	No	Yes	nyis1236
Oriya	3	SOV	No	No	Yes	oriy1255
Pawaia	3	SOV	No	No	Yes	pawa1255
Purki	3	SOV	No	No	Yes	puri1258
Rabha	3		No	No	Yes	rabh1238
Shipibo-Konibo	3	SOV	No	No	Yes	ship1254
Tairora	3	SOV				nort2920
Tamang (Eastern)	3	SOV				east2347
Tshangla	3	SOV	No	No	Yes	tsha1245
Yagaria	3	SOV	No	No	Yes	yaga1260 (huaa1250)
Yuwan	3		No	No	Yes	sout2954

A.10 N-Num-A-Dem

Language	Type	WALS	V-I	V-M	V-F	Glottocode
ǂHőaã	2	SVO	No	Yes	No	hoaa1235
Baggara Arabic	2		No	Yes	No	chad1249
Baraïn (Barayin)	2		No	Yes	No	bare1279
Bebele	2		No	Yes	No	bebe1248
Buma	2	SVO	No	Yes	No	tean1237
Chimpoto	2		No	Yes	No	mpot1240
Erromangan (Sye)	2	SVO	No	Yes	No	siee1239
Esan	2		No	Yes	No	esan1238
Goemai	2	SVO	No	Yes	No	goem1240

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Jejara	2		No	Yes	No	para1302
Kele	2	SVO	No	Yes	No	kele1258
Labwor	2		No	Yes	No	labw1238
Lango	2	SVO	No	Yes	No	lang1324
Larantuka Malay	2		No	Yes	No	lara1260
Lele	2	SVO	No	Yes	No	lele1270
Lou	2	SVO	No	Yes	No	louu1245
Luo	2	SVO	No	Yes	No	luok1236
Makhuwa-Enahara	2		No	Yes	No	makh1264
Migili	2		No	Yes	No	liji1238
Shimwela (Mwera)	2	SVO	No	Yes	No	mwer1248
Sobei	2	SVO	No	Yes	No	sobe1238
Sulka	2	SVO	No	Yes	No	sulk1246
Tiv	2	SVO	No	Yes	No	tivv1240
Inanwatan	2.5	SOV or SVO	No	Yes	Yes	suab1238
(Costa Rica) Teribe	3	SOV	No	No	Yes	teri1250
Buglere	3		No	No	Yes	bugl1243
Dahalo	3		No	No	Yes	daha1245
Dan	3	SOV			Yes	dann1241
Hakha Chin	3		No	No	Yes	haka1240
Kalamang	3		No	No	Yes	kara1499
Lavukaleve	3	SOV	No	No	Yes	lavu1241
Maba	3	SOV	No	No	Yes	maba1277
Makasae	3	SOV			Yes	maka1316
Oromo (Waata)	3	SOV			Yes	waat1238

A.11 Num-A-N-Dem

Language	Type	WALS	V-I	V-M	V-F	Glottocode
(Takituduh) Bunun	1		Yes	Maybe	No	bunu1267
Chiquimulilla	1		Yes	No	No	xinc1242
Chuj	1	VOS	Yes	No	No	chuj1250

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Garifuna	1		Yes	No	No	gari1256
Guazacapán	1		Yes	No	No	xinc1246
Itzá	1		Yes	No	No	itza1241
Jakaltek	1	VSO				kanj1263
Jumaytepeque	1		Yes	No	No	xink1235
Kankanaey	1		Yes	No	No	kank1243
Kavalan	1		Yes	No	No	kava1241
Mayan Quiché	1		Yes	No	No	kich1262
Nisgha	1	VSO	Yes	No	No	nisg1240
Q'anjob'al	1		Yes	No	No	qanj1241
Taushiro	1	VSO	Yes	No	No	taus1253
Tsimshian (Coast)	1	VSO				coas1300
Tsou	1		Yes	No	No	tsou1248
Yupiltepeque	1		Yes	No	No	xinc1243
Tinrin [242]	1.5	SVO or VOS	Yes	No	No	tiri1258
Berbice Dutch Creole	2	SVO (APiCS)	No	Yes	No	berb1259
Bislama	2	SVO (APiCS)	No	Yes	No	bisl1239
Eleme	2		No	Yes	No	elem1253
Gokana	2	SVO				goka1239
Ho	2	SVO (APiCS)	No	No	Yes	sara1340
Ndyuka	2	SVO				suri1272
Nyam	2		No	Yes	No	nyam1285
Ogbronuagum	2	SVO				ogbr1243
Puluwatese	2	SVO	No	Yes	No	pulu1242
Ramoaaina	2		No	Yes	No	ramo1244
Xârâcùù	2	SVO	No	Yes	No	xara1244
Northern Gumuz	2.5		No	Yes	Yes	gumu1244

[242] According to Osumi (1995: 254), in Tinrin the unmarked order is shifting from VOS to SVO.

A.12 Dem-N-Num-A

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Arusa (Maasai superset)	1	VSO	Yes	No	No	masa1300
Bakueri	2	SVO	No	Yes	No	mokp1239
Chamacoco [243]	2		Yes	Yes	No	cham1315
Digo	2		No	Yes	No	digo1243
Ha	2		No	Yes	No	haaa1252
Kupang Malay	2		No	Yes	No	kupa1239
Langi	2	SVO	No	Yes	No	lang1320
Manado Malay	2	SVO	No	Yes	No	mala1481
Xhosa	2	SVO	No	Yes	No	xhos1239
Chimariko	2.5		No	Yes	Yes	chim1301
Fur	3	SOV	No	No	Yes	furr1244
Girawa	3		No	No	Yes	gira1247
Kogui	3		No	No	Yes	cogu1240
Koorete	3		No	No	Yes	koor1239
Lahu	3	SOV	No	No	Yes	lahu1253
Retuarã	3	SOV	No	No	Yes	tani1257
Slave	3	SOV				slav1253
Una	3	SOV	No	No	Yes	unaa1239
Wappo	3	SOV	No	No	Yes	wapp1239
Yidiny	3	SOV	No	No	Yes	yidi1250

A.13 N-Dem-A-Num

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Anêm	2	SVO	No	Yes	No	anem1249
Baga Sitemu	2	SVO	No	Yes	No	baga1272
Diola-Fogny	2	SVO				jola1263

[243] According to Luca Ciucci, Chamacoco has a strict SVO word order. See the classification on SAILS (Hammarstrom 2014) (https://sails.clld.org/languages/ceg).

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Kove	2		No	Yes	No	kove1237
Mor	2	SVO	No	Yes	No	morm1235
Nkore-Kiga	2	SVO	No	Yes	No	chig1238
Noon	2	SVO	No	Yes	No	noon1242
Nyoro	2	SVO	No	Yes	No	nyor1246
Runyankore	2	SVO	No	Yes	No	nyan1307
Sie	2		No	Yes	No	siee1239
Tiene	2		No	Yes	No	tien1242
Koluwawa	3		No	No	Yes	kolu1245
Kom	3		No	No	Yes	komi1270
Konso	3		No	No	Yes	kons1243

A.14 A-N-Num-Dem

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Gude [244]	1	VSO	Yes	Yes	No	gude1246
Apma	2		No	Yes	No	apma1240
Eton	2		No	Yes	No	eton1253
Ibibio	2		No	Yes	No	ibib1240
Izi	2	SVO				izie1238
Makaa	2	SVO	No	Yes	No	maka1304
Mayogo	2	SVO	No	Yes	No	mayo1261
Ngbaka (Ma'bo)	2	SVO	No	Yes	No	ngba1284
Ngbandi (Northen variety)	2	SVO	No	Yes	No	nort2774
Paluai	2		No	Yes	No	balu1257
Sango	2	SVO				sang1328
Sunwadia	2		No	Yes	No	mari1426
Tuki	2	SVO	No	Yes	No	tuki1240

[244] Rijkhoff (2002: 304-305) and Cinque (2005, 2021) follow Hoskinson (1983: 51-53) in arguing that A-N-Num-Dem is the order of U20 found in Gude, which is in turn reported as a VSO (i.e. type 1) language in pragmatically unmarked main clauses not only on Grambank, but also in works like Canac Marquis (1989), Julien (2002: 334) and M. C. Baker (2002: 324), with seemingly SVO alternate orders possible only in questions, with negation, and in pragmatically marked contexts.

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Vaeakau-Taumako	2		No	Yes	No	pile1238
Southern Gumuz	2.5		No	Yes	Yes	sout3236
Lamkang	3		No	No	Yes	lamk1238
Menggwa Dla	3	SOV	No	No	Yes	dera1245
Milang	3		No	No	Yes	mila1245
Oksapmin	3	SOV	No	No	Yes	oksa1245

A.15 A-N-Dem-Num

Language	Type	WALS	V-I	V-M	V-F	Glottocode
Babanki	2		No	Yes	No	baba1266
Bai	2	SVO	No	Yes	No	cent2004, nort2724, sout2730
Baka (in Cameroon)	2	SVO	No	Yes	No	baka1272
Belanda Bor	2		No	Yes	No	bela1256
Gbaya (Northwestern)	2		No	Yes	No	nort2775
Kol	2		No	Yes	No	kolc1235
Linda (Banda-Bambari superset)	2	SVO	No	Yes	No	band1344
Mɔnɔ	2		No	Yes	No	mono1270
Munzombo	2		No	Yes	No	monz1249
Pa'anci	2	SVO	No	Yes	No	paaa1242
Zulgo (Zulgo-Gemzek superset)	2		No	Yes	No	zulg1242
Puxi Qiang	3		No	No	Yes	sout2728

Bibliography

Abeillé, Anne. 2024. The empirical turn and its consequences for theoretical syntax. *Theoretical Linguistics* 50(1–2). 5–21. DOI: 10.1515/tl-2024-2011.

Abeillé, Anne, Kathleen Bishop, Sharon Cote & Yves Schabes. 1990. *A Lexicalized Tree Adjoining Grammar for English*. Technical Reports (CIS) MS-CIS-90-24. Philadelphia (PA): University of Pennsylvania. https://www.researchgate.net/publication/45599403_A_Lexicalized_Tree_Adjoining_Grammar_for_English.

Abeillé, Anne & Owen Rambow. 2000. Tree Adjoining Grammar: An overview. In Anne Abeillé & Owen Rambow (eds.), *Tree Adjoining Grammars: Formalisms, linguistic analyses and processing*, 1–68. Stanford (CA): CSLI Publications.

Abels, Klaus. 2007. Towards a restrictive theory of (remnant) movement! *Linguistic Variation Yearbook* 7(1). 53–120. DOI: 10.1075/livy.7.04abe.

Abels, Klaus. 2015. Word Order. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax - Theory and Analysis: An International Handbook*, vol. 2, 1400–1448. Berlin ; München ; Boston: De Gruyter Mouton. DOI: 10.1515/9783110363708-017.

Abels, Klaus & Ad Neeleman. 2009. Universal 20 without the LCA. In José M. Brucart, Anna Gavarró & Jaume Solà (eds.), *Merging Features: Computation, Interpretation, and Acquisition*, 60–79. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199553266.003.0004.

Abels, Klaus & Ad Neeleman. 2010. Nihilism masquerading as progress. *Lingua* 120(12). 2657–2660. DOI: 10.1016/j.lingua.2010.03.008.

Abels, Klaus & Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15(1). 25–74. DOI: 10.1111/j.1467-9612.2011.00163.x.

Abney, Steven P. 1986. Functional elements and licensing. https://web.archive.org/web/20191124074559/http://vinartus.net/spa/86b.pdf.

Abney, Steven P. 1987. *The English noun phrase in its sentential aspect*. Massachusetts Institute of Technology dissertation. http://hdl.handle.net/1721.1/14638.

Aboh, Enoch Oladé. 2004. The Syntax of Noun Phrases. In Enoch Oladé Aboh (ed.), *The Morphosyntax of Complement-Head Sequences: Clause Structure and Word Order Patterns in Kwa*. Oxford University Press. DOI: 10.1093/acprof:oso/9780195159905.003.0003.

Aboh, Enoch Oladé. 2010. The morphosyntax of the noun phrase. In Enoch Oladé Aboh & James Essegbey (eds.), *Topics in Kwa syntax*, vol. 78, 11–37. Dordrecht: Springer. DOI: 10.1007/978-90-481-3189-1_2.

Ackema, Peter & Ad Neeleman. 2002. Effects of Short-Term Storage in Processing Rightward Movement. In Sieb Nooteboom, Fred Weerman & Frank Wijnen (eds.), *Storage and Computation in the Language Faculty*, vol. 30, 219–256. Dordrecht: Kluwer. DOI: 10.1007/978-94-010-0355-1_8.

Adam, Nina & Andreas Hölzl. 2024. Subjects in Word Order and Alignment Typology. In Artemis Alexiadou, Doreen Georgi, Fabian Heck, Gereon Müller & Florian Schäfer (eds.), *Gisbert Fanselow's Contributions to Syntactic Theory*, 311–322. Leipzig: Institut für Linguistik Universität Leipzig. https://www.philol.uni-leipzig.de/fileadmin/Fakult%C3%A4t_Philo/Linguistik/Forschung/LAB/LAB_96/LAB96.pdf.

Adger, David. 2013. *A syntax of substance*. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/9780262018616.001.0001.

Aissen, Judith L. 1992. Topic and Focus in Mayan. *Language* 68(1). 43–80. DOI: 10.2307/416369.

Alexeyenko, Sascha & Hedde Zeijlstra. 2021a. (Dis)obeying the Head-Final Filter. In D. K. E. Reisinger & Marianne Huijsmans (eds.), *Proceedings of the 37th West Coast Conference on Formal Linguistics*, 50–58. Somerville (MA): Cascadilla Press. https://www.lingref.com/cpp/wccfl/37/abstract3514.html.

Alexeyenko, Sascha & Hedde Zeijlstra. 2021b. Linearization of complex modifiers: Ways of (dis)obeying the Head-Final Filter. Ms. University of Göttingen. https://www.heddezeijlstra.org/_files/ugd/533bc0_3fd3566d02e1454eb20742ca77492ac6.pdf.

Alexiadou, Artemis. 2014. The syntax of adjectives. In Andrew Carnie, Daniel Siddiqi & Yosuke Sato (eds.), *The Routledge Handbook of Syntax*, 107–125. Abingdon ; New York: Routledge. DOI: 10.4324/9781315796604-13.

Alexiadou, Artemis. 2022. Obligatory subjects and the SVO/SOV debate. *Theoretical Linguistics* 48(1–2). 41–48. DOI: 10.1515/tl-2022-2029.

Alexiadou, Artemis & Elena Anagnostopoulou. 2001. The subject-in-situ generalization and the role of case in driving computations. *Linguistic inquiry* 32(2). 193–231. DOI: 10.1162/00243890152001753.

Alexiadou, Artemis & Elena Anagnostopoulou. 2007. The Subject-In-Situ Generalization Revisited. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + Recursion = Language?*, 31–60. Berlin ; New York: De Gruyter Mouton. DOI: 10.1515/9783110207552.31.

Alexiadou, Artemis, Liliane M. V. Haegeman & Melita Stavrou. 2007. *Noun phrase in the generative perspective*. Berlin ; New York: De Gruyter Mouton. DOI: 10.1515/9783110207491.

Alfieri, Luca, Giorgio Francesco Arcodia & Paolo Ramat. 2021. Linguistic categories, language description and linguistic typology – An overview. In Luca Alfieri, Giorgio Francesco Arcodia & Paolo Ramat (eds.), *Linguistic Categories, Language Description and Linguistic Typology*, 1–34. Amsterdam ; Philadelphia: John Benjamins. DOI: 10.1075/tsl.132.01alf.

Alkuwaihes, Ghada Ahmed A. 2017. *The Syntax of Particle-Based Exclamatives in Gulf Arabic*. York: University of York. https://drepo.sdl.edu.sa/handle/20.500.14154/34897.

Allassonnière-Tang, Marc & One-Soon Her. 2020. Numeral base, numeral classifier, and noun: Word order harmonization. *Language and Linguistics.* □□□□□□ 21(4). 511–556. DOI: 10.1075/lali.00069.all.

Arikawa, Koji. 2013. Is word order asymmetry mathematically expressible? *Biolinguistics* 7. 276–300. DOI: 10.5964/bioling.8967.

Arkoh, Ruby & Lisa Matthewson. 2013. A familiar definite article in Akan. *Lingua* 123. 1–30. DOI: 10.1016/j.lingua.2012.09.012.

Artiagoitia, Xabier. 2006. Basque adjectives and the functional structure of the noun phrase. *Anuario del Seminario de Filología Vasca "Julio de Urquijo"* 40(1–2). 107–132. DOI: 10.1387/asju.4376.

Artiagoitia, Xabier. 2008. Some arguments for complement-head order in Basque DPs. *Anuario del Seminario de Filología Vasca "Julio de Urquijo"*. 71–92. https://ojs.ehu.eus/index.php/ASJU/article/download/9767/9199.

Artiagoitia, Xabier. 2010. Basque Genitive Case and Multiple Checking. *Annual Meeting of the Berkeley Linguistics Society*. 1–15. DOI: 10.3765/bls.v36i1.3899.

Baggio, Pietro. 2019. The Life Cycle of Romance Demonstratives: Against a Low Merge Site. Talk. Research Seminars in Linguistics & Cognition series, Università degli Studi di Siena. https://pietrobaggio.wordpress.com/wp-content/uploads/2019/04/16_04_2019___siena_talk.pdf.

Baggio, Pietro. 2021. Introducing arguments and modifiers in the noun phrase. *Linguistic Inquiry* 52(2). 393–407. DOI: 10.1162/ling_a_00381.

Baggio, Pietro. 2022. *In Defence of DP - The Isomorphism between Nominals and Clauses*. LingBuzz: 006754. Pre-published.

Baier, Nicholas B. 2018. *Anti-Agreement*. Berkeley: University of California dissertation. https://escholarship.org/uc/item/2n162261.

Baker, James. 2019. Split intransitivity in English. *English Language and Linguistics* 23(3). 557–589. DOI: 10.1017/S1360674317000533.

Baker, Mark C. 2002. Building and merging, not checking: the nonexistence of (Aux)-SVO languages. *Linguistic Inquiry* 33(2). 321–328. DOI: 10.1162/002438902317406731.

Baker, Mark C. 2008a. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The Limits of Syntactic Variation*, 351–373. Amsterdam ; Philadelphia: John Benjamins. DOI: 10.1075/la.132.16bak.

Baker, Mark C. 2008b. *The syntax of agreement and concord*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511619830.

Baker, Mark C. 2012. Formal generative typology. In Bernd Heine & Heiko Narrog (eds.), *The oxford handbook of linguistic analysis*, 1st edn., 285–312. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199544004.013.0012.

Baker, Mark C. & Jonathan Bobalijk. 2017. On inherent and dependent theories of ergative case. In Jessica Coon, Diane Massam & Lisa deMena Travis (eds.), *The Oxford handbook of ergativity*, 111–134. Oxford Handbooks: online edn, Oxford Academic. DOI: 10.1093/oxfordhb/9780198739371.013.5.

Baltin, Mark R. 1989. Heads and Projections. In Mark R. Baltin & Anthony S. Kroch (eds.), *Alternative Conceptions of Phrase Structure*, 1–16. Chicago: University of Chicago Press. https://archive.org/details/alternativeconce0000unse.

Baltin, Mark R. 2006. Extraposition. In Martin Everaert & Henk C. van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, 1st edn., 237–271. Hoboken: John Wiley & Sons. DOI: 10.1002/9780470996591.ch25.

Bangalore, Srinivas & Aravind K. Joshi. 2010. Introduction. In Srinivas Bangalore & Aravind K. Joshi (eds.), *Supertagging: Using Complex Lexical Descriptions in Natural Language Processing*, 1–32. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/8370.003.0004.

