University of Verona

Department of Foreign Languages and Literatures

Doctoral Programme in Foreign Literatures, Languages and Linguistics

S.S.D. L-LIN/01 XXXIII. cycle
L-LIN/14 2017 - 2021

The Phonetics and Phonology of Sibilants

A Synchronic and Diachronic OT Typology of Sibilant Inventories

Coordinator:	Stefan Rabanus	(University of Verona)
Tutor:	Birgit Alber	(Free University of Bozen)
Co-tutors:	Silke Hamann	(University of Amsterdam)
	Alessandro Vietti	(Free University of Bozen)

Doctoral student: Joachim Kokkelmans

29 April 2021

The Phonetics and Phonology of Sibilants: A Synchronic and Diachronic OT Typology of Sibilant Inventories - Joachim Kokkelmans
Doctoral thesis
Verona, 29 April 2021

Sibilants and compound consonants are exceedingly difficult, if not impracticable, to the unlettered Hawaiian. Had we made the Hawaiian people, as we found them, pass through the Israelitish ordeal of distinguishing and pronouncing correctly the words *Sibboleth* and *Shibboleth*, to save their lives, it is not probable that one in a thousand would have succeeded (…).

Bingham (1847: 155)

Joachim:	Perché ci sono meno bus oggi?
Veronese bus driver:	Lo so però.
Joachim:	Lei sa cosa?
Veronese bus driver:	Lo sopero, ho deto, ts’è lo sopero odzi.

Abstract

This Ph.D. dissertation has a twofold purpose, consisting of a *what* (1) and a *why* (2):

1. To centralise the existing knowledge on sibilants in a structured and comprehensive overview, describing
   (a) *what* all sibilants are, phonetically as well as phonologically, and
   (b) *what* their attested patterns of co-occurrence (= inventories) and patterns of change in natural languages are.
2. To explain, based on these attested inventories and changes,
   (a) *why* sibilants pattern the way they do.
   (b) The reasons *why* they pattern that way are then integrated into several theoretical models (mainly Optimality Theory and BiPhon-Neural Networks) that make predictions about all possible and impossible sibilant inventories.

These four points correspond in fact to the four main chapters of this dissertation:
(1a) → chap. 2: *Phonetic and phonological categorisation of sibilants*
(1b) → chap. 3: *A synchronic and diachronic typological database of sibilant inventories*
(2a) → chap. 4: *The universal principles that shape sibilant inventories*
(2b) → chap. 5: *Typological models of sibilant inventories*

In one sentence, this dissertation is thus about determining *what* the different sibilants and sibilant inventories of the languages of the world are, understanding *why* they are or become structured that way, as well as predicting all possible sibilant inventories and changes between inventories based on these principles.

The main outcomes of this dissertation are:

- An overview addressing almost any aspect of sibilants, approximating a short encyclopedia (i.e. "*everything you always wanted to know about sibilants*")
- A novel explanation for the interaction of constraints and phenomena that restricts the set of possible sibilant inventories (in the concrete form of an OT typology and BiPhon-NN simulations)
- A typological database containing 258 sibilant inventories that provides a detailed insight into the sibilant diversity of the languages of the world
- An experiment that quantifies the human ability to distinguish 3, 5 and 7 places of articulation in sibilants
- Central findings allowing in turn to progress in more peripheral questions (e.g. the sibilant typology can explain the Biblical *Shibboleth*-story)

Acknowledgments

This dissertation, although officially written at the University of Verona exclusively, has largely benefitted from a cooperation and *de facto* co-tutelle with Silke Hamann and Paul Boersma at the University of Amsterdam.

The origin of this dissertation can be traced back to the day in 2017 on which I sat at a table with Birgit Alber, my principal supervisor, and said: "*Ich möchte alles erklären*". We agreed on the spot that explaining 'everything' was a bit too much for a 3-year-long dissertation, and thus decided to only explain all sibilant inventories of the world instead (something for which 3 years are not too much, either). I am therefore very grateful to her for having let me start writing on the topic that interested me most, despite the potentially gargantuesque dimensions of the undertaking. Throughout this Ph.D., Birgit helped me in innumerable ways, one of which is by showing me the power of conciseness in writing (for example, she suggested writing 'a phoneme inventory contrasting thirty-seven places of articulation for sibilant fricatives and/or affricates' as '37s', and must have saved at least one hectare of forest by suggesting this). She fully supported my research regardless of its results or the direction it was taking, and helped me acquire the tools and instinct of the researcher. I might be the person who wrote the pages of this dissertation, but the countless amount of time she spent listening to my ideas, suggesting new literature, reading and correcting chapters of it etc. makes it also somehow 'her' dissertation.