Barbiers, Sjef. 2000. The Right Periphery in SOV Languages: English and Dutch. In Peter Svenonius (ed.), *The Derivation of VO and OV*, 181–218. Amsterdam: John Benjamins. https://www.jbe-platform.com/content/books/9789027299246-la.31.08bar.

Barrie, Michael Jonathan Mathew, Audrey Li, Martina Wiltschko & Jong Un Park. 2021. In Defense of DP (or KP). *Linguistic Research* 38(2). 207–238. https://www.academia.edu/download/79814084/2.pdf.

Barrie, Michael Jonathan Mathew & Éric Mathieu. 2014. Head movement. In Andrew Carnie, Yosuke Sato & Daniel Siddiqi (eds.), *The Routledge Handbook of Syntax*. Abingdon ; New York: Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315796604-9/head-movement-michael-barrie-%C3%A9ric-mathieu?context=ubx&refId=ef3e7cfa-eb3c-4bc5-8c12-921b2371201e.

Basri, Hasan. 1999. *Phonological and syntactic reflections of the morphological structure of Selayarese*. State University of New York at Stony Brook. https://search.proquest.com/openview/f558e674ea96d7d23cb1626467aa0362/1?pq-origsite=gscholar&cbl=18750&diss=y.

Bayer, Josef & Constantin Freitag. 2020. How much verb moves to second position? In Horst Lohnstein & Antonios Tsiknakis (eds.), *Verb Second*, 77–122. Berlin ; Boston: De Gruyter Mouton. DOI: 10.1515/9781501508141-003.

Beermann, Dorothee & Lars Hellan. 2011. Preposition stranding and locative adverbs in German. In Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz & Jan Koster (eds.), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, 31–42. Berlin ; Boston: De Gruyter Mouton. DOI: 10.1515/9783110892994.31.

Belk, Zoë & Ad Neeleman. 2017. AP adjacency as a precedence constraint. *Linguistic Inquiry* 48(1). 1–45. DOI: 10.1162/LING_a_00234.

Belletti, Adriana. 2004. Aspects of the Low IP Area. In Luigi Rizzi (ed.), *The Structure of CP and IP: The Cartography of Syntactic Structures*, vol. 2, 104–131. Oxford University Press. DOI: 10.1093/oso/9780195159486.003.0002.

Bernardi, Emanuele. 2022. Greenberg's Universal 20: New insights and perspectives. In Mihaela Tănase-Dogaru, Alina Tigău, Ioana Stoicescu & Mihaela Zamfirescu (eds.), *New insights into linguistic communicative behaviour*, 194–218. Newcastle-upon-Tyne: Cambridge Scholars Publishing. https://www.cambridgescholars.com/product/978-1-0364-1925-7/.

Bernardi, Emanuele. 2023. Greenberg's U(niversal) 20: A T(ree) A(djoining) G(rammar) perspective. In Mihaela Tănase-Dogaru, Alina Tigău & Ioana Stoicescu (eds.), *Crosslinguistic approaches to language analysis*, 189–221. Newcastle-upon-Tyne: Cambridge Scholars Publishing. https://www.cambridgescholars.com/product/978-1-0364-2066-6/.

Bernardi, Emanuele. 2025. A New Perspective on the Parallels between Greenberg's Universal 20 and the Word Order in the Clause. In Mihaela Tănase-Dogaru, Alina Tigău, Ioana Stoicescu & Ruxandra Dragan (eds.), *Novel Avenues to the Study of Language Structure*, 103–138. Newcastle-upon-Tyne: Cambridge Scholars Publishing. https://www.cambridgescholars.com/product/978-1-0364-5143-1/.

Bernstein, Judy B. 1997. Demonstratives and reinforcers in Romance and Germanic languages. *Lingua* 102(2). 87–113. DOI: 10.1016/S0024-3841(96)00046-0.

Bernstein, Judy B. 2001. The DP Hypothesis: Identifying Clausal Properties in the Nominal Domain. In Mark Baltin & Chris Collins (eds.), *The Handbook of Contemporary Syntactic Theory*, 1st edn., 536–561. Oxford ; Malden (MA): Blackwell. DOI: 10.1002/9780470756416.ch17.

Bernstein, Judy B. 2008. Reformulating the Determiner Phrase Analysis. *Language and Linguistics Compass* 2(6). 1246–1270. DOI: 10.1111/j.1749-818X.2008.00091.x.

Bernstein, Judy B., Francisco Ordóñez & Francesc Roca. 2021. Emphatic elements and the development of definite articles: Evidence for a layered DP in early Romance. *Journal of Historical Syntax* 5(16–25). 1–32. DOI: 10.18148/hs/2021.v5i16-25.63.

Bernstein, Judy B., Francisco Ordóñez & Francesc Roca. Forthcoming. The Role of Person in Ibero- and Gallo-Romance Demonstratives. In Amsterdam ; Philadelphia. John Benjamins.

Besnier, Niko. 2002. *Tuvaluan: A Polynesian Language of the Central Pacific*. Routledge. DOI: 10.4324/9780203027127.

den Besten, Hans. 2010. Is there "preposition stranding in COMP" in Afrikaans? No way! In C. Jan-Wouter Zwart & Mark de Vries (eds.), *Structure Preserved: Studies in syntax for Jan Koster*, 57–64. Amsterdam: John Benjamins. DOI: 10.1075/la.164.07bes.

Bhat, D. N. S. 2013. Third Person Pronouns and Demonstratives. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Zenodo. DOI: 10.5281/zenodo.13950591.

Biberauer, Theresa. 2017a. Probing The Nature Of The Final-Over-Final Condition: The Perspective From Adpositions. In Laura R. Bailey & Michelle L. Sheehan (eds.), *Order and structure in syntax I: Word order and syntactic structure*, 177–216. Berlin: Language Science Press. DOI: 10.5281/ZENODO.1117694.

Biberauer, Theresa. 2017b. The Final-over-Final Condition and Particles. In Michelle L. Sheehan, Theresa Biberauer, Ian Roberts & Anders Holmberg (eds.), *The Final-Over-Final Condition*, 187–296. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/8687.003.0012.

Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45(2). 169–225. DOI: 10.1162/LING_a_00153.

Biberauer, Theresa, Ian Roberts & Michelle Sheehan. 2014. No-choice parameters and the limits of syntactic variation. In Robert E. Santana-LaBarge (ed.), *Proceedings of the 31st West Coast Conference on Formal Linguistics*, 46–55. Somerville (MA): Cascadilla Press. https://www.lingref.com/cpp/wccfl/31/paper3006.pdf.

Biberauer, Theresa & Michelle L. Sheehan. 2013. Theoretical Approaches to Disharmonic Word Order. In Theresa Biberauer & Michelle L. Sheehan (eds.), *Theoretical Approaches to Disharmonic Word Order*, 1–44. Oxford ; New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780199684359.003.0001.

Bickel, Balthasar. 2007. Typology in the 21st century: Major current developments. *Linguistic Typology* 11. 239–251. DOI: 10.1515/LINGTY.2007.018.

Bickel, Balthasar. 2015. Distributional typology: Statistical inquiries into the dynamics of linguistic diversity. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 2nd edn., 901–924. Oxford: Oxford University Press. https://web.archive.org/web/20250315083122/https://www.zora.uzh.ch/id/eprint/109110/.

Bidese, Ermenegildo. 2023. *Sprachkontakt generativ: eine Untersuchung kontaktbedingten syntaktischen Wandels im Zimbrischen*. Berlin ; Boston: De Gruyter. https://library.oapen.org/handle/20.500.12657/63845.

Bittner, Maria & Ken Hale. 1996. The structural determination of case and agreement. *Linguistic inquiry* 27(1). 1–68. JSTOR: 4178925.

Black, Cheryl A. 1994. *Quiegolani zapotec syntax*. Santa Cruz: University of California dissertation. https://www.proquest.com/docview/304114470/abstract/9AA813C8D9B2471DPQ/1.

Bleam, Tonia. 2000. Clitic Climbing and the Power of Tree Adjoining Grammar. In Anne Abeillé & Owen Rambow (eds.), *Tree Adjoining Grammars: Formalisms, linguistic analyses and processing*, 193–220. Stanford (CA): CSLI Publications.

Blevins, Juliette. 2008. Hes and Kus Questions in Yurok: A case for lexeme-specific word order. *Skase Journal of Theoretical Linguistics* 5(1). https://julietteblevins.ws.gc.cuny.edu/files/2013/04/HesKus08.pdf.

Blümel, Andreas & Anke Holler. 2021. DP, NP, or neither? Contours of an unresolved debate. *Glossa: a journal of general linguistics* 7(1). 153. 1–25. DOI: 10.16995/glossa.8326.

Borer, Hagit. 1984. *Parametric syntax, case studies in Semitic and Romance languages*. Dordrecht (The Netherlands) ; Cinnaminson (NJ): Foris. http://archive.org/details/parametricsyntax0000bore.

Borer, Hagit. 1999. Deconstructing the Construct. In Kyle Johnson & Ian Roberts (eds.), *Beyond Principles and Parameters: Essays in Memory of Osvaldo Jaeggli*, 43–89. Dordrecht: Springer Netherlands. DOI: 10.1007/978-94-011-4822-1_3.

Borer, Hagit. 2005. *Structuring Sense Volume I: In Name Only*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199263905.001.0001.

Borer, Hagit. 2013. *Structuring Sense: Volume III: Taking Form*. 1st edn. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199263936.001.0001.

Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59(1). 1–45. DOI: 10.1111/j.1467-9582.2005.00118.x.

Bošković, Željko. 2016a. On the typology of the nominal domain: its syntax, semantics and learnability. In *Talk given at Workshop on Syntax and Semantics of the Nominal Domain. Goethe-Universität Frankfurt*. https://nominal-modification.de/wp-content/uploads/2015/12/ho_boskovic_feb16.pdf.

Bošković, Željko. 2016b. What is sent to spell-out is phases, not phasal complements. *Linguistica* 56(1). 25–66. DOI: 10.4312/linguistica.56.1.25-66.

Boullier, Pierre & Benoît Sagot. 2011. Multi-Component Tree Insertion Grammars. In Philippe de Groote, Markus Egg & Laura Kallmeyer (eds.), *Formal Grammar*, 31–46. Berlin ; Heidelberg: Springer. DOI: 10.1007/978-3-642-20169-1_3.

Brattico, Pauli & Cristiano Chesi. 2020. A top-down, parser-friendly approach to pied-piping and operator movement. *Lingua* 233. 102760. DOI: 10.1016/j.lingua.2019.102760.

Brody, Michael. 2018. Two Advantages of Precedence Syntax. In Huba Bartos, Marcel den Dikken, Zoltán Bánréti & Tamás Váradi (eds.), *Boundaries Crossed, at the Interfaces of Morphosyntax, Phonology, Pragmatics and Semantics*, 319–326. Cham: Springer International. DOI: 10.1007/978-3-319-90710-9_20.

Broekhuis, Hans. 2008. *Derivations and evaluations: Object shift in the Germanic languages*. Berlin ; New York: De Gruyter Mouton. DOI: 10.1515/9783110207200.

Broekhuis, Hans. 2010. A typology of clause structure. *Linguistic Variation Yearbook* 10. 1–31. DOI: 10.1075/livy.10.01bro.

Brown, Jason. 2024. Adjective Order and Adnominal Modification in Naasioi. *Linguistic Inquiry* 55(2). 423–435. DOI: 10.1162/ling_a_00464.

Bruening, Benjamin. 2009. Selectional Asymmetries between CP and DP Suggest that the DP Hypothesis is Wrong. (5). https://repository.upenn.edu/handle/20.500.14332/44728.

Bruening, Benjamin. 2014. Precede-and-Command Revisited. *Language* 90(2). 342–388. DOI: 10.1353/lan.2014.0037.

Bruening, Benjamin. 2018a. An Even Simpler Account of ”Dem Num A N” Orders. Rough Draft. University of Delaware. https://copland.udel.edu/~bruening/Downloads/DemNumANorder1.pdf.

Bruening, Benjamin. 2018b. CPs Move Rightward, Not Leftward. *Syntax* 21(4). 362–401. DOI: 10.1111/synt.12164.

Bruening, Benjamin. 2020. The head of the nominal is N, not D: N-to-D Movement, Hybrid Agreement, and conventionalized expressions. *Glossa-an International Journal of Linguistics* 5(1). 15. 1–19. DOI: 10.5334/gjgl.1031.

Bruening, Benjamin. 2021a. Generalizing the Presuppositional Approach to the Binding Conditions. *Syntax* 24(4). 417–461. DOI: 10.1111/synt.12221.

Bruening, Benjamin. 2021b. Spelling out the Numeration, Part 1: Selection by Itself Accounts for Synthetic-Periphrastic Alternations. Ms. University of Delaware. https://copland.udel.edu/~bruening/Downloads/NumerationPeriphrasis4.pdf.

Bruening, Benjamin. 2021c. Spelling out the Numeration, Part 2: A Left-to-Right Derivation Accounts for Elements that Attach to the First Item of a Certain Type. Ms. University of Delaware. https://copland.udel.edu/~bruening/Downloads/NumerationLocation1.pdf.

Bruening, Benjamin. 2022. Hebrew nominals do not require functional structure above the NP. *Glossa: a journal of general linguistics* 7(1). 1–45. DOI: 10.16995/glossa.5763.

Bruening, Benjamin & Eman Al Khalaf. 2019. No argument–adjunct asymmetry in reconstruction for Binding Condition C. *Journal of Linguistics* 55(2). 247–276. DOI: 10.1017/S0022226718000324.

Bruening, Benjamin, Xuyen Dinh & Lan Kim. 2018. Selection, idioms, and the structure of nominal phrases with and without classifiers. *Glossa: a journal of general linguistics* 3(1). 42. 1–46. DOI: 10.5334/gjgl.288.

Brugè, Laura. 1996. Demonstrative Movement in Spanish : A comparative approach. In Laura Brugè (ed.), *University of Venice Working Papers in Linguistics*, vol. 6, 1–54. Venice: Centro Linguistico Interfacoltà Università Ca' Foscari. https://phaidra.cab.unipd.it/api/object/o:458856/diss/Content/get.

Butt, Miriam. 2006. *Theories of Case*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139164696.

Canac Marquis, Réjean. 1989. Word orders in Gude and the VSO Parameter. In Zygmunt Frajzyngier (ed.), *Current Progress in Chadic Linguistics: Proceedings of the International Symposium on Chadic Linguistics, Boulder, Colorado, 1–2 May 1987*, 55–86. Amsterdam: John Benjamins. DOI: 10.1075/cilt.62.04can.

Candito, Marie-Hélène & Sylvain Kahane. 1998. Can the TAG derivation tree represent a semantic graph? An answer in the light of Meaning-Text Theory. In Anne Abeillé, Tilman

Becker, Giorgio Satta & K. Vijay-Shanker (eds.), *Proceedings of the Fourth International Workshop on Tree Adjoining Grammars and Related Frameworks (TAG+4)*, 21–24. University of Pennsylvania: Institute for Research in Cognitive Science. https://aclanthology.org/W98-0106/.

Cardinaletti, Anna & Giuliana Giusti. 2006. The Syntax of Quantified Phrases and Quantitative Clitics. In Martin Everaert & Henk Van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, 1st edn., 23–93. Hoboken: Wiley. DOI: 10.1002/9780470996591.ch71.

Castillo, Juan Carlos, John E. Drury & Kleanthes Grohmann. 2009–1999. Merge Over Move and the Extended Projection Principle: MOM and the EPP Revisited. *Iberia: An International Journal of Theoretical Linguistics* 1(1). 53–114. https://scholarworks.uni.edu/facpub/6495.

Catasso, Nicholas. 2024. A Cartographic Approach to Verb Movement and Two Types of FinP V2 in German. *Languages* 9(1). 21. DOI: 10.3390/languages9010021.

Cecchetto, Carlo. 2013. Backward dependencies must be short. In Theresa Biberauer & Ian Roberts (eds.), *Challenges to Linearization*, 57–92. Berlin: De Gruyter Mouton. DOI: 10.1515/9781614512431.57.

Chen-Main, Joan. 2012. Delayed Tree Locality and the Status of Derivation Structure. In *Proceedings of the 11th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+ 11)*, 99–107. https://aclanthology.org/W12-4612.pdf.

Chen-Main, Joan, Tonia Bleam & Aravind K. Joshi. 2012. Delayed Tree-Locality, Set-locality, and Clitic Climbing. In Giorgio Satta & Chung-Hye Han (eds.), *Proceedings of the 11th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+11)*, 1–9. Paris, France. https://aclanthology.org/W12-4601.

Chen-Main, Joan & Aravind K. Joshi. 2007. Multi-Component Tree Adjoining Grammars, Dependency Graph Models, and Linguistic Analyses. In Timothy Baldwin, Mark Dras, Julia Hockenmaier, Tracy Holloway King & Gertjan van Noord (eds.), *ACL 2007 Workshop on Deep Linguistic Processing*, 1–8. Prague, Czech Republic: Association for Computational Linguistics. https://aclanthology.org/W07-1201/.

Chen-Main, Joan & Aravind K. Joshi. 2008. Flexible Composition, Multiple Adjoining and Word Order Variation. In Claire Gardent & Anoop Sarkar (eds.), *Proceedings of the Ninth International Workshop on Tree Adjoining Grammar and Related Frameworks (TAG+9)*,

9–16. Tübingen: Association for Computational Linguistics. https://aclanthology.org/W08-2302.

Chen-Main, Joan & Aravind K. Joshi. 2014. A dependency perspective on the adequacy of tree local multi-component tree adjoining grammar. *Journal of Logic and Computation* 24(5). 989–1022. DOI: 10.1093/logcom/exs012.

Chesi, Cristiano. 2015. On Directionality of Phrase Structure Building. *Journal of Psycholinguistic Research* 44(1). 65–89. DOI: 10.1007/s10936-014-9330-6.

Chesi, Cristiano. 2024. *Is It the End of (Generative) Linguistics as We Know It?* DOI: 10.48550/arXiv.2412.12797. Pre-published.

Chesi, Cristiano, Dan Nieto & Roberto Zamparelli. 2022. Constraints on Greenberg's Universal 20: Agreement and expectations. https://agenda.uib.es/_files/_event/_74834/_editorFiles/file/POSTERS/ChesiNieto_Zamparelli.pdf.

Chiang, David & Tatjana Scheffler. 2008. Flexible Composition and Delayed Tree-Locality. In Claire Gardent & Anoop Sarkar (eds.), *Proceedings of the Ninth International Workshop on Tree Adjoining Grammar and Related Frameworks (TAG+9)*, 17–24. Tübingen, Germany: Association for Computational Linguistics. https://aclanthology.org/W08-2303.

Chierchia, Gennaro. 1998. Reference to kinds across language. *Natural language semantics* 6(4). 339–405. DOI: 10.1023/A:1008324218506.

Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174. 99–149. DOI: 10.1007/s11229-009-9686-6.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge (MA): MIT Press. JSTOR: j.ctt17kk81z.

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick Jacobs & Peter Rosenbaum (eds.), *Readings in English Transformational Grammar*, 184–221. Waltham (MA): Ginn. https://archive.org/details/readingsinenglis0000jaco.

Chomsky, Noam. 1979. *The Logical Structure of Linguistic Theory*. Chicago: The University Of Chicago Press. http://archive.org/details/logicalstructure00noam.

Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris. https://archive.org/details/lecturesongovern0000chom.

Chomsky, Noam. 1986. *Barriers*. Cambridge (MA): MIT Press. https://archive.org/details/barriers0000chom.

Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory. In Kenneth Hale & Samuel J. Keyser (eds.), *The view from Building 20*, 1–52. Cambridge (MA): MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge (MA): MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: essays on minimalist syntax in honor of howard lasnik*, 89–155. Cambridge (MA): MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In Michael J. Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/4056.003.0004.

Chomsky, Noam. 2002–1957. *Syntactic Structures*. 2nd edn. Berlin ; New York: De Gruyter Mouton. https://www.ling.upenn.edu/courses/ling5700/Chomsky1957.pdf.

Chomsky, Noam. 2004. *The Generative Enterprise Revisited: Discussions with Riny Huybregts, Henk van Riemsdijk, Naoki Fukui and Mihoko Zushi*. Berlin ; New York: De Gruyter Mouton. DOI: 10.1515/9783110902440.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic inquiry* 36(1). 1–22. JSTOR: 4179307.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces+ recursion= language? Chomsky's minimalism and the view from syntax-semantics*, 1–29. London ; New York: De Gruyter Mouton. DOI: 10.1515/9783110207552.1.

Chomsky, Noam. 2008. On Phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory*, 133–166. The MIT Press. DOI: 10.7551/mitpress/7713.003.0009.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI: 10.1016/j.lingua.2012.12.003.

Chomsky, Noam. 2014–1995. *The Minimalist Program*. 20th anniversary edition. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/9780262527347.001.0001.

Chomsky, Noam. 2015. Problems of projection: Extensions. In Elisa Di Domenico, Cornelia Hamann & Simona Matteini (eds.), *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, 3–16. Amsterdam ; Philadelphia: John Benjamins. DOI: 10.1075/la.223.01cho.

Chomsky, Noam. 2020. The UCLA Lectures (April 29–May 2, 2019). LingBuzz: 005485.

Chomsky, Noam & George A. Miller. 1963. Introduction to the formal analysis of natural languages. In Duncan R. Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of Mathematical Psychology*, vol. 2, 269–321. New York: Wiley. https://archive.org/details/handbookofmathem017893mbp/page/269/.

Chomsky, Noam & Marcel Paul Schützenberger. 1963. The Algebraic Theory of Context-Free Languages. In Paul Louis Braffort & Daniel S. Hirschberg (eds.), *Studies in Logic and the Foundations of Mathematics*, 118–161. Elsevier. DOI: 10.1016/S0049-237X(08)72023-8.

Chomsky, Noam, T. Daniel Seely, Robert C. Berwick, Sandiway Fong, M. A. C. Huybregts, Hisatsugu Kitahara, Andrew McInnerney & Yushi Sugimoto. 2023. *Merge and the Strong Minimalist Thesis*. 1st edn. Cambridge University Press. DOI: 10.1017/9781009343244.

Chung, Chin-Wan. 2003. The Linear Constituent Order of the Noun Phrase: An Optimality Theoretic Account. *English Language & Literature Teaching* 9(1). 23–48. https://koreascience.kr/article/JAKO200317357304278.page.

Chung, Sandra. 2006. Properties of VOS Languages. In Martin Everaert & Henk C. van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, 1st edn., 685–720. Hoboken: John Wiley & Sons. DOI: 10.1002/9780470996591.ch52.

Cinque, Guglielmo. 1990a. Ergative adjectives and the lexicalist hypothesis. *Natural Language & Linguistic Theory* 8(1). 1–39. DOI: 10.1007/BF00205530.

Cinque, Guglielmo. 1990b. *Types of a'-dependencies*. Cambridge (MA). https://iris.unive.it/handle/10278/7437?mode=complete.

Cinque, Guglielmo. 1994. On the evidence for partial N-movement in the Romance DP. In Jan Koster, Raffaella Zanuttini, Guglielmo Cinque & Jean-Yves Pollock (eds.), *Paths towards Universal Grammar: Studies in honor of Richard S. Kayne*, 85–110. Washington (DC): Georgetown University Press. https://hdl.handle.net/10278/10825.

Cinque, Guglielmo. 1996. The 'antisymmetric'programme: theoretical and typological implications1. *Journal of Linguistics* 32(2). 447–464. DOI: 10.1017/S0022226700015966.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780195115260.001.0001.

Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and Its Exceptions. *Linguistic Inquiry* 36(3). 315–332. DOI: 10.1162/0024389054396917.

Cinque, Guglielmo. 2009. The Fundamental Left-Right Asymmetry of Natural Languages. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of Language Today*, 165–184. Dordrecht: Springer Netherlands. DOI: 10.1007/978-1-4020-8825-4_9.

Cinque, Guglielmo. 2010. *The syntax of adjectives: a comparative study*. Cambridge (MA): MIT Press.

Cinque, Guglielmo. 2013. Word-Order Typology: A Change of Perspective. In Theresa Biberauer & Michelle Sheehan (eds.), *Theoretical Approaches to Disharmonic Word Order*, 47–73. Oxford ; New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780199684359.003.0002.

Cinque, Guglielmo. 2014. Extraction from DP in Italian Revisited. In Enoch Oladé Aboh, Maria Teresa Guasti & Ian Roberts (eds.), *Locality*, 1st edn., 86–103. Oxford ; New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780199945269.003.0004.

Cinque, Guglielmo. 2017a. A Microparametric Approach To The Head-Initial/Head-Final Parameter. 41(3–4). 309–366. https://www.linguisticanalysis.com/downloads/a-microparametric-approach-to-the-head-initial-head-final-parameter-page-309-366-download/.

Cinque, Guglielmo. 2017b. On the status of functional categories (heads and phrases). *Language and Linguistics* 18(4). 521–576. DOI: 10.1075/lali.18.4.01cin.

Cinque, Guglielmo. 2021. *A Database on the Cross-Linguistic Order of Demonstrative, Numeral, Adjective and Noun*. ARCA. https://hdl.handle.net/10278/3750226.

Cinque, Guglielmo. 2023. *On linearization. Towards a restrictive theory*. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/14681.001.0001.

Cinque, Guglielmo. Forthcoming. Again on the Order of Demonstrative, Numeral, Adjective, and Noun. Ms. University of Venice.

Citko, Barbara. 2014. *Phase Theory: An Introduction*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139644037.

Clem, Emily. 2022. Disharmony and the Final-Over-Final Condition in Amahuaca. *Linguistic Inquiry* 53(4). 809–822. DOI: 10.1162/ling_a_00425.

Clemens, Lauren E. & Maria Polinsky. 2017. Verb-initial word orders (primarily in Austronesian and Mayan languages). In Martin Everaert & Henk C. van Riemsdijk (eds.), *The*

Wiley Blackwell companion to syntax, 2nd edn., 1–50. Hoboken: John Wiley & Sons. DOI: 10.1002/9781118358733.wbsyncom056.

Comrie, Bernard. 2013. Numeral Bases (v2020.3). In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Zenodo. DOI: 10.5281/zenodo.7385533.

Corver, Norbert. 2013. Lexical categories and (extended) projection. In Marcel den Dikken (ed.), *The Cambridge Handbook of Generative Syntax*, 353–424. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511804571.015.

Crisma, Paola. 1991. Functional categories inside the noun phrase: A study on the distribution of nominal modifiers. Ms. University of Venice.

Croft, William. 2003. *Typology and Universals*. 2nd edn. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511840579.

Culbertson, Jennifer. 2017. New approaches to greenbergian word order dependencies. In N. J. Enfield (ed.), *Dependencies in language: On the causal ontology of linguistic systems*, 23–38. Berlin: Language Science Press. https://zenodo.org/records/573778.

Culbertson, Jennifer & David Adger. 2014. Language learners privilege structured meaning over surface frequency. *Proceedings of the National Academy of Sciences* 111(16). 5842–5847. DOI: 10.1073/pnas.1320525111.

Culbertson, Jennifer, Julie Franck, Guillaume Braquet, Magda Barrera Navarro & Inbal Arnon. 2020. A learning bias for word order harmony: Evidence from speakers of non-harmonic languages. *Cognition* 204. 104392. DOI: 10.1016/j.cognition.2020.104392.

Cysouw, Michael. 2010. Dealing with diversity: Towards an explanation of NP-internal word order frequencies. *Linguistic Typology* 14(2–3). DOI: 10.1515/lity.2010.010.

Danon, Gabi. 2012. Two structures for numeral-noun constructions. *Lingua* 122(12). 1282–1307. DOI: 10.1016/j.lingua.2012.07.003.

Davies, William D. 1999. Madurese and Javanese as Strict Word-Order Languages. *Oceanic Linguistics* 38(1). 152–167. DOI: 10.2307/3623397.

Davies, William D. 2005. Madurese Prolepsis and Its Implications for a Typology of Raising. *Language* 81(3). 645–665. DOI: 10.1353/lan.2005.0121.

Davis, Colin Pierce Bryon. 2020. *The linear limitations of syntactic derivations*. Cambridge (MA): Massachusetts Institute of Technology dissertation. https://dspace.mit.edu/handle/1721.1/129122.

De Santo, Aniello, Alëna Aksënova & Thomas Graf. 2016. An Alternate View on Strong Lexicalization in TAG. In David Chiang & Alexander Koller (eds.), *Proceedings of the 12th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+12)*, 93–102. Düsseldorf, Germany. https://aclanthology.org/W16-3310/.

Deal, Amy Rose. 2015. Ergativity. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax - Theory and Analysis: An International Handbook*, vol. 1, 654–708. Berlin ; München ; Boston: De Gruyter Mouton. DOI: 10.1515/9783110377408.654.

Deal, Amy Rose. 2016. Syntactic Ergativity: Analysis and Identification. *Annual Review of Linguistics* 2(1). 165–185. DOI: 10.1146/annurev-linguistics-011415-040642.

Déchaîne, Rose-Marie & Victor Manfredi. 1998. SVO ergativity and abstract ergativity. *Recherches linguistiques de Vincennes* (27). 71–94. https://people.bu.edu/manfredi/RLV.pdf.

Dékány, Éva. 2018. Approaches to head movement: A critical assessment. *Glossa: a journal of general linguistics* 3(1). 65. 1–43. DOI: 10.5334/gjgl.316.

Dékány, Éva. 2021. *The Hungarian Nominal Functional Sequence*. Cham: Springer International Publishing. DOI: 10.1007/978-3-030-63441-4.

DelBusso, Natalie R. 2018. *Typological structure and properties of property theory*. University of Rutgers dissertation. DOI: doi:10.7282/T3T43X8T.

DelBusso, Natalie R. 2020. The Final-over-Final Condition, Stringency, and Typological Structure. *Linguistic Inquiry* 51(4). 765–784. DOI: 10.1162/ling_a_00357.

Delfitto, Denis. 2006. Adverb Classes and Adverb Placement. In Martin Everaert & Henk C. van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, 1st edn., 83–120. Hoboken: John Wiley & Sons. DOI: 10.1002/9780470996591.ch4.

Diercks, Michael, Beatrice Santorini & Anthony S. Kroch. 2025. *The Syntax of Natural Language*. 2nd edn. Claremont: Scholarship@Claremont. https://scholarship.claremont.edu/pomona_facbooks/50/.

Diessel, Holger. 1999. *Demonstratives: form, function and grammaticalization*. Amsterdam: John Benjamins. DOI: 10.1075/tsl.42.

den Dikken, Marcel. 2009. Arguments for successive-cyclic movement through SpecCP. *Linguistic Variation Yearbook* 9(1). 89–126. DOI: 10.1075/livy.9.03dik.

den Dikken, Marcel. 2018. *Dependency and directionality*. Cambridge: Cambridge University Press. DOI: 10.1017/9781316822821.

Dimitrova-Vulchanova, Mila & Giuliana Giusti. 1998. Fragments of Balkan nominal structure. In Artemis Alexiadou & Chris Wilder (eds.), *Possessors, predicates and movement in the determiner phrase*, 333–360. Amsterdam: John Benjamins. DOI: 10.1075/la.22.12dim.

Diouf, Jean-Léopold. 2009. *Grammaire du wolof contemporain: Édition revue et complétée*. Paris: L'Harmattan. https://www.torrossa.com/gs/resourceProxy?an=5095795&publisher=FZ2990.

Dixon, Robert Malcolm Ward. 1988. *A grammar of Boumaa Fijian*. Chicago: University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/G/bo3643579.html.

Doetjes, Jenny. 2021a. Number and Numeral Classifiers. In Patricia Cabredo Hofherr & Jenny Doetjes (eds.), *The Oxford Handbook of Grammatical Number*, 220–241. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780198795858.013.10.

Doetjes, Jenny. 2021b. Number and Quantity Expressions. In Patricia Cabredo Hofherr & Jenny Doetjes (eds.), *The Oxford Handbook of Grammatical Number*, 65–100. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780198795858.013.4.

Donohue, Mark. 2011. *A grammar of Tukang Besi*. Berlin ; New York: De Gruyter Mouton. DOI: 10.1515/9783110805543.

Douglas, Jamie. 2017. Unifying the that-trace and anti-that-trace effects. *Glossa: a journal of general linguistics* 2(1). 60. 1–28. DOI: 10.5334/gjgl.312.

Doyle, Aidan. 2001. *Irish*. Vol. 201. München: Lincom Europa.

Dryer, Matthew S. 1991. SVO languages and the OV:VO typology. *Journal of Linguistics* 27(2). 443–482. JSTOR: 4176126.

Dryer, Matthew S. 1992. The Greenbergian Word Order Correlations. *Language* 68(1). 81–138. DOI: 10.2307/416370.

Dryer, Matthew S. 1997. On the Six-Way Word Order Typology. *Studies in Language* 21(1). 69–103. DOI: 10.1075/sl.21.1.04dry.

Dryer, Matthew S. 2007. Word order. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, 2nd edn., 61–131. Cambridge University Press. DOI: 10.1017/CBO9780511619427.002.

Dryer, Matthew S. 2009. The branching direction theory of word order correlations revisited. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of language today*, 185–207. Dordrecht: Springer. DOI: 10.1007/978-1-4020-8825-4_10.

Dryer, Matthew S. 2013a. On the six-way word order typology, again. *Studies in Language. International Journal Sponsored by the Foundation "Foundations of Language"* 37(2). 267–301. DOI: 10.1075/sl.37.2.02dry.

Dryer, Matthew S. 2013b. Order of Adjective and Noun (v2020.3). In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Zenodo. DOI: 10.5281/zenodo.7385533.

Dryer, Matthew S. 2013c. Order of Demonstrative and Noun (v2020.3). In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Zenodo. DOI: 10.5281/zenodo.7385533.

Dryer, Matthew S. 2013d. Order of Numeral and Noun (v2020.3). In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Zenodo. DOI: 10.5281/zenodo.7385533.

Dryer, Matthew S. 2013e. Order of Relative Clause and Noun (v2020.3). In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Zenodo. DOI: 10.5281/zenodo.7385533.

Dryer, Matthew S. 2013f. Order of Subject, Object and Verb (v2020.3). In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Zenodo. DOI: 10.5281/zenodo.7385533.

Dryer, Matthew S. 2018a. On the order of demonstrative, numeral, adjective, and noun. *Language* 94(4). 798–833. DOI: 10.1353/lan.2018.0054.

Dryer, Matthew S. 2018b. On the order of demonstrative, numeral, adjective, and noun: Supplemental Material. *Language* 94(4). DOI: 10.1353/lan.2018.0072.

Dryer, Matthew S. 2019. Grammaticalization accounts of word order correlations. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 63–95. Berlin: Language Science Press. DOI: 10.5281/zenodo.2583809.

Dryer, Matthew S. & Martin Haspelmath. 2013. *The World Atlas of Language Structures Online*. Version v2020.3. DOI: 10.5281/ZENODO.7385533.

Duguine, Maia. 2022. FOFC as a PF phenomenon: Evidence from Basque clausal embedding. *Glossa: a journal of general linguistics* 7(1). 1–46. DOI: 10.16995/glossa.5745.

Dyer, William, Charles Torres, Gregory Scontras & Richard Futrell. 2023. Evaluating a Century of Progress on the Cognitive Science of Adjective Ordering. *Transactions of the Association for Computational Linguistics* 11. 1185–1200. DOI: 10.1162/tacl_a_00596.

Eberhard, David M., Gary F. Simons & Charles D. Fennig. 2024. *Ethnologue: Languages of the World*. Twenty-Sixth. Version 27th. Dallas, Texas: SIL International. http://www.ethnologue.com.

Edmiston, Daniel; Potsdam & Eric Potsdam. 2017. Linearization at PF: Evidence from Malagasy Extraposition. DOI: 10.6082/07RS-BH88.

Elordieta, Arantzazu. 2013. On the Relevance of the Head Parameter in a Mixed OV Language. In Theresa Biberauer & Michelle Sheehan (eds.), *Theoretical Approaches to Disharmonic Word Order*, 306–329. Oxford University Press. DOI: 10.1093/acprof:oso/9780199684359.003.0011.

Elordieta, Arantzazu & Bill Haddican. 2018. Truncation feeds intervention. *Natural Language & Linguistic Theory* 36(2). 403–443. DOI: 10.1007/s11049-017-9381-0.

Emonds, Joseph. 2012. Which Is the Extended Projection Above N? In Laura Brugé, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro & Cecilia Poletto (eds.), *Functional Heads*, 325–337. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199746736.003.0025.

Engelfriet, Joost & Andreas Maletti. 2017. Multiple Context-Free Tree Grammars and Multi-component Tree Adjoining Grammars. In Ralf Klasing & Marc Zeitoun (eds.), *Fundamentals of Computation Theory*, 217–229. Berlin; Heidelberg: Springer. DOI: 10.1007/978-3-662-55751-8_18.

Epps, Patience, Claire Bowern, Cynthia A. Hansen, Jane H. Hill & Jason Zentz. 2012. On numeral complexity in hunter-gatherer languages. 16(1). 41–109. DOI: 10.1515/lity-2012-0002.

Erlewine, Michael Yoshitaka. 2017. Low sentence-final particles in Mandarin Chinese and the Final-over-Final Constraint. *Journal of East Asian Linguistics* 26(1). 37–75. DOI: 10.1007/s10831-016-9150-9.

Erlewine, Michael Yoshitaka. 2020. Anti-locality and subject extraction. *Glossa: a journal of general linguistics* 5(1). 84. 1–38. DOI: 10.5334/gjgl.1079.

Ernst, Thomas. 2014. The syntax of adverbs. In Andrew Carnie, Yosuke Sato & Daniel Siddiqi (eds.), *The Routledge handbook of syntax*, 108–130. Abingdon ; New York: Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315796604-7/syntax-adverbs-thomas-ernst?context=ubx&refId=0206210d-a3bb-4dda-a3d5-f86094e2b6e7.

Etxepare, Ricardo & Bill Haddican. 2017. Repairing Final-Over-Final Condition Violations: Evidence From Basque Verb Clusters. In Laura R. Bailey & Michelle L. Sheehan (eds.), *Order and structure in syntax I: Word order and syntactic structure*, 135–157. Berlin: Language Science Press. DOI: 10.5281/ZENODO.1117702.

Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32(5). 429–448. DOI: 10.1017/S0140525X0999094X.

Eynde, Frank Van. 2020. Agreement, disagreement and the NP vs. DP debate. *Glossa: a journal of general linguistics* 5(1). 65. 1–23. DOI: 10.5334/gjgl.1119.

Eynde, Frank Van. 2021. Nominal structures. In Stefan Müller, Anne Abeillé & Robert D. Borsley (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 275–313. Berlin: Language Science Press. DOI: 10.5281/ZENODO.5599832.

Feldcamp, Zachary. 2022. Linkers and the licensing of dependents. *Toronto Working Papers in Linguistics*. https://twpl.library.utoronto.ca/index.php/twpl/article/download/41005/31378.