My thanks go to Silke Hamann for her co-supervision and help in many parts of this dissertation (e.g. the learnability experiment, computations with Neural Networks, DFT and a large amount of phonology in general). Many parts of this dissertation benefitted from her feedback and guidance, and she made sure I could feel at home at the University of Amsterdam.

I am furthermore grateful to Alessandro Vietti and Paul Boersma, who also (officially and *de facto*, respectively) co-supervised this dissertation and provided feedback especially for the phonetic, statistical and Neural Network parts. I am grateful to Alan Prince (who commented extensively on the OT part), Nazarré Merchant (who saved my initially hopelessly inaccurate OT typology by suggesting a stringency hierarchy), Birgit Alber, Martin Krämer, Will Bennett and other participants of the 3rd and 4th SOTA meetings (Eirini Apostolopoulou, Tor Håvard Solhaug, Olga Urek, Eric Baković, Junko Itô, Armin Mester, Nick Kalivoda, Jennifer Bellik, Max Tarlov, Richard Bibbs, Nick Van Handel, Anna Mai, Brett Hyde). In Amsterdam, I benefitted from important feedback from Silke Hamann, Paul Boersma, Klaas Seinhorst, Chao Zhou, Xinyu Zhang and Camilla Horslund. In Verona, Maria Vender helped me with the design of the learnability experiment. Nina Topintzi also gave important feedback on several parts of this dissertation.

I also thank Marta Tagliani, Jelena Živojinović, Piero Renato Costa León, Eirini Apostolopoulou, Giorgos Markopoulos, Anne Kruijt, Michela Redolfi and Tor Håvard Solhaug for linguistic discussions (and the good time we had during this Ph.D.). Additional thanks go to Piero Cossu, who helped me understand the Central Italian deaffrication pattern, Eva Zimmermann for providing a reduplication example used in this dissertation, Vivek Notani, who convinced me to write this dissertation in LaTeX and helped for the mathematical formulas in 8.3, and Jacek Kudera, who helped me understand Polish dialectal isoglosses and helped in searching participants for the experiment. Urszula Paprocka-Piotrowska, Stefano Redaelli, Jacopo Saturno, Daniele Artoni and Stefano Aloe must also be thanked for helping me find Polish participants. I am of course grateful to all anonymous participants of the experiment.

Nawel Es Sobhi (حبيبتي), my parents, family and friends are to be thanked for the moral support they provided, and for listening more or less attentively when I bombarded them with facts about sibilants (for which they had not necessarily asked). I thank Toscane Drujon d'Astros for providing various interesting examples of sibilant fronting. I am furthermore indebted to Gustav Mahler and Leonard Bernstein for providing examples of perseveration, hard work, passion and the faculty to constantly question one's work.

Contents

Abbreviations

37s	A phoneme inventory containing thirty-seven contrasting sibilant fricatives or affricates according to their place of articulation
ADT	Adaptive Dispersion Theory (Liljencrants and Lindblom 1972; Flemming 1995)
AudF	Auditory Form (as used e.g. in Boersma et al. 2020)
BiPhon	Bidirectional Phonology and Phonetics (Boersma 2011*b*)
CoG	Centre of Gravity
dB	Decibel (unit of measure of loudness)
DFT	Distinctive Feature Theory (Hall 2001; Hall and Mielke 2011)
FacTyp	Factorial Typology (in Optimality Theory)
GLA	Gradual Learning Algorithm (Boersma 1998; Boersma and Hayes 2001)
Hz	Hertz (unit of measure of frequency)
IE	Indo-European
NN	(Bidirectional) Neural Networks (Boersma et al. 2020)
OT	Optimality Theory (Prince and Smolensky 1993/2004)
PA	Property Analysis (in Optimality Theory; Alber and Prince 2016)
PG	Proto-Germanic
PIE	Proto-Indo-European
PoA	Place of articulation
SF	Surface Form (as used e.g. in Boersma et al. 2020)
SibInv	The sibilant inventory database, typologies and analyses presented in this dissertation (e.g. 'the SibInv database', 'the SibInv OT typology')
a sibinv	A sibilant fricative/affricate inventory (i.e. the sibilants used in a language)
TTCA	Tongue-Tip Constriction Area (Gafos 1999)
UF	Underlying Form (as used e.g. in Boersma et al. 2020)
VT	Violation Tableau (in Optimality Theory)