Ferrer-i-Cancho, Ramon. 2024. *The Optimal Placement of the Head in the Noun Phrase. The Case of Demonstrative, Numeral, Adjective and Noun*. DOI: 10.48550/arXiv.2402.10311. Pre-published.

Findlay, Jamie Y. 2023. LFG and Tree-Adjoining Grammar. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*, 2069–2125. Berlin: Language Science Press. DOI: 10.5281/ZENODO.10186054.

Finer, Daniel L. 1997. Contrasting Ā-Dependencies in Selayarese. *Natural Language and Linguistic Theory* 15(4). 677–728. DOI: 10.1023/A:1005801115342.

Forkel, Robert & Harald Hammarström. 2022. Glottocodes: Identifiers linking families, languages and dialects to comprehensive reference information. *Semantic Web* 13(6). 917–924. DOI: 10.3233/SW-212843.

Franco, Ludovico, Leonardo M. Savoia & Maria Rita Manzini. 2015. Linkers and agreement. *The Linguistic Review* 32(2). 277–332. DOI: 10.1515/tlr-2014-0024.

Frank, Robert. 2002. *Phrase structure composition and syntactic dependencies*. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/5366.001.0001.

Frank, Robert. 2004. Restricting grammatical complexity. *Cognitive Science* 28(5). 669–697. DOI: 10.1207/s15516709cog2805_3.

Frank, Robert. 2006. Phase theory and Tree Adjoining Grammar. *Lingua. International review of general linguistics. Revue internationale de linguistique générale* 116. 145–202. DOI: 10.1016/j.lingua.2005.02.005.

Frank, Robert. 2010. Lexicalized Syntax and Phonological Merge. In Srinivas Bangalore & Aravind K. Joshi (eds.), *Supertagging: Using Complex Lexical Descriptions in Natural Language Processing*, 373–406. Cambridge (MA): The MIT Press. DOI: 10.7551/mitpress/8370.003.0023.

Frank, Robert. 2013. Tree Adjoining Grammar. In Marcel den Dikken (ed.), *The Cambridge Handbook of Generative Syntax*, 226–262. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511804571.011.

Frank, Robert & Tim Hunter. 2021. Variation in mild context-sensitivity: Derivational state and structural monotonicity. *Evolutionary Linguistic Theory* 3(2). 181–214. DOI: 10.1075/elt.00033.fra.

Frank, Robert & Dennis Ryan Storoshenko. 2012. The Shape of Elementary Trees and Scope Possibilities in STAG. In Giorgio Satta & Chung-Hye Han (eds.), *Proceedings of the 11th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+11)*, 232–240. Paris, France. https://aclanthology.org/W12-4627.

Fukui, Naoki & Hiroki Narita. 2014. Merge, labeling, and projection. In Andrew Carnie, Yosuke Sato & Daniel Siddiqi (eds.), *The Routledge Handbook of Syntax*, 1st edn., 3–23. Abingdon ; New York: Routledge. DOI: 10.4324/9781315796604.

Fukui, Naoki & Yuji Takano. 1998. Symmetry in syntax: Merge and demerge. *Journal of East Asian Linguistics* 7(1). 27–86. DOI: 10.1023/A:1008240710949.

Fukui, Naoki & Yuji Takano. 2000. Nominal Structure: An Extension of the Symmetry Principle. In Peter Svenonius (ed.), *The Derivation of VO and OV*, 219–254. Amsterdam: John Benjamins. https://www.jbe-platform.com/content/books/9789027299246-la.31.09fuk.

Gardent, Claire & Shashi Narayan. 2015. Multiple Adjunction in Feature-Based Tree-Adjoining Grammar. *Computational Linguistics* 41(1). 41–70. DOI: 10.1162/COLI_a_00217.

Gärtner, Hans-Martin. 2022. Copies from "Standard Set Theory"? A Note on the Foundations of Minimalist Syntax in Reaction to Chomsky, Gallego and Ott (2019). *Journal of Logic, Language and Information* 31(1). 129–135. DOI: 10.1007/s10849-021-09342-x.

Gast, Volker & Łukasz Jędrzejowski. 2022. Managing Data in a Typological Study. In Andrea L. Berez-Kroeker, Bradley McDonnell, Eve Koller & Lauren B. Collister (eds.), *The Open Handbook of Linguistic Data Management*, 597–608. The MIT Press. DOI: 10.7551/mitpress/12200.003.0058.

van Gelderen, Elly. 2011. *The linguistic cycle: Language change and the language faculty*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199756056.001.0001.

van Gelderen, Elly. 2017. *Syntax: An Introduction to Minimalism*. Amsterdam: John Benjamins. DOI: 10.1075/z.214.

Gell-Mann, Murray & Merritt Ruhlen. 2011. The origin and evolution of word order. *Proceedings of the National Academy of Sciences* 108(42). 17290–17295. DOI: 10.1073/pnas.1113716108.

Georgi, Doreen. 2019. Intermediate reflexes of movement: A problem for TAG? In J. M. M. Brown, Andreas Schmidt & Marta Wierzba (eds.), *Of trees and birds: a Festschrift for Gisbert Fanselow*, 77–94. Potsdam: University Press Potsdam. DOI: 10.25932/publishup-43221.

Georgi, Doreen & Gereon Müller. 2010. Noun-Phrase Structure by Reprojection. *Syntax* 13(1). 1–36. DOI: 10.1111/j.1467-9612.2009.00132.x.

Gerdes, Kim, Sylvain Kahane & Xinying Chen. 2021. Typometrics: From Implicational to Quantitative Universals in Word Order Typology. *Glossa: a journal of general linguistics* 6(1). 17. 1–31. DOI: 10.5334/gjgl.764.

Gianollo, Chiara, Cristina Guardiano & Giuseppe Longobardi. 2008. Three fundamental issues in parametric linguistics. In Theresa Biberauer (ed.), *The Limits of Syntactic Variation*, 109–142. Amsterdam ; Philadelphia: John Benjamins. DOI: 10.1075/la.132.05gia.

Gianollo, Chiara, Klaus von Heusinger & Maria Napoli. 2021. Reference and Quantification in Nominal Phrases: The Current Landscape and the Way Ahead. In Chiara Gianollo, Klaus von Heusinger & Maria Napoli (eds.), *Determiners and Quantifiers: Functions, Variation, and Change*, 1–28. Leiden (The Netherlands): Brill. DOI: 10.1163/9789004473324_002.

Gil, David. 2001. 92. Quantifiers. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, vol. 2, 2 vols., 1275–1294. Berlin ; New York: De Gruyter Mouton. DOI: 10.1515/9783110194265-029.

Giorgi, Alessandra & Fabio Pianesi. 1997. *Tense And Aspect: From Semantics to Morphosyntax*. New York: Oxford University Press. DOI: 10.1093/oso/9780195091922.001.0001.

Giusti, Giuliana. 2002. The Functional Structure of Noun Phrases: A Bare Phrase Structure Approach. In Guglielmo Cinque (ed.), *Functional Structure in DP and IP: The Cartography of Syntactic Structures, Volume 1*, 54–90. Oxford University Press. DOI: 10.1093/oso/9780195148794.003.0003.

Giusti, Giuliana. 2018. Demonstratives as arguments and modifiers of N. In Marco Coniglio, Andrew Murphy, Eva Schlachter & Tonjes Veenstra (eds.), *Atypical Demonstratives*, 23–56. Berlin ; Boston: De Gruyter Mouton. DOI: 10.1515/9783110560299-002.

Giusti, Giuliana. in press. Arguments for the universality of dp and determiners. In Martina Wiltschko & Solveiga Armoskaite (eds.), *The oxford handbook of determiners*. https://hdl.handle.net/10278/3739477.

Glaude, Herby. 2013. *Aspects de la syntaxe de l'haïtien*. Paris: Anibwé. https://pure.uva.nl/ws/files/1989164/116959_thesis.pdf.

Gorn, Saul. 1967. Explicit Definitions and Linguistic Dominoes. In John Hart & Satoru Takasu (eds.), *Systems and Computer Science*, 77–115. University of Toronto Press. DOI: 10.3138/9781487592769-008.

Graffi, Giorgio. 1980. "Universali di Greenberg" e grammatica generativa. *Lingua e Stile* 15(3). 371–387. https://iris.univr.it/handle/11562/428803.

Graffi, Giorgio. 2001. *200 years of syntax: A critical survey*. Amsterdam: John Benjamins. DOI: 10.1075/sihols.98.

Greenberg, Joseph Harold. 1966–1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In Joseph Harold Greenberg (ed.), *Universals of Language: Report of a Conference Held at Dobbs Ferry, New York, April 13-15, 1961. 2nd ed*, 2nd edn., 73–113. Cambridge (MA): MIT Press. https://www.fulcrum.org/epubs/1n79h7060?locale=en#page=15.

Greenberg, Joseph Harold. 1990–1972. Numeral Classifiers and Substantival Number: Problems in the Genesis of a Linguistic Type. In Keith Denning & Suzanne Kemmer (eds.), *On Language: Selected Writings of Joseph H. Greenberg*, 166–193. Redwood City: Stanford University Press. DOI: 10.1515/9781503623217-009.

Greenberg, Joseph Harold. 1990. Two approaches to language universals. In Keith Denning & Suzanne Kemmer (eds.), *On Language: Selected Writings of Joseph H. Greenberg*, 702–722. Redwood City: Stanford University Press. DOI: 10.1515/9781503623217-039.

Grimshaw, Jane B. 1990. *Argument structure*. Cambridge (MA): MIT Press. http://archive.org/details/argumentstructur0000grim.

Grimshaw, Jane B. 1991. Extended projections. Ms. Brandeis University.

Grimshaw, Jane B. 2005. Extended projection. In *Words and structure*, 1–73. Stanford: CSLI Publications. https://csli.sites.stanford.edu/publications/csli-lecture-notes/words-and-structure.

Grohmann, Kleanthes K. 2003. *Prolific Domains: On the Anti-Locality of movement dependencies*. Amsterdam: John Benjamins. DOI: 10.1075/la.66.

Grohmann, Kleanthes K. 2012. Anti-Locality: Too-Close Relations in Grammar. In Cedric Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199549368.013.0012.

Guardiano, Cristina. 2012. Demonstratives, word order and the DP between syntax and semantics: crosslinguistic remarks. In Despina Papadopoulou, Anthi Revithiadou & Alex Tantos (eds.), *Studies in Greek Linguistics 32*, 100–115. Aristotle University of Thessaloniki: Institùto neoellinikòn spudòn. https://hdl.handle.net/11380/743718.

Guimarães, Maximiliano. 2000. In defense of vacuous projections in bare phrase structure. *University of Maryland Working Papers in Linguistics* 9. 90–115. LingBuzz: 000463/current.pdf?_s=2KOzTceiMvhm2QH7.

Guimarães, Maximiliano. 2008. A note on the strong generative capacity of standard Antisymmetry. *Snippets* (18). https://www.ledonline.it/snippets/allegati/snippets18001.pdf.

Guzmán Naranjo, Matías & Laura Becker. 2022. Statistical bias control in typology. *Linguistic Typology* 26(3). 605–670. DOI: 10.1515/lingty-2021-0002.

Haegeman, Liliane & Henk C. van Riemsdijk. 1986. Verb projection raising, scope, and the typology of rules affecting verbs. *Linguistic inquiry* 17(3). 417–466. JSTOR: 4178499.

Haegeman, Liliane & Barbara Ürögdi. 2010. Referential CPs and DPs: An operator movement account. *Theoretical Linguistics* 36(2–3). DOI: 10.1515/thli.2010.008.

Haider, Hubert. 1992. Branching and Discharge. In *Working Papers of the SFB 340*. Stuttgart. https://www.plus.ac.at/wp-content/uploads/2021/02/BRANCHING__Discharge.pdf.

Haider, Hubert. 1997. Precedence among predicates. *Journal of Comparative Germanic Linguistics* 1(1). 3–41. DOI: 10.1023/A:1009713319586.

Haider, Hubert. 2000. OV is More Basic than VO. In Peter Svenonius (ed.), *The derivation of OV and VO*, 45–68. Amsterdam: John Benjamins. DOI: 10.1075/la.31.03hai.

Haider, Hubert. 2010. *The Syntax of German*. 1st edn. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511845314.

Haider, Hubert. 2013. *Symmetry Breaking in Syntax*. 1st edn. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139084635.

Haider, Hubert. 2015. Head directionality – in syntax and morphology. In Antonio Fabregas, Jaume Mateu & Michael T. Putnam (eds.), *Contemporary Linguistic Parameters*, 73–97. London: Bloomsbury Academic. https://uni-salzburg.elsevierpure.com/en/publications/head-directionality-in-syntax-and-morphology.

Haider, Hubert. 2018. *Phrase-Structure Based Typology*. DOI: 10.13140/RG.2.2.34341.76008. Pre-published.

Haider, Hubert. 2019. On absent, expletive, and non-referential subjects. In *Semantic and syntactic aspects of impersonality*, 11–46. Hamburg: Helmut Buske. https://uni-salzburg.elsevierpure.com/en/publications/on-absent-expletive-and-non-referential-subjects.

Haider, Hubert. 2020a. **Final-over-Final – *Initial-over-Final – The Universal Asymmetry of Functional Projections*. DOI: 10.13140/RG.2.2.23705.85600. Pre-published.

Haider, Hubert. 2020b. VO-/OV-Base Ordering. In B. Richard Page & Michael T. Putnam (eds.), *The Cambridge Handbook of Germanic Linguistics*, 339–364. Cambridge: Cambridge University Press. DOI: 10.1017/9781108378291.016.

Haider, Hubert. 2022a. Subjects across alignment systems. Ms. LingBuzz: 006840.

Haider, Hubert. 2022b. The Left-Left Constraint: A structural constraint on adjuncts. In Ulrike Freywald, Stefan Müller & Horst J. Simon (eds.), *Headedness and/or grammatical anarchy?*, vol. 11. Berlin: Language Science Press. DOI: 10.5281/ZENODO.7142712.

Haider, Hubert. 2023. «OVS» – a misnomer for SVO languages with ergative alignment. Submitted paper. LingBuzz: 005680.

Haider, Hubert & Luka Szucsich. 2022a. Slavic languages are Type 3 languages: replies. *Theoretical Linguistics* 48(1–2). 113–142. DOI: 10.1515/tl-2022-2036.

Haider, Hubert & Luka Szucsich. 2022b. Slavic Languages–«SVO» languages without SVO qualities? *Theoretical Linguistics* 48(1–2). 1–39. DOI: doi.org/10.1515/tl-2022-2035.

Hall, David. 2015. *Spelling Out the Noun Phrase: Interpretation, Word Order, and the Problems of 'Meaningless Movement'*. London: Queen Mary University dissertation. https://www.qmul.ac.uk/sllf/media/sllf-new/department-of-linguistics/documents/35)-David-Hall_PhDThesis.pdf.

Hammarstrom, Harald. 2014. *SAILS Online*. Jena: Max Planck Institute for the Science of Human History. https://sails.clld.org/.

Hammarström, Harald. 2007. Complexity in numeral systems with an investigation into pidgins and creoles. In Fred Karlsson, Matti Miestamo & Kaius Sinnemäki (eds.), *Language Complexity: Typology, Contact, Change*, 287–304. Amsterdam ; Philadelphia: John Benjamins. https://urn.kb.se/resolve?urn=urn%3Anbn%3Ase%3Auu%3Adiva-336573.

Hammarström, Harald. 2015. The basic word order typology: An exhaustive study. http://www.eva.mpg.de/fileadmin/content_files/linguistics/conferences/2015-diversity-linguistics/Hammarstroem_slides.pdf.

Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2024. *Glottolog 5.0*. Leipzig: Max Planck Institute for Evolutionary Anthropology. DOI: 10.5281/zenodo.10804357.

Han, Chung-hye, David Potter & Dennis Ryan Storoshenko. 2010. Non-local Right Node Raising: an Analysis using Delayed Tree-Local MC-TAG. In Srinivas Bangalore, Robert Frank & Maribel Romero (eds.), *Proceedings of the 10th International Workshop on Tree Adjoining Grammar and Related Frameworks (TAG+10)*, 9–16. Yale University: Linguistic Department, Yale University. https://aclanthology.org/W10-4402.

Hanafiou, H. S. 2007. *Grammaire comparée français-songay-zarma*. Niamey (Niger): Université Abdou Moumouni de Niamey. https://www.songhay.org/static/documents/grammar/grammaire_comparee.pdf.

Harizanov, Boris. 2019. Head movement to specifier positions. *Glossa: a journal of general linguistics* 4(1). 140. 1–36. DOI: 10.5334/gjgl.871.

Harizanov, Boris & Vera Gribanova. 2019. Whither head movement? *Natural Language & Linguistic Theory* 37(2). 461–522. DOI: 10.1007/s11049-018-9420-5.

Harriehausen, Bettina. 1990. *Hmong njua: syntaktische analyse einer gesprochenen sprache mithilfe datenverarbeitungstechnischer mittel und sprachvergleichende beschreibung des südostasiatischen sprachraumes*. Berlin: Max Niemeyer Verlag. DOI: 10.1515/9783111356440.

Hawkins, John A. 1983. *Word order universals*. New York: Academic Press. DOI: 10.1016/C2009-0-21896-4.

Hawkins, John A. 2013. Disharmonic word orders from a processing-efficiency perspective. In Theresa Biberauer & Michelle Sheehan (eds.), *Theoretical approaches to disharmonic word orders*, 391–406. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199684359.003.0014.

Hayward, Dick. 1984. *The Arbore Language: A First Investigation (including a vocabulary)*. Hamburg: Helmut Buske Verlag.

Heck, Fabian. 2016. *Non-monotonic derivations*. Universität Leipzig Habilitation. https://home.uni-leipzig.de/heck/papiere/heck-2016-nonmon.pdf.

Heck, Fabian. 2023. Towards a Unified Explanation of Apparent Cases of Counter-cyclicity. In Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.), *Cyclicity*, 297–328. Leipzig: Institut für Linguistik Universität Leipzig. https://home.uni-leipzig.de/~heck/papiere/heck-cyc.pdf.

Heck, Fabian & Gereon Müller. 2007. Extremely local optimization. In *Proceedings of WECOL 2006*, vol. 26, 170–183. Fresno: California State University. https://home.uni-leipzig.de/~heck/papiere/wecolpapier.pdf.

Hengeveld, Kees, Jan Rijkhoff & Anna Siewierska. 2004. Parts-of-speech systems and word order. *Journal of Linguistics* 40(3). 527–570. DOI: 10.1017/S0022226704002762.

Her, One-Soon. 2017. Deriving classifier word order typology, or Greenberg's Universal 20A and Universal 20. *Linguistics* 55(2). 265–303. DOI: 10.1515/ling-2016-0044.

Hicks, Glyn. 2005. Deriving Condition A. *York Papers in Linguistics* 3(series 2). 91–131. https://eprints.soton.ac.uk/64786/.

Higginbotham, James. 1985. On Semantics. *Linguistic Inquiry* 16(4). 547–593. JSTOR: 4178457.

Höhn, Georg FK. 2017. *Non-possessive person in the nominal domain*. University of Cambridge dissertation. https://www.researchgate.net/profile/Georg-Hoehn/publication/380462272_Non-possessive_person_in_the_nominal_domain/links/663d775c35243041538539bc/Non-possessive-person-in-the-nominal-domain.pdf.

Holmberg, Anders. 2000. Deriving OV Order in Finnish. In Peter Svenonius (ed.), *The Derivation of VO and OV*, 123–152. Amsterdam: John Benjamins. DOI: 10.1075/la.31.06hol.

Holmberg, Anders. 2015. Verb Second. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and Analysis: An International Handbook*, vol. 1, 342–383. Berlin ; München ; Boston: De Gruyter Mouton. DOI: 10.1515/9783110377408.342.

Holmberg, Anders. 2017a. Introduction. In Michelle Sheehan, Theresa Biberauer, Ian Roberts & Andres Holmberg (eds.), *The Final-Over-Final Condition: A Syntactic Universal*, 1–10. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/8687.003.0004.

Holmberg, Anders. 2017b. The Final-over-Final Condition and Linearization in Generative Grammar. In Michelle L. Sheehan, Theresa Biberauer, Ian Roberts & Andres Holmberg (eds.), *The Final-Over-Final Condition: A Syntactic Universal*, 43–78. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/8687.003.0007.

Hoot, Bradley & Shane Ebert. 2021. The That-Trace Effect: Evidence from Spanish–English Code-Switching. *Languages* 6(4). 189. DOI: 10.3390/languages6040189.

Horvath, Julia. 2006. Pied-Piping. In Martin Everaert & Henk C. van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, 1st edn., 569–630. Hoboken: John Wiley & Sons. DOI: 10.1002/9780470996591.ch50.

Hoskinson, James. 1983. *A grammar and dictionary of the Gude language*. The Ohio State University dissertation. https://web.archive.org/web/20181117192920/https://etd.ohiolink.edu/!etd.send_file?accession=osu1487239392028787&disposition=inline.

Hsu, Brian. 2017. Verb second and its deviations: An argument for feature scattering in the left periphery. *Glossa: a journal of general linguistics* 2(1). 35. 1–33. DOI: 10.5334/gjgl.132.

Hsu, Brian. 2021. Coalescence: A Unification of Bundling Operations in Syntax. *Linguistic Inquiry* 52(1). 39–87. DOI: 10.1162/ling_a_00372.

Hsu, Brian & Saurov Syed. 2020. Cooccurrence of nonarticle determiners as evidence for split DPs and bundled projections. *Language* 96(3). e114–e134. DOI: 10.1353/lan.2020.0051.

Hualde, José Ignacio & Jon Ortiz de Urbina (eds.). 2003. *A grammar of Basque*. Berlin ; New York: De Gruyter Mouton. DOI: 10.1515/9783110895285.

Huber, Magnus & the APiCS Consortium. 2013. Order of subject, object, and verb. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press. https://apics-online.info/parameters/1.

Hucklebridge, Shay. 2023. Learning and the typology of word order: A model of the Final-over-Final Condition. *Glossa: a journal of general linguistics* 8(1). 1–50. DOI: 10.16995/glossa.6566.

Ionin, Tania & Ora Matushansky. 2018. *Cardinals: The syntax and semantics of cardinal-containing expressions*. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/8703.001.0001.

Iordăchioaia, Gianina. 2020. D and N are different nominalizers. *Glossa: a journal of general linguistics* 5(1). 53. 1–25. DOI: 10.5334/gjgl.1111.

Ivani, Jessica K. & Balthasar Bickel. 2023. *Databases for Comparative Syntactic Research*. DOI: 10.48550/arXiv.2310.11187. Pre-published.

Jackendoff, Ray. 1977. *X'-syntax: A study of phrase structure*. Cambridge (MA): MIT Press. http://archive.org/details/xsyntaxstudyofph00jack.

Jäger, Gerhard & James Rogers. 2012. Formal language theory: refining the Chomsky hierarchy. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1598). 1956–1970. DOI: 10.1098/rstb.2012.0077.

Jenks, Peter. 2015. Two kinds of definites in numeral classifier languages. *Semantics and Linguistic Theory* 25. 103. DOI: 10.3765/salt.v25i0.3057.

Jenks, Peter. 2018. Articulated Definiteness without Articles. *Linguistic Inquiry* 49(3). 501–536. DOI: 10.1162/ling_a_00280.

Jenks, Peter & Rassidatou Konate. 2022. Indexed definiteness. *Glossa: a journal of general linguistics* 7(1). 1–44. DOI: 10.16995/glossa.5717.

Joo, Kum-Jeong. 2015. Factors for Preferred Internal Orderings of Korean Noun Phrases. *Second Language Research* 51. 383–400. https://s-space.snu.ac.kr/bitstream/10371/94839/1/05.%ec%a3%bc%ea%b8%88%ec%a0%95(383-400).pdf.

Joshi, Aravind K. 1985. Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In David R. Dowty, Lauri Karttunen & Arnold M. Zwicky (eds.), *Natural Language Parsing*, 1st edn., 206–250. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511597855.007.

Joshi, Aravind K. 1987. An introduction to tree adjoining grammars. *Mathematics of language* 1. 87–115. DOI: 10.1075/z.35.07jos.

Joshi, Aravind K. 2003. A note on the strong and weak generative powers of formal systems. *Theoretical Computer Science* 293(2). 243–259. DOI: 10.1016/S0304-3975(01)00347-4.

Joshi, Aravind K. 2014. Recursion in Language: Is It Indirectly Constrained? In Tom Roeper & Margaret Speas (eds.), *Recursion: Complexity in Cognition*, 139–158. Cham: Springer International Publishing. DOI: 10.1007/978-3-319-05086-7_7.

Joshi, Aravind K., Tilman Becker & Owen Rambow. 2000. Complexity of scrambling: A new twist to the competence-performance distinction. In Anne Abeillé & Owen Rambow (eds.), *Tree Adjoining Grammars: Formalisms, linguistic analyses and processing*, 167–181. Stanford (CA): CSLI Publications.

Joshi, Aravind K., Laura Kallmeyer & Maribel Romero. 2003. Flexible composition in LTAG: Quantifier scope and inverse linking. In Harry Bunt & Reinhard Muskens (eds.), *Comput-*

ing Meaning, vol. 3, 233–256. Dordrecht: Springer Netherlands. https://core.ac.uk/outputs/14509792/.

Joshi, Aravind K., Leon S. Levy & Masako Takahashi. 1975. Tree adjunct grammars. *Journal of Computer and System Sciences* 10(1). 136–163. DOI: 10.1016/S0022-0000(75)80019-5.

Joshi, Aravind K. & Owen Rambow. 2003. A Formalism for Dependency Grammar Based on Tree Adjoining Grammar. In *Proceedings of the Conference on Meaning-text Theory*, 207–216. Paris: MTT. https://web.archive.org/web/20240414025337/http://www.cs.columbia.edu/~rambow/papers/joshi-rambow-2003.pdf.

Joshi, Aravind K. & Yves Schabes. 1991. Tree-adjoining grammars and lexicalized grammars. *Technical Reports (CIS)* 445. https://repository.upenn.edu/handle/20.500.14332/7375.

Joshi, Aravind K. & Yves Schabes. 1997. Tree-Adjoining Grammars. In Grzegorz Rozenberg & Arto Salomaa (eds.), *Handbook of Formal Languages: Volume 3 Beyond Words*, 69–123. Berlin ; Heidelberg: Springer. DOI: 10.1007/978-3-642-59126-6_2.

Julien, Marit. 2002. *Syntactic heads and word formation*. Oxford ; New York: Oxford University Press. DOI: 10.1093/oso/9780195149500.001.0001.

Kallmeyer, Laura. 2009. A Declarative Characterization of Different Types of Multicomponent Tree Adjoining Grammars. *Research on Language and Computation* 7(1). 55–99. DOI: 10.1007/s11168-009-9064-z.

Kallmeyer, Laura. 2010. *Parsing beyond Context-Free Grammars*. Heidelberg ; New York: Springer. http://archive.org/details/parsingbeyondcon0000kall.

Kallmeyer, Laura & Aravind Joshi. 2003. Factoring Predicate Argument and Scope Semantics: Underspecified Semantics with LTAG. *Research on Language and Computation* 1(1). 3–58. DOI: 10.1023/A:1024564228892.

Kallmeyer, Laura & Marco Kuhlmann. 2012. A Formal Model for Plausible Dependencies in Lexicalized Tree Adjoining Grammar. In Giorgio Satta & Chung-Hye Han (eds.), *Proceedings of the 11th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+11)*, 108–116. Paris, France. https://aclanthology.org/W12-4613.

Kallmeyer, Laura, Rainer Osswald & Robert D. Van Valin. 2013. Tree Wrapping for Role and Reference Grammar. In Glyn Morrill & Mark-Jan Nederhof (eds.), *Formal Grammar*, 175–190. Berlin ; Heidelberg: Springer. DOI: 10.1007/978-3-642-39998-5_11.

Kallmeyer, Laura & Tatjana Scheffler. 2004. LTAG Analysis for Pied-Piping and Stranding of wh-Phrases. In Owen Rambow & Matthew Stone (eds.), *Proceedings of the 7th International Workshop on Tree Adjoining Grammar and Related Formalisms*, 32–39. Vancouver, Canada: Simon Fraser University. https://aclanthology.org/W04-3305/.

Kantarovich, Jessica. 2020. *Argument structure in language shift: Morphosyntactic variation and grammatical resilience in Modern Chukchi*. The University of Chicago dissertation. https://search.proquest.com/openview/5168bf9ae2071caefb05546a0abf5420/1?pq-origsite=gscholar&cbl=18750&diss=y.

Kashyap, Abhishek Kumar. 2019. Language Typology. In David Schönthal, Geoff Thompson, Lise Fontaine & Wendy L. Bowcher (eds.), *The Cambridge Handbook of Systemic Functional Linguistics*, 767–792. Cambridge: Cambridge University Press. DOI: 10.1017/9781316337936.031.

Kayne, Richard S. 1984. *Connectedness and binary branching*. Berlin ; Boston: De Gruyter Mouton. DOI: 10.1515/9783111682228.

Kayne, Richard S. 1989. Null Subjects and Clitic Climbing. In Osvaldo A. Jaeggli & Kenneth J. Safir (eds.). Red. by Joan Maling & Luigi Rizzi, *The Null Subject Parameter*, vol. 15, 239–261. Dordrecht: Springer Netherlands. DOI: 10.1007/978-94-009-2540-3_8.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge (MA): MIT Press. https://archive.org/details/antisymmetryofsy0000kayn.

Kayne, Richard S. 2008. Some Notes on Comparative Syntax, with Special Reference to English and French. In Guglielmo Cinque & Richard S. Kayne (eds.), *The Oxford Handbook of Comparative Syntax*, 3–69. Oxford ; New York: Oxford University Press. DOI: 10.1093/oxfordhb/9780195136517.013.0001.

Kayne, Richard S. 2013. Why Are There No Directionality Parameters? In Theresa Biberauer & Michelle Sheehan (eds.), *Theoretical Approaches to Disharmonic Word Order*, 219–244. Oxford ; New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780199684359.003.0008.

Keine, Stefan & Hedde Zeijlstra. 2023. Clause-internal successive cyclicity: phasality or DP intervention? https://stefankeine.com/papers/KeineZeijlstra-2024.pdf.

Keizer, Evelien. 2022. Modifier-numeral word order in the English NP: An FDG analysis. *Open Linguistics* 8(1). 545–572. DOI: 10.1515/opli-2022-0207.

Kim, Kyumin & Paul B. Melchin. 2018. On the complementary distribution of plurals and classifiers in East Asian classifier languages. *Language and Linguistics Compass* 12(4). e12271. DOI: 10.1111/lnc3.12271.

Kim, Min-Joo. 2019. *The Syntax and Semantics of Noun Modifiers and the Theory of Universal Grammar: A Korean Perspective*. Cham: Springer International Publishing. DOI: 10.1007/978-3-030-05886-9.

Kiss, Katalin É. 2003. Argument Scrambling, Operator Movement, and Topic Movement in Hungarian. In *Word Order and Scrambling*, 22–43. Malden (MA): Blackwell. DOI: 10.1002/9780470758403.ch2.

Klockmann, Heidi. 2020. The article a(n) in English quantifying expressions: A default marker of cardinality. *Glossa: a journal of general linguistics* 5(1). 85. 1–31. DOI: 10.5334/gjgl.1151.

Kobele, Gregory M. 2023. Cyclicity in Minimalist Syntax. In Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.), *Cyclicity*, 73–82. Leipzig: Institut für Linguistik Universität Leipzig. https://www.philol.uni-leipzig.de/fileadmin/Fakult%C3%A4t_Philo/Linguistik/Forschung/LAB/LAB_95/lab95_2_Kobele.pdf.

Koeneman, Olaf & Hedde Zeijlstra. 2014. The Rich Agreement Hypothesis Rehabilitated. *Linguistic Inquiry* 45(4). 571–615. DOI: 10.1162/LING_a_00167.

Koopman, Hilda. 2000a. The Spec Head Configuration. In *The Syntax of Specifiers and Heads*, 323–356. London ; New York: Routledge. DOI: 10.4324/9780203027868.

Koopman, Hilda. 2000b. Unifying Predicate Cleft Constructions. In *The Syntax of Specifiers and Heads*, 357–374. London ; New York: Routledge. DOI: 10.4324/9780203027868.

Koopman, Hilda & Cristina Guardiano. 2022. Managing Data in TerraLing, a Large-Scale Cross-Linguistic Database of Morphological, Syntactic, and Semantic Patterns. In Andrea L. Berez-Kroeker, Bradley McDonnell, Eve Koller & Lauren B. Collister (eds.), *The Open Handbook of Linguistic Data Management*, 617–630. The MIT Press. DOI: 10.7551/mitpress/12200.003.0060.

Koopman, Hilda & Dominique Sportiche. 1991. The position of subjects. *Lingua* 85(2–3). 211–258. DOI: 10.1016/0024-3841(91)90022-W.

Körtvélyessy, Lívia. 2017. *Essentials of language typology*. Košice: UPJŠ. https://unibook.upjs.sk/img/cms/2017/ff/esentials-language-final.pdf.

Kosta, Peter & Diego Gabriel Krivochen. 2012. Some thoughts on language diversity, ug and the importance of language typology: scrambling and non-monotonic merge of adjuncts and specifiers in czech and german. *Zeitschrift für Slawistik* 57(4). 377–407. DOI: 10.1524/slaw.2012.0031.

Kotowski, Sven. 2016. *Adjectival Modification and Order Restrictions: The Influence of Temporariness on Prenominal Word Order*. Berlin ; Boston: De Gruyter. DOI: 10.1515/9783110478457.

Köylü, Yılmaz. 2021. An overview of the NP versus DP debate. *Language and Linguistics Compass* 15(3). e12406. DOI: 10.1111/lnc3.12406.

Krivochen, Diego Gabriel. 2018a. *Aspects of emergent cyclicity in language and computation*. University of Reading dissertation. https://centaur.reading.ac.uk/77331/.

Krivochen, Diego Gabriel. 2018b. *On Folding and Twisting (and Whatknot): Towards a Characterization of Workspaces in Syntax*. DOI: http://arxiv.org/abs/1809.07853. Pre-published.

Krivochen, Diego Gabriel. 2019. On trans-derivational operations: generative semantics and tree adjoining grammar. *Language Sciences* 74. 47–76. DOI: 10.1016/j.langsci.2019.04.002.

Krivochen, Diego Gabriel. 2021. Mixed computation. Grammar up and down the Chomsky hierarchy. *Evolutionary Linguistic Theory* 3(2). 215–244. DOI: 10.1075/elt.00034.kri.

Krivochen, Diego Gabriel. 2022. Sobre la syntaxis de las oraciones de relativo: Un análisis de adjunción con estructura compartida. *Borealis – An International Journal of Hispanic Linguistics* 11(3). 305–377. DOI: 10.7557/1.11.3.6661.

Krivochen, Diego Gabriel. 2023a. Raising to object: A graph-theoretic analysis. *Evolutionary Linguistic Theory* 5(2). 128–161. DOI: 10.1075/elt.00050.kri.

Krivochen, Diego Gabriel. 2023b. *Syntax on the Edge: A Graph-Theoretic Analysis of Sentence Structure*. Leiden (The Netherlands): Brill. DOI: 10.1163/9789004542310.

Krivochen, Diego Gabriel. 2023c. The Search for Minimal Search: A Graph-Theoretic Approach. *Biolinguistics* 17. 1–50. DOI: 10.5964/bioling.9793.

Krivochen, Diego Gabriel. 2023d. Towards a theory of syntactic workspaces: Neighbourhoods and distances in a lexicalised grammar. *The Linguistic Review* 40(2). 311–360. DOI: 10.1515/tlr-2023-2004.

Krivochen, Diego Gabriel. 2024. Constituents, arrays, and trees: two (more) models of grammatical description. *Folia Linguistica* 58(3). 699–727. DOI: 10.1515/flin-2024-2025.

Krivochen, Diego Gabriel & Andrea Padovan. 2021. Lexicalised locality: Local domains and non-local dependencies in a lexicalised tree-adjoining grammar. *Philosophies* 6(3). DOI: 10.3390/philosophies6030070.

Kroch, Anthony S. 1987. Unbounded Dependencies and Subjacency in a Tree Adjoining Grammar. In Alexis Manaster-Ramer (ed.), *Mathematics of Language: Proceedings of a conference held at the University of Michigan, Ann Arbor, October 1984*, 143–172. Amsterdam: John Benjamins. DOI: 10.1075/z.35.09kro.

Kroch, Anthony S. 1989. Asymmetries in Long-Distance Extraction in a Tree-Adjoining Grammar. In Mark R. Baltin & Anthony S. Kroch (eds.), *Alternative conceptions of phrase structure*, 66–98. Chicago: University Press. https://archive.org/details/alternativeconce0000unse/page/66/.

Kroch, Anthony S. & Aravind K. Joshi. 1985. *The Linguistic Relevance of Tree Adjoining Grammar*. Technical Reports (CIS) MS-CIS-85-16. Philadelphia (PA): University of Pennsylvania. https://repository.upenn.edu/handle/20.500.14332/7618.

Kroch, Anthony S. & Beatrice Santorini. 1991. The derived constituent structure of the West-Germanic verb-raising construction. In Robert Freidin (ed.), *Principles and Parameters in Comparative Grammar*, 269–338. Cambridge (MA): MIT Press. https://www.ling.upenn.edu/~kroch/papers/kroch-and-santorini-1991.pdf.

Kuhlmann, Marco & Giorgio Satta. 2012. Tree-Adjoining Grammars Are Not Closed Under Strong Lexicalization. *Computational Linguistics* 38(3). 617–629. DOI: 10.1162/COLI_a_00090.

Kulick, Seth, Robert Frank & K. Vijay-Shanker. 2000. Defective Complements in Tree Adjoining Grammar. (10). https://repository.upenn.edu/handle/20.500.14332/45506.

Kulick, Seth Norman. 2000. *Constraining non-local dependencies in tree-adjoining grammar: computational and linguistic perspectives*. USA: University of Pennsylvania dissertation. https://www.proquest.com/dissertations-theses/constraining-non-local-dependencies-tree/docview/304640613/se-2.

Laenzlinger, Christopher. 2005. French adjective ordering: perspectives on DP-internal movement types. *Lingua* 115(5). 645–689. DOI: 10.1016/j.lingua.2003.11.003.

Laenzlinger, Christopher. 2017. A View of the CP/DP-(non)parallelism from the Cartographic Perspective. *Languages* 2(4). 18. DOI: 10.3390/languages2040018.

Laka, Itziar. 1990. *Negation in syntax–on the nature of functional categories and projections*. Cambridge (MA): Massachusetts Institute of Technology dissertation. https://dspace.mit.edu/bitstream/handle/1721.1/13667/24302516-MIT.pdf?sequence=2.

Lamontagne, Greg & Lisa deMena Travis. 1987. The syntax of adjacency. In *Proceedings of the Sixth West Coast Conference on Formal Linguistics*, 173–186. Stanford: CSLI Publications. https://www.researchgate.net/publication/248030092_The_syntax_of_adjacency.