1. Introduction

1.1 State of the art and aims

Sibilants (i.e. sounds like <s, z, ch, j> in French; <s, (t)z, (t)sch> in German; <ж, з, с, ц, ч, ш, щ> in Russian) are widespread in the world's languages (Nartey 1979: 4; Maddieson 1984, 2013*b*) and well-known to all linguists, each of whom is typically aware of the existence of a certain number of different sibilants (e.g. /s, z, ʃ, ʒ, t͡ʃ/, or more rarely /ɕ, s̠, d͡z̝, z̡̠/) in a certain number of languages. Few, however, are familiar with *all* existing sibilants, as listed in (1) (not including all possible apical or laminal specifications):

(1) t͡s̻ d͡z̻ - t͡s d͡z - t͡s̠ d͡z̠ - t͡s̝ d͡z̝ - t͡ɕ d͡ʑ - t͡ʃ d͡ʒ - t͡ṣ d͡ẓ - t͡s̡̠ d͡z̡̠
s̻ z̻ - s z - s̠ z̠ - s̝ z̝ - ɕ ʑ - ʃ ʒ - ṣ ẓ - s̡̠ z̡̠

Besides this, very few linguists have addressed the question 'What are all possible co-occurrences of sibilants in a language (i.e. what are the possible sibilant inventories)?'. Indeed, as Hall (1997*b*: 85) notes in his *Phonology of Coronals*: "[o]ne issue phonologists have tended to ignore in recent years is the question of what a possible inventory is". Since 1997, several typological accounts of sibilant inventories have been added to the literature (e.g. Boersma and Hamann 2008; Seinhorst 2012; Flemming 2018; Boersma et al. 2020), although they provide partial answers to the question above compared to the full typology I propose in this dissertation, as shown in detail in 5.1. While many have worked on particular aspects of sibilants, for example in 1 or 2 languages or on 1 type of sibilants (e.g. Jongman et al. 2000 on English, Shosted 2006 on whistled sibilants, Żygis, Pape and Jesus 2012 on Slavic, Baroni 2014*b* on Italian, Stevens and Harrington 2016 on English, Ruch 2018 on Spanish, as a small random sample of the vast sibilant-related literature) and very few have worked on sibilant inventories as a whole system, there is a lack of a comprehensive understanding of *all* sibilants and *all* possible sibilant inventories.

A first **aim** of this dissertation is therefore to summarise and extend the knowledge about sibilants in a comprehensive way: their phonetic nature, their behaviour (as a class as well as within the class) in phonology, their possible contrasts and co-occurrences in sibilant inventories. Understanding the full typology of sibilant inventories allows researchers in turn to better analyse single patterns in natural languages (for example, using adequate IPA signs avoids suggesting to other researchers that e.g. Basque <s> is retroflex like Polish <sz>, or that Polish <sz> is a plain palatoalveolar [ʃ], that Hollandic Dutch <s> is phonetically identical to Flemish Dutch <s> etc.).

As highlighted in the last example, the literature abounds with what could be called ‘misrepresentations’ of sibilants, typically a confusion between PoAs such as retracted alveolar, alveopalatal, palatoalveolar, retroflex and subapical retroflex, but also misunderstandings about sibilant-related sound shifts, where e.g. preconsonantal s-retraction is misleadingly called ‘palatalisation’ (e.g. in Lorenzetti 2018) or where palatalisation is thought to yield almost any direct phonetic output (e.g. /ki/ → /t͡s/ in Telfer 2006). A non-exhaustive account of misrepresentations that stand in the way of a reliable overview of sibilants is provided in 3.1.5 for segments and 3.3.2 for sound shifts. A second **aim** is thus to shed light on such misrepresentations and to provide a more accurate unitary view on the nature of sibilants and the workings of sound shifts that target them.

This dissertation thus aims to constitute a reference for sibilant-related work (centralising the existing knowledge about sibilants) and to provide an explanation for the attested structures of sibilant inventories allied with the predictive power of linguistic frameworks (mainly Optimality Theory, Prince and Smolensky 1993/2004 and Neural Networks in Bidirectional Phonology and Phonetics, Boersma et al. 2020). I aim to show that these frameworks are complementary: NN simulations exhibit the emergent dispersion effect absent in the OT typology, and OT provides a detailed typology with sibilantless languages, voicing, affricates and the constraint interaction that transparently illustrates the action of different forces in the grammar.