Landau, Idan. 2016. DP-internal semantic agreement: A configurational analysis. *Natural Language & Linguistic Theory* 34(3). 975–1020. DOI: 10.1007/s11049-015-9319-3.

Larson, Richard K. 1988. On the double object construction. *Linguistic inquiry* 19(3). 335–391. JSTOR: 25164901.

Larson, Richard K. 2014. *On Shell Structure*. New York: Routledge. DOI: 10.4324/9780203429204.

Larson, Richard K. 2019. The DP Hypothesis and (A)Symmetries Between DP and CP. *Linguistic Analysis* 42(3–4). 507–548. https://www.linguisticanalysis.com/downloads/the-dp-hypothesis-and-asymmetries-between-dp-and-cp-page-507-548-download/.

Lasnik, Howard & Terje Lohndal. 2013. Brief overview of the history of generative syntax. In Marcel den Dikken (ed.), *The Cambridge Handbook of Generative Syntax*, 26–60. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511804571.004.

Ledgeway, Adam. 2012a. From Latin to Romance: Configurationality, Functional Categories and Head-Marking. *Transactions of the Philological Society* 110(3). 422–442. DOI: 10.1111/j.1467-968X.2012.01310.x.

Ledgeway, Adam. 2012b. *From Latin to Romance: Morphosyntactic Typology and Change*. Oxford ; New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780199584376.001.0001.

Lehmann, Winfred P. 1973. A Structural Principle of Language and Its Implications. *Language* 49(1). 47. DOI: 10.2307/412102.

Lehmann, Winfred P. 1978. The Great Undelying Ground-Plans. In *Syntactic typology: studies in the phenomenology of language*, 3–55. Austin: University of Texas press. https://lrc.la.utexas.edu/books/typology/index.

Leivada, Evelina. 2020. Mid-level generalizations of generative linguistics: Experimental robustness, cognitive underpinnings and the interdisciplinarity paradox. *Zeitschrift für Sprachwissenschaft* 39(3). 357–374. DOI: 10.1515/zfs-2020-2021.

Leivada, Evelina. 2022. Determining the cognitive biases behind a viral linguistic universal: the order of multiple adjectives. *Humanities and Social Sciences Communications* 9(1). 436. DOI: 10.1057/s41599-022-01440-w.

Leivada, Evelina & Marit Westergaard. 2019. Universal linguistic hierarchies are not innately wired. Evidence from multiple adjectives. *PeerJ* 7. e7438. DOI: 10.7717/peerj.7438.

Leu, Thomas. 2014. *The Architecture of Determiners*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199945238.001.0001.

Levin, Theodore Frank. 2015. *Licensing without case*. Massachusetts Institute of Technology dissertation. https://dspace.mit.edu/handle/1721.1/101451.

Levinson, Stephen C. 2022. *A Grammar of Yélî Dnye: The Papuan Language of Rossel Island*. Berlin ; Boston: De Gruyter Mouton. DOI: 10.1515/9783110733853.

Lichte, Timm. 2015. *Syntax und valenz: zur modellierung kohärenter und elliptischer strukturen mit baumadjunktionsgrammatiken*. Berlin: Language Science Press. DOI: 10.26530/OAPEN_603338.

Lichte, Timm. 2021. Against strict headedness in syntax. *Journal of Language Modelling* 9(2). 291–348. https://jlm.ipipan.waw.pl/index.php/JLM/article/view/249.

Lichte, Timm & Laura Kallmeyer. 2016. Same Syntax, Different Semantics: A Compositional Approach to Idiomaticity in Multi-word Expressions. In Christopher Piñón (ed.), *Empirical Issues in Syntax and Semantics 11*, 111–140. Paris: CSSP Publications. http://www.cssp.cnrs.fr/eiss11/eiss11_lichte-and-kallmeyer.pdf.

Lichte, Timm & Laura Kallmeyer. 2017. *Tree-Adjoining Grammar: A Tree-Based Constructionist Grammar Framework for Natural Language Understanding*. SS-17-02. https://cdn.aaai.org/ocs/15330/15330-68218-1-PB.pdf.

Lindenbergh, Charlotte & Jan-Wouter Zwart. 2021. Rethinking alignment typology. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences III: Inside syntax*, 23–50. Berlin: Language Science Press. https://zenodo.org/record/4680298.

Little, Carol-Rose, Mary Moroney & Justin Royer. 2022. Classifiers can be for numerals or nouns: Two strategies for numeral modification. *Glossa: a journal of general linguistics* 7(1). 1–35. DOI: 10.16995/glossa.8437.

Longenbaugh, Nicholas Steven. 2014. *Tree Adjoining Grammar at the Interfaces*. Harvard College BA thesis. http://nrs.harvard.edu/urn-3:HUL.InstRepos:12717344.

Longobardi, Giuseppe. 1994. Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic inquiry*. 609–665. JSTOR: 4178880.

Longobardi, Giuseppe, Paola Crisma & Cristina Guardiano. 2020. Syntactic diversity and language learnability. *Studi e Saggi Linguistici* 58(2). 99–130. https://www.studiesaggilinguistici.it/ssl/article/download/265/226.

Lynch, John. 2000. *A grammar of Anejom*. DOI: 10.15144/PL-507.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511605789.

Mahajan, Anoop. 1997. Universal grammar and the typology of ergative languages. In Artemis Alexiadou & Tracy Alan Hall (eds.), *Studies on Universal Grammar and Typological Variation*, 35–57. DOI: 10.1075/la.13.03mah.

Mahowald, Kyle, Dan Jurafsky & Mark Norris. 2021. Concord begets concord: A Bayesian model of nominal concord typology. *Proceedings of the Linguistic Society of America* 6(1). 541–555. DOI: 10.3765/plsa.v6i1.4988.

Malčić, Justin. 2019. The asymmetry and antisymmetry of syntax: A relational approach to displacement. In *Proceedings of ULAB IX*, 260–290. London: Undergraduate Linguistics Association of Britain. DOI: 10.5281/zenodo.3773332.

Maldonado, Violeta Vázquez Rojas. 2012. *The syntax and semantics of Purépecha noun phrases and the mass/count distinction*. New York University dissertation. https://

search.proquest.com/openview/312d9ce6e549bbce245e5ed7ef718905/1?pq-origsite=gscholar&cbl=18750.

Maletti, Andreas & Joost Engelfriet. 2012. Strong Lexicalization of Tree Adjoining Grammars. In Haizhou Li, Chin-Yew Lin, Miles Osborne, Gary Geunbae Lee & Jong C. Park (eds.), *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 506–515. Jeju Island, Korea: Association for Computational Linguistics. https://aclanthology.org/P12-1053/.

Mallinson, Graham & Barry J. Blake. 1981. *Language typology: Cross-linguistic studies in syntax*. Amsterdam: North-Holland. https://archive.org/details/languagetypology0000mall.

Manning, Christopher David. 2003. Probabilistic Syntax. In Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probabilistic Linguistics*. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/5582.003.0011.

Manova, Stela. 2021. The linear order of elements in prominent linguistic sequences: Deriving Tns-Asp-Mood orders and Greenberg's Universal 20 with n-grams. Ms. University of Vienna. DOI: 10.13140/RG.2.2.17715.55843.

Manzini, Maria Rita. 2021. Chomsky's (2020) Links and linker phenomena. *Quaderni di Linguistica e Studi Orientali* 7. 89–102. DOI: 10.13128/qulso-2421-7220-12004.

Manzini, Maria Rita & Diego Pescarini. 2022. The clitic string as a Pair Merge sequence. *Glossa: a journal of general linguistics* 45(1). 1–37. DOI: 10.16995/glossa.6571.

Martin, Alexander, Klaus Abels, David Adger & Jennifer Culbertson. 2019. Do learners' word order preferences reflect hierarchical language structure? In *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*, 2303–2309.

Martin, Alexander, Annie Holtz, Klaus Abels, David Adger & Jennifer Culbertson. 2020. Experimental evidence for the influence of structure and meaning on linear order in the noun phrase. *Glossa: a journal of general linguistics* 5(1). 97. 1–21. DOI: 10.5334/gjgl.1085.

Martins, Marci Fileti. 2003. *Descrição e análise de aspectos da gramática do Guarani Mbyá*. Campinas: UNICAMP dissertation. http://etnolinguistica.wdfiles.com/local--files/tese%3Amartins-2003/Martins_2003_Guarani_Mbya.pdf.

Maurits, Luke & Thomas Griffiths. 2014. Tracing the roots of syntax with bayesian phylogenetics. *Proceedings of the National Academy of Sciences* 111(37). 13576–13581. DOI: 10.1073/pnas.1319042111.

May, Robert. 1985. *Logical form: Its structure and derivation*. Cambridge (MA): MIT press. https://archive.org/details/logicalformitsst0000mayr.

McCloskey, James. 2000. Quantifier Float and Wh-Movement in an Irish English. *Linguistic Inquiry* 31(1). 57–84. DOI: 10.1162/002438900554299.

McInnerney, Andrew. 2022. *The argument/adjunct distinction and the structure of prepositional phrases*. University of Michigan dissertation. https://deepblue.lib.umich.edu/handle/2027.42/174399.

McLaughlin, Fiona. 2004. Is There an Adjective Class in Wolof? In R M W Dixon Robert Malcolm Ward & Alexandra Y Aikhenvald (eds.), *Adjective Classes: A Cross-Linguistic Typology*, 242–262. Oxford University Press. DOI: 10.1093/oso/9780199270934.003.0010.

Medeiros, David P. 2018. ULTRA: Universal Grammar as a Universal Parser. *Frontiers in Psychology* 9. 155. DOI: 10.3389/fpsyg.2018.00155.

Medeiros, David P. 2021. A parsimonious method for generating syntactic structure. Ms. https://lingbuzz.net/lingbuzz/005589/.

Meier, Sabrina C. 2020. *Topics in the grammar of Mono-Alu (Oceanic)*. Newcastle (Australia): University of Newcastle. http://hdl.handle.net/1959.13/1414765.

Melchin, Paul. 2019. *The semantic basis for selectional restrictions*. Université d'Ottawa/University of Ottawa dissertation. https://ruor.uottawa.ca/handle/10393/38831.

Merlan, Francesca C. 1994. *A Grammar of Wardaman: A Language of the Northern Territory of Australia*. Berlin ; New York: De Gruyter Mouton. DOI: 10.1515/9783110871371.

Merlo, Paola. 1996. *Parsing with Principles and Classes of Information*. Gennaro Chierchia, Pauline Jacobson & Francis J. Pelletier (red.). Dordrecht: Springer Netherlands. DOI: 10.1007/978-94-009-1708-8.

Merlo, Paola & Sarah Ouwayda. 2018. Movement and structure effects on Universal 20 word order frequencies: A quantitative study. *Glossa: a journal of general linguistics* 3(1). 84. 1–35. DOI: 10.5334/gjgl.149.

Merrill, John T. M. & Viktoria Apel. 2021. The gender system of Noon: insights into the reorganization of agreement. *STUF - Language Typology and Universals* 74(2). 347–368. DOI: 10.1515/stuf-2021-1036.

Milway, Daniel. 2022. A Parallel Derivation Theory of Adjuncts. *Biolinguistics* 16(e9313). DOI: 10.5964/bioling.9313.

Mittendorf, Ingo & Louisa Sadler. 2005. Numerals, nouns and number in Welsh NPs. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2005 conference*, 294–312. Stanford (CA): CSLI Publications. https://web.stanford.edu/group/cslipublications/cslipublications/LFG/10/pdfs/lfg05.pdf.

Moro, Andrea. 2000. *Dynamic antisymmetry*. Cambridge (MA): MIT Press. https://ia801503.us.archive.org/12/items/hard-drive-new-content-6/026213375X.MIT_Press.Dynamic_Asymmetry.Moro%2C_Andrea.Sep.2000.pdf.

Moro, Andrea. 2008. *The boundaries of Babel. The brain and the enigma of impossible languages*. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/9780262134989.001.0001.

Moro, Andrea & Ian Roberts. 2024. The duality of syntax: Unstable structures, labelling and linearisation. *Natural Language & Linguistic Theory* 42(2). 609–631. DOI: 10.1007/s11049-023-09588-z.

Morzycki, Marcin. 2016. *Modification*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511842184.

Müller, Gereon. 2010. On deriving CED effects from the PIC. *Linguistic inquiry* 41(1). 35–82. DOI: 10.1162/ling.2010.41.1.35.

Müller, Gereon. 2011. *Constraints on displacement: A phase-based approach*. Amsterdam: John Benjamins. DOI: 10.1075/lfab.7.

Müller, Gereon. 2023. Challenges for cyclicity. In Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.), *Cyclicity*, 3–72. Leipzig: Institut für Linguistik Universität Leipzig. https://www.philol.uni-leipzig.de/fileadmin/Fakult%C3%A4t_Philo/Linguistik/Forschung/LAB/LAB_95/lab95_1_M%C3%BCller.pdf.

Müller, Stefan. 2023a. *Germanic syntax: A constraint-based view*. Berlin: Language Science Press. DOI: 10.5281/zenodo.7733033.

Müller, Stefan. 2023b. *Grammatical theory: From transformational grammar to constraint-based approaches.* 5th revised. Berlin: Language Science Press. DOI: 10.5281/zenodo.7628029.

Munro, Pamela. 2020. Universal 20 Word Order and Its Occasional Variations. Ms. Universal 20 Fest, AIS, 25 August, UCLA.

Nakajima, Heizo. 1992. Another Type of Antecedent Government. *Linguistic Inquiry* 23(2). 313–328. JSTOR: 4178770.

Nchare, Abdoulaye Laziz. 2012. *The Grammar of Shupamem*. New York University dissertation. https://lingbuzz.net/lingbuzz/001441/.

Neeleman, Ad. 2017. PP-over-V meets Universal 20. *The Journal of Comparative Germanic Linguistics* 20(1). 3–47. DOI: 10.1007/s10828-017-9085-1.

Neeleman, Ad & Kriszta Szendröi. 2008. Case morphology and radical pro-drop. In Theresa Biberauer (ed.), *The Limits of Syntactic Variation*, 331–348. Amsterdam ; Philadelphia: John Benjamins. DOI: 10.1075/la.132.14nee.

Nesson, Rebecca & Stuart Shieber. 2009. Efficiently Parsable Extensions to Tree-Local Multicomponent TAG. In Mari Ostendorf, Michael Collins, Shri Narayanan, Douglas W. Oard & Lucy Vanderwende (eds.), *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 92–100. Boulder, Colorado: Association for Computational Linguistics. https://aclanthology.org/N09-1011.

Nesson, Rebecca Nancy. 2009. *Synchronous and multicomponent tree-adjoining grammars: complexity, algorithms and linguistic applications*. USA: Harvard University dissertation. https://www.proquest.com/dissertations-theses/synchronous-multicomponent-tree-adjoining/docview/304893595/se-2.

Newmeyer, Frederick J. 1998. *Language Form and Language Function*. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/4115.001.0001.

Newmeyer, Frederick J. 2005. *Possible and probable languages: a generative perspective on linguistic typology*. Oxford ; New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780199274338.001.0001.

Newmeyer, Frederick J. 2010. Accounting for rare typological features in formal syntax: Three strategies and some general remarks. In Jan Wohlgemuth & Michael Cysouw (eds.),

Rethinking Universals, 195–222. Berlin ; New York: De Gruyter Mouton. DOI: 10.1515/9783110220933.195.

Nguyễn, Hùng Tưởng. 2004. *The structure of the vietnamese noun phrase*. Boston University dissertation. https://www.proquest.com/docview/305224709/abstract/ED3782F980445E3PQ/1.

Norris, Mark. 2017a. Description and analyses of nominal concord (Pt I). *Language and Linguistics Compass* 11(11). e12266. DOI: 10.1111/lnc3.12266.

Norris, Mark. 2017b. Description and analyses of nominal concord (Pt II). *Language and Linguistics Compass* 11(11). e12267. DOI: 10.1111/lnc3.12267.

Osborne, Timothy. 2021. NPs, not DPs: The NP vs. DP debate in the context of dependency grammar. *Acta Linguistica Academica*. DOI: 10.1556/2062.2021.00001.

Osterholzer, Johannes. 2017. *New Results on Context-Free Tree Languages*. Dresden: Technische Universität Dresden dissertation. New%20Results%20on%20Context-Free%20Tree%20Languages.

Osumi, Midori. 1995. Tinrin Grammar. *Oceanic Linguistics Special Publications* (25). i–304. JSTOR: 20006758.

Ouwayda, Sarah & Ur Shlonsky. 2017. Word order variation in Lebanese Arabic DPs: In support of low numerals. *Linguistic Inquiry* 48(1). 181–193. DOI: 10.1162/LING_a_00240.

Park, Dongmyung, Fadzli Irwan Bahrudin & Ji Han. 2020. Circular Reasoning for the Evolution of Research through a Strategic Construction of Research Methodologies. *International Journal of Quantitative and Qualitative Research Methods* 8(3). 1–23. https://eajournals.org/ijqqrm/vol-8-issue-3-september-2020/circular-reasoning-for-the-evolution-of-research-through-a-strategic-construction-of-research-methodologies/.

Paul, Ileana. 2009. On the presence versus absence of determiners in Malagasy. In Jila Ghomeshi, Ileana Paul & Martina Wiltschko (eds.), *Linguistik Aktuell/Linguistics Today*, vol. 147, 215–242. Amsterdam: John Benjamins. DOI: 10.1075/la.147.07pau.

Payne, Doris L. & Thomas E. Payne. 1990. Syntax of the Sentence or the Clause. In Desmond C. Derbyshire & Geoffrey Keith Pullum (eds.), *Volume 2 Handbook Amazonian Languages: Volume 2*, 253–347. Berlin ; Boston: De Gruyter Mouton. DOI: 10.1515/9783110860382.253.

Pedrós Caballero, Toni. 2014. Grammar sketch of Bugis. Final work for the MA course ”Fieldwork methods”, at Leiden University, September-December 2013. https://www.academia.edu/30880134/Grammar_sketch_of_Bugis.

Pereltsvaig, Asya M. 2006. Head movement in Hebrew nominals: A reply to Shlonsky. *Lingua* 116(8). A1–A40. DOI: 10.1016/j.lingua.2005.11.001.

Pereltsvaig, Asya M. 2021. Noun phrases, big and small. In Sabine Laszakovits & Shen Zheng (eds.), *The size of things I: Structure building*, 323–327. Berlin: Language Science Press. DOI:%2010.5281/ZENODO.5524304.

Petzell, Erik M. 2021. Agreement inflection and word order in Viskadalian Swedish. In Ida Larsson & Erik M. Petzell (eds.), *Morphosyntactic change in Late Modern Swedish*, 277–317. Berlin: Language Science Press. DOI: 10.5281/ZENODO.5792963.

Philip, Joy Naomi. 2013. (Dis) harmony, the Head-Proximate Filter, and linkers. *Journal of Linguistics* 49(1). 165–213. DOI: 10.1017/S0022226712000163.

Piantadosi, Steven T. & Edward Gibson. 2014. Quantitative Standards for Absolute Linguistic Universals. *Cognitive Science* 38. 736–756. DOI: 10.1111/cogs.12088.

Pleshak, Polina & Maria Polinsky. Forthcoming. Functional typology and formal comparative syntax. In Maria Polinsky, Sjef Barbiers & Norbert Corver (eds.), *The Cambridge Handbook of Comparative Syntax*. http://www.mariapolinsky.com/wp-content/uploads/2022/10/Ch-6-PleshakPolinsky-revised-1.pdf.

du Plessis, Hans. 1977. Wh movement in Afrikaans. *Linguistic Inquiry*. 723–726. JSTOR: 4178017.