The following parameters pertaining to sibilants will be taken into consideration in this dissertation:

- Place of articulation (from front to back, anterior vs. posterior sibilants)
- Distributedness (apical vs. laminal sibilants)
- Manner (sibilant fricatives vs. sibilant affricates)
- Voicing (voiceless sibilants vs. voiced sibilants)
- Gaps (voicing gaps and manner gaps)

while the following will *not* be taken into consideration:

- Secondary articulations and modifications (i.e. /s^h, s^j, s^w, s:, $s^ʕ$, s’/ all count as /s/)
- Acoustically neighbouring non-sibilant obstruents

The former of these two parameters does not contribute significantly to understanding sibilant inventories and rather pertains to sibilant-inventory-external aspects of phonology, i.e. they fall outside the scope of this dissertation. The latter exerts an influence on the sibilant inventory only limited to phonetic detail (cfr. e.g. Pétursson 1971, who shows that the precise phonetic nature of /θ/ interacts with that of /s/). At the phonological level however, their interaction is not significant, which explains why non-sibilant obstruents have been left out here (although included in the SibInv database for descriptive precision, cfr. 3.1.3).

This dissertation *does* take into account languages that have no sibilant, to which I refer as 0-sibilant inventories (abbreviated 0s). They are relevant among others in the OT typologies of chapter 5, where they result from a high ranking of sibilant markedness, and in chapter 3, where it is shown that sibilants can still surface allophonically in some sibilantless languages. This dissertation shows in fact that 0s are not the mere absence of sibilants, but an active avoidance of sibilants as phonemes, which is relevant for a typology of sibilants.

To summarise what is taken into account in this dissertation and what is not, the sibilant inventory (abbreviated sibinv) of e.g. standard Russian is represented as follows:[1]

(2) Russian (Padgett and Żygis 2007: 296; Kochetov 2017: 322)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.	t͡s̪			t͡ɕ			
Sib. fri.	s̪ z̪			ɕ		ṣ ẓ	

where place of articulation (3 PoAs), distributedness (laminal /ɕ/ vs. apical /ṣ/), manner (fricative /s̪/ vs. affricate /t͡s̪/), voicing (voiceless /s̪/ vs. voiced /z̪/) and gaps (at /d͡z̪/, /ʑ/, /d͡ʑ/, /t͡ṣ/ and /d͡ẓ/) are considered, but not e.g. the secondarily palatalised /sʲ/ (which counts as /s/), the geminate /ɕː/ (which counts as /ɕ/) or the non-sibilant fricatives /x, xʲ/.

1.2 Hypotheses

Although the inherently descriptive parts of this dissertation (the ones that centralise and summarise existing knowledge) expectedly have no particular hypothesis behind them, there are a few hypotheses that surface elsewhere in this work. The main and most important hypothesis regards the *finiteness* and *restrictedness* of sibilant inventories, both in the typological description of existing sibinvs and in the predictions made by the OT sibilant typology:

[1] In this representation of the sibilant inventory, the abbreviations in grey stand for 'sibilant affricate', 'sibilant fricative', 'dental', 'alveolar', 'retracted alveolar', 'alveopalatal', 'palatoalveolar', 'plain retroflex' and 'subapical retroflex', respectively.

All diacritics used for sibilants in this dissertation except one ([S̩]) are part of the IPA: the subscripts [S̪] for dental, [S̠] for retracted, [S̺] for apical, [S̻] for laminal and superscript [t͡s] for affricates/unitary sound combinations are from the standard IPA (International Phonetic Association 2015) and [S̫] for whistled is from the extended IPA (Ball et al. 2017). The two retroflex series, consisting of plain retroflexes such as [Ṣ] in Polish, Mandarin Chinese or Sanskrit versus the more posterior subapical or sublaminal retroflex [S̢] in Toda, are distinguished (or even not distinguished) in the literature with varying diacritics (Hamann 2003: 18-22). Ladefoged and Maddieson (1996: 15) use a subscript dot for the plain retroflexes and the hook diacritic for the subapical retroflexes, while Ball and Rahilly (1999: 56) use the retraction subscript to distinguish subapical retroflexes from plain retroflexes. To make the distinction even clearer, I use Ladefoged and Maddieson's subscript dot for plain and Ball and Rahilly's subscript retraction bar together with Ladefoged and Maddieson's hook for subapical retroflexes (i.e. [Ṣ] vs. [S̢̠]).

(3) Not any *a priori* logically possible sibilant inventory can exist in natural languages. Constraints on articulation and perception reduce the immense number of logically possible inventories to a much smaller subset of predicted sibilant inventories, thereby excluding the unpredicted ones.

For example, I predict an inventory such as /ʒ, ʑ, t͡ʃ, d͡ʐ/ to be impossible, even if it is logically possible *a priori*. This hypothesis goes opposite to a kind of 'null hypothesis' in which 'anything goes' and any sibilant inventory is expected to be possibly attested. Section 4.1 in particular argues against this 'null hypothesis', and chapter 4 in general argues based on various types of evidence (statistical, phonological, acquisitional etc.) that the attested inventories are indeed subjected to a certain number of universal principles. Figures I in this section and CXVII in the conclusion of this dissertation illustrate to what extent the predicted sibinvs of the OT and NN typology are a minuscule proportion of logically possible inventories.

The second most important hypothesis, addressed in 5.4.2, states that:

(4) There are sibilant inventories in *grammar*, and sibilant inventories as observed at the *surface*. Every surface inventory is an observable realisation of a grammar inventory, and several surface inventories can be different realisations of one and the same grammar inventory.

For example, the surface inventories /s̠, z̠, t͡s̠/ and /s̠, t͡s̠, d͡z̠/ are two different realisations of one and the same underlying 1-sibilant-inventory with voicing and affricates. They have a different gap (at /d͡z̠/ and /z̠/, respectively), but nothing distinguishes them in the underlying grammar. As children abstract the underlying grammar from the observed surface inventory during acquisition, this bears out the prediction that a native speaker of one of the two surface inventories will be able to learn and pronounce the sibilants of the other without any difficulty (e.g. a speaker of /s̠, z̠, t͡s̠/ will pronounce /d͡z̠/ precisely as [d͡z̠]). If two speakers have a different grammar inventory in their native languages, they are predicted to experience difficulties when learning to produce sibilants pertaining to the part of the grammar inventory that they do not have natively (e.g. a PoA or voicing contrasts). For example, as the two citations at the beginning of this dissertation illustrate, a native speaker of the sibilantless Hawaiian language is predicted to experience difficulties in producing any sibilant, whereas a native speaker of Veronese Italian is predicted to realise as retracted alveolar what would be palatoalveolar sibilants in standard Italian and as dental what would be alveolar affricates:

(5) Standard Italian (Vietti 2019)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ		
Sib. fri.		s z			ʃ		

(6) Veronese regional standard Italian (Avesani et al. 2017)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠				
Sib. fri.			s̠ z̠				

This explains why [s̠ɔpero] ‘strike’ uttered by the native speaker of Veronese Italian in the quote above corresponds to standard Italian [ʃɔpero], becoming homophonous with *so però* ‘I however know’. A view opposite to this grammar vs. surface inventory hypothesis would argue that the observed inventories are each also a distinct inventory in grammar (i.e. no difference between surface and grammar); this alternate view is refuted in 8.2.1.2 and 8.2.1.3, where it is shown that it allows for plenty of irrealistic languages.

To summarise these two main hypotheses, figure I illustrates how this dissertation splits the logically possible inventories into predicted and unpredicted ones, both in the grammar and at the surface. It will be shown in 5.4.2.2 that the surface inventories are