Pogodalla, Sylvain. 2017. A syntax-semantics interface for Tree-Adjoining Grammars through Abstract Categorial Grammars. *Journal of Language Modelling* 5(3). 527–605. DOI: 10.15398/jlm.v5i3.193.

Polinsky, Maria. 2017. Syntactic Ergativity. In *The Wiley Blackwell Companion to Syntax*, 2nd edn., 1–37. Hoboken: John Wiley & Sons. DOI: 10.1002/9781118358733.wbsyncom051.

Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20(3). 365–424. JSTOR: 4178634.

Postal, Paul M. 1969. On So-Called ”Pronouns” in English. In *Modern Studies in English: Readings in Transformational Grammar*, 201–224. Englewood Cliffs (NJ): Prentice-Hall. https://archive.org/details/modernstudiesine00reib/page/200/.

Postal, Paul M. 1972. On some rules that are not successive cyclic. *Linguistic Inquiry* 3(2). 211–222. JSTOR: 4177702.

Potsdam, Eric & Jeffrey T. Runner. 2001. Richard returns: Copy Raising and its implications. In Mary Andronis, Christopher Ball, Heidi Elston & Sylvain Neuvel (eds.), *Chicago Linguistic Society: Papers from the 37th Regional Meeting*, vol. 37–1, 453–468. https://www.researchgate.net/publication/2365997_Richard_returns_Copy_Raising_and_its_implications.

Preminger, Omer. 2020. Functional structure in the noun phrase: revisiting Hebrew nominals. *Glossa: a journal of general linguistics* 5(1). 68. 1–8. DOI: 10.5334/gjgl.1244.

Primus, Beatrice. 2001. 64. Word order typology. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, vol. 2, 2 vols., 855–873. Berlin ; New York: De Gruyter Mouton. DOI: 10.1515/9783110194265-001.

Prince, Alan & Paul Smolensky. 2004. Optimality Theory: Constraint interaction in generative grammar. In John J. McCarthy (ed.), *Optimality Theory in Phonology: A Reader*, 1–71. Oxford ; Malden (MA): Blackwell. DOI: 10.1002/9780470756171.ch1.

Prince, Alan S. & Paul Smolensky. 2002–1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Tech Report RuCCS-TR-2; CU-CS-696-93. Rutgers University. DOI: 10.7282/T34M92MV.

Pullum, Geoffrey Keith & Philip Miller. 2022. NPs versus DPs: Why Chomsky was right. Ms. University of Edinburgh & Universitè Paris Cité. https://lingbuzz.net/lingbuzz/006845/.

Punske, Jeffrey. 2014. Functional structure inside nominal phrases. In Andrew Carnie, Yosuke Sato & Daniel Siddiqi (eds.), *The Routledge Handbook of Syntax*, 1st edn., 65–88. Abingdon ; New York: Routledge. DOI: 10.4324/9781315796604.

Putnam, Michael T. 2020. One Feature—One Head: Features as Functional Heads in Language Acquisition and Attrition. In Pedro Guijarro-Fuentes & Cristina Suárez-Gómez (eds.), *New Trends in Language Acquisition Within the Generative Perspective*, 3–25. Dordrecht: Springer Netherlands. DOI: 10.1007/978-94-024-1932-0_1.

Rackowski, Andrea & Norvin W. Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic inquiry* 36(4). 565–599. DOI: 10.1162/002438905774464368.

Radford, Andrew. 2004. *Minimalist syntax: exploring the structure of english*. Cambridge: Cambridge University Press. www.cambridge.org/9780521834971.

Ramchand, Gillian & Peter Svenonius. 2014. Deriving the functional hierarchy. *Language Sciences* 46. 152–174. DOI: 10.1016/j.langsci.2014.06.013.

Rauth, Philipp & Augustin Speyer. 2021. Adverbial reinforcement of demonstratives in dialectal German. *Glossa: a journal of general linguistics* 6(1). 4. 1–24. DOI: 10.5334/gjgl.1166.

Reinhart, Tanya. 1981. Definite NP Anaphora and C-Command Domains. *Linguistic Inquiry* 12(4). 605–635. JSTOR: 4178245.

Richards, Norvin W. 1997. *What moves where when in which languages?* Massachusetts Institute of Technology Thesis. https://dspace.mit.edu/handle/1721.1/10236.

Richards, Norvin W. 2014. A-bar movement. In Andrew Carnie, Yosuke Sato & Daniel Siddiqi (eds.), *The Routledge Handbook of Syntax*, 1st edn., 167–191. Abingdon ; New York: Routledge. DOI: 10.4324/9781315796604.

Richards, Norvin W. 2016. *Contiguity theory*. Cambridge (MA): The MIT Press. DOI: 10.7551/mitpress/10542.001.0001.

van Riemsdijk, Henk C. 1998. Categorial feature magnetism: The endocentricity and distribution of projections. *The Journal of Comparative Germanic Linguistics* 2(1). 1–48. DOI: 10.1023/A:1009763305416.

Rijkhoff, Jan. 2002. *The Noun Phrase*. Oxford ; New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780198237822.001.0001.

Rijkhoff, Jan. 2015. Word Order. In James D. Wright (ed.), *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*, 644–656. Oxford: Elsevier. DOI: 10.1016/B978-0-08-097086-8.53031-1.

Rijkhoff, Jan. 2016. Crosslinguistic categories in morphosyntactic typology: Problems and prospects. *Linguistic Typology* 20(2). 333–363. DOI: 10.1515/lingty-2016-0010.

Rijkhoff, Jan & Dik Bakker. 1998. Language sampling. *Linguistic Typology* 2(3). 263–314. DOI: 10.1515/lity.1998.2.3.263.

Rijkhoff, Jan, Dik Bakker, Kees Hengeveld & Peter Kahrel. 1993. A Method of Language Sampling. *Studies in Language* 17(1). 169–203. DOI: 10.1075/sl.17.1.07rij.

Ritter, Elizabeth. 1991. Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew. In Susan Rothstein (ed.), *Perspectives on Phrase Structure: Heads and Licensing*, 37–62. Leiden (The Netherlands): Brill. DOI: 10.1163/9789004373198_004.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110883718.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge (MA): MIT Press. https://mitpress.mit.edu/9780262680615/relativized-minimality/.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In Liliane Haegeman (ed.), *Elements of Grammar*, 281–337. Dordrecht: Springer Netherlands. DOI: 10.1007/978-94-011-5420-8_7.

Rizzi, Luigi. 2001. Relativized Minimality Effects. In Mark Baltin & Chris Collins (eds.), *The Handbook of Contemporary Syntactic Theory*, 1st edn., 89–110. Oxford ; Malden (MA): Blackwell. DOI: 10.1002/9780470756416.ch4.

Rizzi, Luigi. 2013. Notes on cartography and further explanation. *Probus* 25(1). 197–226. DOI: 10.1515/probus-2013-0010.

Rizzi, Luigi. 2016. Labeling, maximality and the head – phrase distinction. *The Linguistic Review* 33(1). 103–127. DOI: 10.1515/tlr-2015-0016.

Rizzi, Luigi & Giuliano Bocci. 2017. Left Periphery of the Clause: Primarily Illustrated for Italian. In Martin Everaert & Henk C. van Riemsdijk (eds.), *The Wiley Blackwell Companion to Syntax*, 2nd edn., 1–30. Hoboken: John Wiley & Sons. DOI: 10.1002/9781118358733.wbsyncom104.

Rizzi, Luigi & Guglielmo Cinque. 2016. Functional Categories and Syntactic Theory. *Annual Review of Linguistics* 2(1). 139–163. DOI: 10.1146/annurev-linguistics-011415-040827.

Roberts, Ian G. 2017a. Harmony, Symmetry, and Dominance in Word Order Universals. In Michelle L. Sheehan, Theresa Biberauer, Ian G. Roberts & Andres Holmberg (eds.), *The Final-Over-Final Condition: A Syntactic Universal*, 27–42. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/8687.003.0006.

Roberts, Ian G. 2017b. The Final-over-Final Condition in DP: Universal 20 and the Nature of Demonstratives. In Michelle L. Sheehan, Theresa Biberauer, Ian G. Roberts & Andres Holmberg (eds.), *The Final-Over-Final Condition: A Syntactic Universal*, 151–186. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/8687.003.0011.

Roberts, Ian G. 2017c. Uniformitarianism. In Adam Ledgeway & Ian G. Roberts (eds.), *The Cambridge Handbook of Historical Syntax*, 338–359. Cambridge: Cambridge University Press. DOI: 10.1017/9781107279070.017.

Roberts, Ian G. 2018. No-Choice Parameters, Phi-Features, and the Structure of DP. In Ángel J. Gallego & Roger Martin (eds.), *Language, Syntax, and the Natural Sciences*, 114–136. Cambridge: Cambridge University Press. DOI: 10.1017/9781316591529.007.

Robins, Robert Henry. 1958. *The Yurok Language*. Berkeley ; Los Angeles: University of California Press. https://archive.org/details/rosettaproject_yur_book-1.

Rocca, Roberta, Kristian Tylén & Mikkel Wallentin. 2019. This shoe, that tiger: Semantic properties reflecting manual affordances of the referent modulate demonstrative use. *PloS one* 14(1). e0210333. DOI: 10.1371/journal.pone.0210333.

Roehrs, Dorian. 2010. Demonstrative-reinforcer constructions. *The Journal of Comparative Germanic Linguistics* 13(3). 225–268. DOI: 10.1007/s10828-010-9038-4.

Roehrs, Dorian. 2020. The Structure of Noun (NP) and Determiner Phrases (DP). In Michael T. Putnam & B. Richard Page (eds.), *The Cambridge Handbook of Germanic Linguistics*, 537–563. Cambridge: Cambridge University Press. DOI: 10.1017/9781108378291.024.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge (MA): Massachusetts Institute of Technology dissertation. https://dspace.mit.edu/handle/1721.1/15166.

Rothstein, Susan. 2017. *Semantics for Counting and Measuring*. Cambridge: Cambridge University Press. DOI: 10.1017/9780511734830.

Routamaa, Judith. 1997. *Events and participants in Kamula discourse*. Ukarumpa (Papua New Guinea): Summer Institute of Linguistics. https://www.sil.org/resources/archives/31172.

Rouveret, Alain. 2017. VSO Word Order in the Celtic Languages. In *The Wiley Blackwell Companion to Syntax*, 2nd edn., 1–41. New York: Wiley. DOI: 10.1002/9781118358733.wbsyncom092.

Royer, Justin. 2022. Decomposing definiteness: Evidence from Chuj. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 67(3). 383–418. DOI: 10.1017/cnj.2022.23.

Ruhlen, Merritt. 1975. *A guide to the languages of the world*. Stanford: Language Universals Project, Stanford University. https://archive.org/details/ruhlen-a-guide-to-the-languages-of-the-world-1975/.

Rutkowski, Paweł. 2006. Why Polish Numerals Should Not Be Analyzed as Nouns. In Changguk Yim (ed.), *Minimalist Views on Language Design: Proceedings of the 8th Seoul International Conference on Generative Grammar*, 249–263. Seoul: Hankook/Korean Generative Grammar Circle. https://www.plm.uw.edu.pl/wp-content/uploads/2012/10/Rutkowski-Why-Polish-Numerals-Should-Not-Be-Analyzed-as-Nouns.pdf.

Safir, Ken. 2008. Coconstrual and Narrow Syntax. *Syntax* 11(3). 330–355. DOI: 10.1111/j.1467-9612.2008.00117.x.

Sailor, Craig & Anoop Mahajan. 2012. Toward a derivational typology. In Thomas Graf, Denis Paperno, Anna Szabolcsi & Jos Tellings (eds.), *Theories of Everything. In Honor of Ed Keenan*, 367–381. https://phonetics.linguistics.ucla.edu/wpl/issues/wpl17/papers/42_sailor.pdf.

Saito, Mamoru. 2018. Kase as a weak head. Ms. Nanzan University. http://people.linguistics.mcgill.ca/~mcgwpl/McGWPL/2018v25n01/2018-25-1_Saito.pdf.

Salahuddin, Salahuddin. 2023. Voice systems in Manggarai: How semantic roles behave in the clauses. *Retorika: Jurnal Bahasa, Sastra, dan Pengajarannya* 16(2). DOI: 10.26858/retorika.v16i2.43382.

Salzmann, Martin. 2020. The NP vs. DP debate. Why previous arguments are inconclusive and what a good argument could look like. Evidence from agreement with hybrid nouns. *Glossa: a journal of general linguistics* 5(1). 83. 1–46. DOI: 10.5334/gjgl.1123.

Salzmann, Martin. 2022. The NP vs. DP-debate and notions of headedness. In Ulrike Freywald, Horst J. Simon & Stefan Müller (eds.), *Headedness and/or grammatical anarchy*, 55–71. Berlin: Language Science Press. https://library.oapen.org/bitstream/handle/20.500.12657/60564/1/external_content.pdf#page=63.

Sanches, Mary & Linda Slobin. 1973. Numeral classifiers and plural marking: An implicational universal. *Working Papers on Language Universals* 11. 1–22. JSTOR: 1264317.

Sauppe, Sebastian, Åshild Næss, Giovanni Roversi, Martin Meyer, Ina Bornkessel-Schlesewsky & Balthasar Bickel. 2023. An agent-first preference in a patient-first language during sentence comprehension. *Cognitive Science* 47. 13340. DOI: 10.1111/cogs.13340.

Schabes, Yves & Stuart M. Shieber. 1992. An alternative conception of tree-adjoining derivation. In *Proceedings of the 30th annual meeting on Association for Computational Linguistics*, 167–176. USA: Association for Computational Linguistics. DOI: 10.3115/981967.981989.

Schabes, Yves & Stuart M. Shieber. 1994. An alternative conception of tree-adjoining derivation. *Computational Linguistics* 20(1). 91–124. https://aclanthology.org/J94-1004.pdf.

Schabes, Yves & Richard C. Waters. 1995. Tree Insertion Grammar: A Cubic-Time, Parsable Formalism that Lexicalizes Context-Free Grammar without Changing the Trees Produced. *Computational Linguistics* 21(4). 479–513. https://aclanthology.org/J95-4002/.

Scheffler, Tatjana. 2003. Semantics of Pied-Piping and Stranding of Embedded wh-Phrases in LTAG. Ms. https://tscheffler.github.io/papers/npextraction.pdf.

Schmerling, Susan F. 2018. *Sound and Grammar: A Neo-Sapirian Theory of Language*. Leiden (The Netherlands): Brill. https://brill.com/display/title/38785.

Schwarz, Florian. 2009. *Two Types of Definites in Natural Language*. University of Massachusetts dissertation. https://florianschwarz.net/FSDiss/FS-Diss_gradschoolformat.pdf.

Schwarz, Florian. 2013. Two Kinds of Definites Cross-linguistically. *Language and Linguistics Compass* 7(10). 534–559. DOI: 10.1111/lnc3.12048.

Scontras, Gregory. 2022. On the semantics of number morphology. *Linguistics and Philosophy* 45(5). 1165–1196. DOI: 10.1007/s10988-021-09345-8.

Scontras, Gregory. 2023. Adjective Ordering Across Languages. *Annual Review of Linguistics* 9. 357–376. DOI: 10.1146/annurev-linguistics-030521-041835.

Scontras, Gregory, Judith Degen & Noah D. Goodman. 2017. Subjectivity Predicts Adjective Ordering Preferences. *Open Mind* 1(1). 53–66. DOI: 10.1162/OPMI_a_00005.

Scott, Gary-John. 2002. Stacked adjectival modification and the structure of nominal phrases. In Guglielmo Cinque (ed.), *Functional Structure in DP and IP: The Cartography of Syntactic Structures, Volume 1*, 91–120. Oxford: Oxford University Press. DOI: 10.1093/oso/9780195148794.003.0004.

Sessarego, Sandro. 2014. *The Afro-Bolivian Spanish Determiner Phrase*. Colombus: Ohio State University Press. https://muse.jhu.edu/pub/30/oa_monograph/book/36171/pdf.

Shcherbakova, Olena, Volker Gast, Damián E. Blasi, Hedvig Skirgård, Russell D. Gray & Simon J. Greenhill. 2023. A quantitative global test of the complexity trade-off hypothesis: The case of nominal and verbal grammatical marking. *Linguistics Vanguard* 9(s1). 155–167. DOI: 10.1515/lingvan-2021-0011.

Sheehan, Michelle L. 2013. Some Implications of a Copy Theory of Labeling. *Syntax* 16(4). 362–396. DOI: 10.1111/synt.12010.

Sheehan, Michelle L. 2017. The Final-over-Final Condition and the Head-Final Filter. In Michelle L. Sheehan, Theresa Biberauer, Ian G. Roberts & Anders Holmberg (eds.), *The Final-Over-Final Condition*. The MIT Press. DOI: 10.7551/mitpress/8687.003.0010.

Sheehan, Michelle L., Theresa Biberauer, Ian G. Roberts & Anders Holmberg. 2017. *The Final-Over-Final Condition: A Syntactic Universal*. Cambridge (MA): MIT Press. DOI: 10.7551/mitpress/8687.001.0001.

Sheehan, Michelle L. & Jenneke van der Wal. 2016. Do we need abstract Case? In Kyeong-min Kim, Pocholo Umbal, Trevor Block, Queenie Chan, Tanie Cheng, Kelli Finney, Mara Katz, Sophie Nickel-Thompson & Lisa Shorten (eds.), *Proceedings of the 33rd West Coast Conference on Formal Linguistics*, 351–360. Somerville (MA): Cascadilla Press. https://www.lingref.com/cpp/wccfl/33/paper3254.pdf.

Sheehan, Michelle L. & Jenneke van der Wal. 2018. Nominal licensing in caseless languages. *Journal of Linguistics* 54(3). 527–589. DOI: 10.1017/S0022226718000178.

Shieber, Stuart M. 1985. Evidence Against the Context-Freeness of Natural Language. In Walter J. Savitch, Emmon Bach, William Marsh & Gila Safran-Naveh (eds.). Red. by Robin Cooper, Elisabet Engdahl & Richard Grandy, *The Formal Complexity of Natural Language*, 320–334. Dordrecht: Springer Netherlands. DOI: 10.1007/978-94-009-3401-6_12.

Shieber, Stuart M. 2014. Bimorphisms and synchronous grammars. *Journal of Language Modelling* 2(1). DOI: 10.15398/jlm.v2i1.84.

Shim, Ji Young. 2021. *OV and VO variation in code-switching*. Berlin: Language Science Press. DOI: 10.5281/ZENODO.4456792.

Shlonsky, Ur. 2004. The form of Semitic noun phrases. *Lingua* 114(12). 1465–1526. DOI: 10.1016/j.lingua.2003.09.019.

Shlonsky, Ur. 2012. On some properties of nominals in Hebrew and Arabic, the construct state and the mechanisms of AGREE and MOVE. *Italian Journal of Linguistics* 24(2). 267–

286. https://access.archive-ouverte.unige.ch/access/metadata/ec901f9d-1474-41fb-a109-8e3c861fd74a/download.

Shlonsky, Ur. 2020. A note on the order of constituents in the Mehri Noun Phrase. *Brill's Journal of Afroasiatic Languages and Linguistics* 12(2). 151–163. DOI: 10.1163/18776930-01202004.

Shlonsky, Ur & Giuliano Bocci. 2019. Syntactic Cartography. In *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. DOI: 10.1093/acrefore/9780199384655.013.310.

Siddiqi, Daniel. 2014. The morphology–syntax interface. In Andrew Carnie, Yosuke Sato & Daniel Siddiqi (eds.), *The Routledge Handbook of Syntax*, 345–364. Abingdon ; New York: Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315796604-20/morphology%E2%80%93syntax-interface-daniel-siddiqi?context=ubx&refId=17495e9b-2232-4962-9414-3cc1eb247e0e.

Simpson, Andrew & Saurov Syed. 2020. 4 Parallels in the structure of phases in clausal and nominal domains. In Jeroen van Craenenbroeck, Cora Pots & Tanja Temmerman (eds.), *Recent Developments in Phase Theory*, 61–88. Berlin ; Boston: De Gruyter Mouton. DOI: 10.1515/9781501510199-004.

Singer, Ruth. 2006. *Agreement in Mawng: Productive and lexicalised uses of agreement in an Australian language*. University of Melbourne, Department of Linguistics and Applied Linguistics dissertation. https://minerva-access.unimelb.edu.au/bitstream/handle/11343/39232/67270_00003242_01_Agreement_in_Mawng.pdf?sequence=1.

Sischo, William R. 1979. Michoacán Nauhal. In Ronald W. Langacker (ed.), *Modern Aztec grammatical sketches: Studies in Uto-Aztecan grammar 2*, vol. 2, 307–380. Dallas: Summer Institute of Linguistics. https://www.sil.org/resources/archives/8871.

Sischo, William R. & Elena Erickson de Hollenbach. 2005. Gramática breve del Náhuatl de Michoacán. https://www.sil.org/resources/archives/63110.

Skirgård, Hedvig, Hannah J. Haynie, Damián E. Blasi, et al. 2023. Grambank reveals the importance of genealogical constraints on linguistic diversity and highlights the impact of language loss. *Science Advances* 9(16). eadg6175. DOI: 10.1126/sciadv.adg6175.

Skirgård, Hedvig, Hannah J. Haynie, Harald Hammarström, et al. 2023. *Grambank v1.0*. Version v1.0. DOI: 10.5281/ZENODO.7740140.

Slonimska, Anita, Asli Özyürek & Olga Capirci. 2022. Simultaneity as an emergent property of efficient communication in language: A comparison of silent gesture and sign language. *Cognitive Science* 46(e13133). DOI: 10.1111/cogs.13133.

Sorace, Antonella. 2000. Gradients in Auxiliary Selection with Intransitive Verbs. *Language* 76(4). 859–890. DOI: 10.2307/417202.

Soukka, Maria. 1999. *A descriptive grammar of Noon, a Cangin language of Senegal*. University of London dissertation. DOI: 10.25501/SOAS.00028800.

Sparing-Chávez, Margarethe. 2012. *Aspects of Amahuaca grammar: An endangered language of the Amazon basin*. Dallas: SIL International. https://www.sil.org/resources/publications/entry/51672.

Sportiche, Dominique. 1998. *Partitions and Atoms of Clause Structure: Subjects, Agreement, Case and Clitics*. London: Routledge. DOI: 10.4324/9780203982372.

Sportiche, Dominique. 2017. Reconstruction, Binding, and Scope. In Martin Everaert & Henk C. van Riemsdijk (eds.), *The Wiley Blackwell Companion to Syntax*, 2nd edn., 1–58. Hoboken: John Wiley & Sons. DOI: 10.1002/9781118358733.wbsyncom002.

Sproat, Richard & Chilin Shih. 1988. Prenominal Adjectival Ordering in English and Mandarin. In James P. Blevins & Juli Carter (eds.), *Proceedings of NELS 18*, 465–489. University of Massachussets Amherst: GLSA. https://hdl.handle.net/20.500.14394/36556.

Stabler, Edward P. 1997. Derivational minimalism. In Christian Retoré (ed.), *Logical Aspects of Computational Linguistics*, 68–95. Berlin ; Heidelberg: Springer. DOI: 10.1007/BFb0052152.

Stabler, Edward P. 2019. Three Mathematical Foundations for Syntax. *Annual Review of Linguistics* 5(1). 243–260. DOI: 10.1146/annurev-linguistics-011415-040658.

Starke, Michal. 2001. *Move Dissolves into Merge: a Theory of Locality* dissertation. http://www.ai.mit.edu/projects/dm/starke-diss.pdf.

Steddy, Sam & Vieri Samek-Lodovici. 2011. On the Ungrammaticality of Remnant Movement in the Derivation of Greenberg's Universal 20. *Linguistic Inquiry* 42(3). 445–469. DOI: 10.1162/LING_a_00053.

Steedman, Mark. 2020. A Formal Universal of Natural Language Grammar. *Language* 96(3). 618–660. DOI: 10.1353/lan.2020.0043.

Steele, Susan. 1978. Word order variation. A typological study. In Joseph Harold Greenberg, Edith A. Moravcsik & Charles A. Ferguson (eds.), *Universals of human language*, vol. 4: Syntax, 585–623. Stanford: Stanford University Press.

Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domains. *Syntax* 10. 80–126. DOI: 10.1111/j.1467-9612.2007.00094.x.

Stepanov, Arthur & Wei-Tien Dylan Tsai. 2008. Cartography and licensing of wh-adjuncts: a cross-linguistic perspective. *Natural Language & Linguistic Theory* 26(3). 589–638. DOI: 10.1007/s11049-008-9047-z.

Storoshenko, Dennis Ryan. 2006. Reconsidering Raising and Experiencers in English. In Tilman Becker & Laura Kallmeyer (eds.), *Proceedings of the Eighth International Workshop on Tree Adjoining Grammar and Related Formalisms*, 159–164. Sydney, Australia: Association for Computational Linguistics. https://aclanthology.org/W06-1524.

Stowell, Timothy Angus. 1981. *Origins of phrase structure*. Cambridge (MA): Massachusetts Institute of Technology dissertation. http://hdl.handle.net/1721.1/15626.

Stowell, Timothy Angus. 1989. Subjects, Specifiers, and X-bar Theory. In Mark R. Baltin & Anthony S. Kroch (eds.), *Alternative Conceptions of Phrase Structure*, 232–262. Chicago: University of Chicago Press. https://archive.org/details/alternativeconce0000unse.

Stroh-Wollin, Ulla. 2016. The emergence of definiteness marking in Scandinavian - new answers to old questions. *Arkiv för nordisk filologi* 131. 129–169. DOI: 10.63420/anf.v131i.27741.

Svenonius, Peter. 1994. The Structural Location of the Attributive Adjective. In Eric Duncan, Donka Farkas & Philip Spaelti (eds.), *Proceedings of the 12th West Coast Conference on Formal Linguistics (April 1993)*, 439–454. Stanford (CA): CSLI Publications. https://web.stanford.edu/group/cslipublications/cslipublications/site/188152633X.shtml.

Svenonius, Peter. 2024. Word Order Universals and Their Relationship to Structure. *Annual Review of Linguistics*. DOI: 10.1146/annurev-linguistics-011724-121650.

Syed, Saurov & Andrew Simpson. 2017. On the DP/NP status of nominal projections in Bangla: Consequences for the theory of phases. *Glossa: a journal of general linguistics* 2(1). 68. 1–24. DOI: 10.5334/gjgl.186.

Szabolcsi, Anna. 1983. The possessor that ran away from home. *The Linguistic Review* 3(1). DOI: 10.1515/tlir.1983.3.1.89.

Szabolcsi, Anna. 1994. The noun phrase. In *The syntactic structure of Hungarian*, 179–274. Leiden (The Netherlands): Brill. DOI: 10.1163/9789004373174_004.

Talić, Aida. 2017. Long-Form and Short-Form Prenominal Adjectives Are Not Reduced Relative Clauses in Bosnian/Croatian/Serbian. *Linguistic Inquiry* 48(1). 194–212. DOI: 10.1162/LING_a_00241.

Tang, Marc & One-Soon Her. 2019. Insights on the Greenberg-Sanches-Slobin generalization: Quantitative typological data on classifiers and plural markers. *Folia Linguistica* 53(2). 297–331. DOI: 10.1515/flin-2019-2013.

Tatsumi, Yuta. 2020. (Im)possible Constituent Orders: Nominals, Numerals, Classifiers and Ordinal Markers. *University of Pennsylvania Working Papers in Linguistics* 26(1). https://repository.upenn.edu/handle/20.500.14332/45279.

Tatsumi, Yuta. 2021. The architecture of complex cardinals in relation to numeral classifiers. *Formal approaches to number in Slavic and beyond* 5. 323. https://library.oapen.org/bitstream/handle/20.500.12657/51044/external_content.pdf?sequence=1#page=335.

Terrill, Angela. 1998. *Biri*. München: Lincom Europa. https://www.mpi.nl/publications/item_679655.

Thieberger, Nicholas. 2006. *A grammar of South Efate: an Oceanic language of Vanuatu*. University of Hawaii Press. JSTOR: j.ctvvn3t2.

Tollan, Rebecca. 2019. *Cross-linguistic effects of subjecthood, case, and transitivity in syntax and sentence processing*. University of Toronto. https://www.proquest.com/dissertations-theses/cross-linguistic-effects-subjecthood-case/docview/2317708589/se-2.

Tollan, Rebecca & Lauren Clemens. 2022. Syntactic ergativity as a constraint on crossing dependencies: The perspective from Mayan. *Linguistic Inquiry* 53(3). 459–499. https://direct.mit.edu/ling/article-abstract/53/3/459/97469.

Toyoshima, Takashi. 2013. Traversal Parameter at the PF Interface: Graph-Theoretical Linearization of Bare Phrase Structure. In Theresa Biberauer & Michelle Sheehan (eds.), *Theoretical Approaches to Disharmonic Word Order*, 358–517. Oxford University Press. DOI: 10.1093/acprof:oso/9780199684359.003.0013.

deMena Travis, Lisa. 1984. *Parameters and effects of word order variation*. Cambridge (MA): Massachusetts Institute of Technology dissertation. https://dspace.mit.edu/bitstream/handle/1721.1/15211/13089587-MIT.pdf?sequence=2.

deMena Travis, Lisa. 2010. *Inner Aspect: The articulation of VP*. Dordrecht: Springer Netherlands. DOI: 10.1007/978-90-481-8550-4.

deMena Travis, Lisa. 2014. The integration, proliferation, and expansion of functional categories: An overview. In Andrew Carnie, Yosuke Sato & Daniel Siddiqi (eds.), *The Routledge handbook of syntax*, 42–64. Abingdon ; New York: Routledge. https://api.taylorfrancis.com/content/chapters/edit/download?identifierName=doi&identifierValue=10.4324/9781315796604-4&type=chapterpdf.

Tvica, Seid. 2017. *Agreement and verb movement: The Rich Agreement Hypothesis from a typological perspective*. Utrecht: LOT dissertation. https://dare.uva.nl/search?identifier=bb78dd5b-3cb7-44b5-82d7-340dca162ea7.

Ura, Hiroyuki. 2001. Case. In *The Handbook of Contemporary Syntactic Theory*, 334–373. Oxford ; Malden (MA): Blackwell. DOI: 10.1002/9780470756416.ch11.

Uriagereka, Juan. 1999. Multiple Spell-Out. In Samuel David Epstein & Norbert Hornstein (eds.), *Working Minimalism*, 244–275. Cambridge (MA): The MIT Press. DOI: 10.7551/mitpress/7305.003.0012.

Uriagereka, Juan. 2012. *Spell-out and the Minimalist Program*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199593521.001.0001.

van de Velde, Mark. 2022. The AMAR mechanism: nominal expressions in the Bantu languages are shaped by apposition and reintegration. *Linguistics* 60(3). 899–931. DOI: 10.1515/ling-2020-0132.

Vennemann, Theo. 1974. Analogy in generative grammar: The origin of word order. In Luigi Heilmann (ed.), *Proceedings of the eleventh international congress of linguists*, vol. 2, 79–83. Bologna: il Mulino. https://www.academia.edu/27786036/Analogy_in_Generative_Grammar_The_origin_of_Word_order_1972_1975_.

Vennemann, Theo. 1977. Categorial grammar and the order of meaningful elements. In Alphonse Juilland (ed.), *Linguistic studies offered to Joseph Greenberg: On the occasion of his sixtieth birthday*, vol. 3, 615–634. Saratoga: Anma Libri. https://www.academia.edu/9491400/Categorial_grammar_and_the_order_of_meaningful_elements_1976_.

Verbeke, Saartje. 2017. V2 Word Order. In Saartje Verbeke (ed.), *Argument structure in Kashmiri: Form and function of pronominal suffixation*, 97–138. Leiden (The Netherlands): Brill. DOI: 10.1163/9789004346789_005.

Vijay-Shanker, K. 1987. *A study of tree adjoining grammars*. Ann Arbor: University of Pennsylvania dissertation. https://www.proquest.com/docview/303610666/abstract/D28ACFB1A5449AAPQ/1.

Vijay-Shanker, K. 1992. Using descriptions of trees in a tree adjoining grammar. *Computational Linguistics* 18(4). 481–518. https://aclanthology.org/J92-4004/.

Villa-García, Julio. 2019. Clitic climbing (or lack thereof) and the Copy Theory of Movement. *Glossa: a journal of general linguistics* 4(1). 84. 1–19. DOI: 10.5334/gjgl.965.

Walkden, George. 2019. The many faces of uniformitarianism in linguistics. *Glossa: a journal of general linguistics* 4(1). 52. 1–17. DOI: 10.5334/gjgl.888.

Walkden, George & Hannah Booth. 2020. Reassessing the historical evidence for embedded Verb Second. In Sam Wolfe & Rebecca Woods (eds.), *Rethinking verb second*, 536–554. Oxford University Press. DOI: 10.1093/oso/9780198844303.003.0022.

Wamsley, James Christian. 2019. Demonstratives and the Noun Phrase Structure of Hakha Chin. *Indiana Working Papers in South Asian Languages and Cultures* 1(1). DOI: 10.14434/iwpsalc2019.v1i1.27454.

Weir, David Jeremy. 1988. *Characterizing mildly context-sensitive grammar formalisms*. Ann Arbor: University of Pennsylvania dissertation. https://www.proquest.com/docview/303735833/abstract/2A5A5009F5B748B8PQ/1.

Whitman, John. 2008. The Classification of Constituent Order Generalizations and Diachronic Explanation. In Jeff Good (ed.), *Linguistic Universals and Language Change*, 233–252. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199298495.003.0010.

Wier, Thomas. 2016. A Sketch Grammar of Tonkawa. https://www.academia.edu/9242027/A_Sketch_Grammar_of_Tonkawa.

Wilhelmsen, Vera. 2019. Mbugwe (F34). In Mark van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn., 533–562. London: Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315755946-18/mbugwe-f34-vera-wilhelmsen.

Williams, Edwin. 1980. Predication. *Linguistic inquiry* 11(1). 203–238. JSTOR: 4178153.

Willis, David. 2006. Against N-raising and NP-raising analyses of Welsh noun phrases. *Lingua* 116(11). 1807–1839. DOI: 10.1016/j.lingua.2004.09.004.

Wilson, Aidan. 2013. *Tiwi Revisited: A reanalysis of Traditional Tiwi verb morphology*. University of Melbourne, Department of Linguistics and Applied Linguistics MA thesis. https://www.academia.edu/7356422/Tiwi_Revisited_A_reanalysis_of_Traditional_Tiwi_verb_morphology.

Wilson, Patricia R. 1980. *Ambulas Grammar*. Vol. 26. Ukarumpa (Papua New Guinea): The Summer Institute of Linguistics. 477. http://www.sil.org/pacific/png/abstract.asp?id=21876.

Wiltschko, Martina (ed.). 2014. *The Universal Structure of Categories: Towards a formal typology*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139833899.009.

Wiltschko, Martina. 2021a. The Syntax of Number Markers. In Patricia Cabredo Hofherr & Jenny Doetjes (eds.), *The Oxford Handbook of Grammatical Number*, 164–196. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780198795858.013.8.

Wiltschko, Martina. 2021b. Universal underpinnings of language-specific categories: A useful heuristic for discovering and comparing categories of grammar and beyond. In Luca Alfieri, Giorgio Francesco Arcodia & Paolo Ramat (eds.), *Linguistic Categories, Language Description and Linguistic Typology*, 59–100. Amsterdam ; Philadelphia: John Benjamins. DOI: 10.1075/tsl.132.03wil.

Witkoś, Jacek. 2021. Undoing Wh-Movement: On the Need for Multiple Copies. *Philosophies* 6(4). 99. DOI: 10.3390/philosophies6040099.

Witkoś, Jacek & Dominika Dziubała-Szrejbrowska. 2018. Cardinal numerals and complex numerals as specifiers. *Poznan Studies in Contemporary Linguistics* 54(4). 605–635. DOI: 10.1515/psicl-2018-0024.

Woidich, Manfred. 2006. *Das kairenisch-arabische: eine grammatik*. Wiesbaden: Harrassowitz. https://www.harrassowitz-verlag.de/titel_3090.ahtml.

Woolford, Ellen. 2006. Lexical Case, Inherent Case, and Argument Structure. *Linguistic Inquiry* 37(1). 111–130. DOI: 10.1162/002438906775321175.

Wright, Douglas Scott. 1995. *The verbal system of Logo (Logoti)*. The University of Texas at Arlington. https://search.proquest.com/openview/

c3f1fc806008fe0b2efc910b50ebe148 / 1 ? pq - origsite = gscholar&cbl = 18750&diss=y.

Wu, Jiun-Shiung & One-Soon Her. 2021. Taxonomy of numeral classifiers. In Chungmin Lee, Young-Wha Kim & Byeong-Uk Yi (eds.), *Numeral Classifiers and Classifier Languages*, 40–71. DOI: 10.4324/9781315166308-3.

Wurmbrand, Susanne. 1999. Modal verbs must be raising verbs. In Sonya Bird, Andrew Carnie, Jason D. Haugen & Peter Norquest (eds.), *Proceedings of the 18th West Coast Conference on Formal Linguistics (WCCFL 18)*. Somerville (MA): Cascadilla Press. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8d48b0f4e4cc1cd0421ac346019263b389f2f74a.

Wurmbrand, Susanne. 2001. *Infinitives: Restructuring and Clause Structure*. Berlin ; Boston: De Gruyter Mouton. DOI: 10.1515/9783110908329.

Wurmbrand, Susanne & Magdalena Lohninger. 2023. An implicational universal in complementation: theoretical insights and empirical progress. In Jutta M. Hartmann & Angelika Wöllstein (eds.), *Propositional arguments in cross-linguistic research: theoretical and empirical issues*, 183–229. Tübingen: Gunter Narr. DOI: 10.24053/9783823394105.

XTAG Research Group. 2001. *A Lexicalized Tree Adjoining Grammar for English*. IRCS-01-03. IRCS, University of Pennsylvania. https://www.cis.upenn.edu/~xtag/tech-report/.

Zamparelli, Roberto. 2000. *Layers in the Determiner Phrase*. New York: Routledge. DOI: 10.4324/9781315054360.

Zanuttini, Raffaella, Thomas Leu & Richard Kayne (eds.). 2013. *An Annotated Syntax Reader - Lasting Insights and Questions*. 1st edn. Oxford ; Malden (MA): John Wiley & Sons. DOI: 10.1002/9781394260140.

Zeijlstra, Hedde. 2007. Negation in Natural Language: On the Form and Meaning of Negative Elements. *Language and Linguistics Compass* 1(5). 498–518. DOI: 10.1111/j.1749-818X.2007.00027.x.

Zeijlstra, Hedde. 2023. FOFC and what left–right asymmetries may tell us about syntactic structure building. *Journal of Linguistics* 59(1). 179–213. DOI: 10.1017/S002222672200007X.

Zelenskii, Dmitrii. 2020. Redefining feature percolation. *Working Papers of the Linguistics Circle* 30(1). 21–30. https://journals.uvic.ca/index.php/WPLC/article/view/19373.

Zwart, Jan-Wouter. 2011. Recursion in Language: A Layered-Derivation Approach. *Biolinguistics* 5(1–2). 43–56. DOI: 10.5964/bioling.8829.