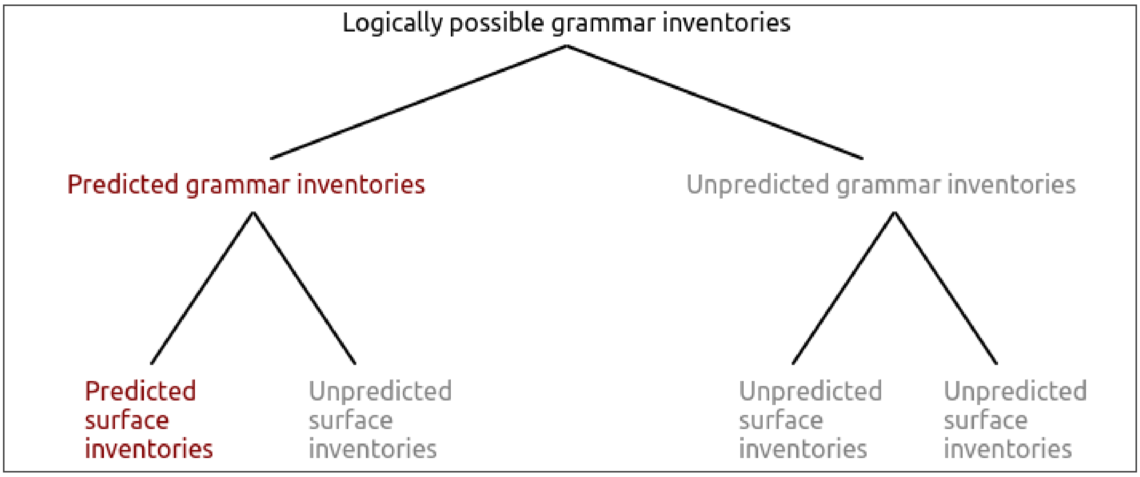


Figure I: *Possible and predicted grammar and surface sibilant inventories.*

much more numerous than the grammar inventories, and the unpredicted inventories much more numerous than the predicted ones (115 vs. 21 in grammar, 4362627495 vs. 10570 at the surface).

A fundamental assumption of this dissertation is that sibilant affricates behave as a class together with sibilant fricatives, i.e. that sibilant affricates fully behave as sibilants. In the literature, the concept of sibilant has sometimes been mingled with that of sibilant fricative (cfr. e.g. Crystal 2008: 435f., who defines a sibilant as a kind of fricative), to the exclusion or oblivion of sibilant affricates. Some views on affricates also consider them as stops without a [+strident] component, as explained by Berns (2013: 19f.), something which goes against the patterning together of sibilant fricatives and affricates illustrated in this dissertation. Nevertheless, convincing arguments have been brought forward in the literature in favour of sibilant affricates as an integral part of sibilant inventories (e.g. Hall and Żygis 2010, Berns 2013: 79). This has to do with the fact that in sibilant affricates, the phonetic stop part assimilates to the place of articulation of the sibilant part rather than the other way around (Recasens and Espinosa 2007: 146). I show in detail in 4.7 that sibilant affricates do pattern as sibilants together with sibilant fricatives.

Turning now to peripheral hypotheses in this dissertation, a hypothesis described in 3.3.2 (*Attested sibilant sound shifts*) states that when sound shifts give birth to new sibilants, these sibilants can surface *a posteriori* with different phonetic realisations (e.g. the product of stop palatalisation as [t͡s] in a language, [t͡ʃ] in another; cfr. Hall and Żygis 2010: 13), but the relevant sound shift always yields one precise phonetic output at first (e.g. [c] → [t͡ɕ] for stop palatalisation). The structure of the sibilant inventory at the time of occurrence of the sound shift then helps to predict what its phonetic output will be (e.g. stop palatalisation typically yields [t͡s̪] in a 1s but [t͡ʃ] in a well-dispersed 2s; see Kokkelmans in prep.).

Another hypothesis, tested in section 2.3 with a learnability experiment, is that sibilant inventories allow at most 4 distinct contrasting PoAs in natural languages. The data resulting from the experiment is incomplete (due to the COVID outbreak), but provisionally speaks in favour of this hypothesis.

1.3 Outline of the dissertation

The next four chapters constitute the main chapters of this dissertation. They are named:

2: *Phonetic and phonological categorisation of sibilants*

3: *A synchronic and diachronic typological database of sibilant inventories*

4: *The universal principles that shape sibilant inventories*

5: *Typological models of sibilant inventories*

Chapter 2 starts by introducing central descriptive aspects of sibilants: what they are in terms of phonetic production, acoustics and perception, how they are analysed in phonology and what role they play in sociolinguistics. The learnability experiment is also part of chapter 2.

Chapter 3 is also predominantly descriptive, yet introduces a transition towards the subsequent argumentative chapters by bringing the attention on recurrent patterns of sibilant inventories that are unlikely to result from chance rather than the action of several universal principles. The SibInv database of sibilant inventories is first presented, before addressing recurrent misanalyses of sibilants in the literature. The synchronic typology of attested sibilant inventories is then detailed, followed by a diachronic typology that illustrates attested transitions between sibilant inventories as well as sound shifts that affect sibilant inventories.

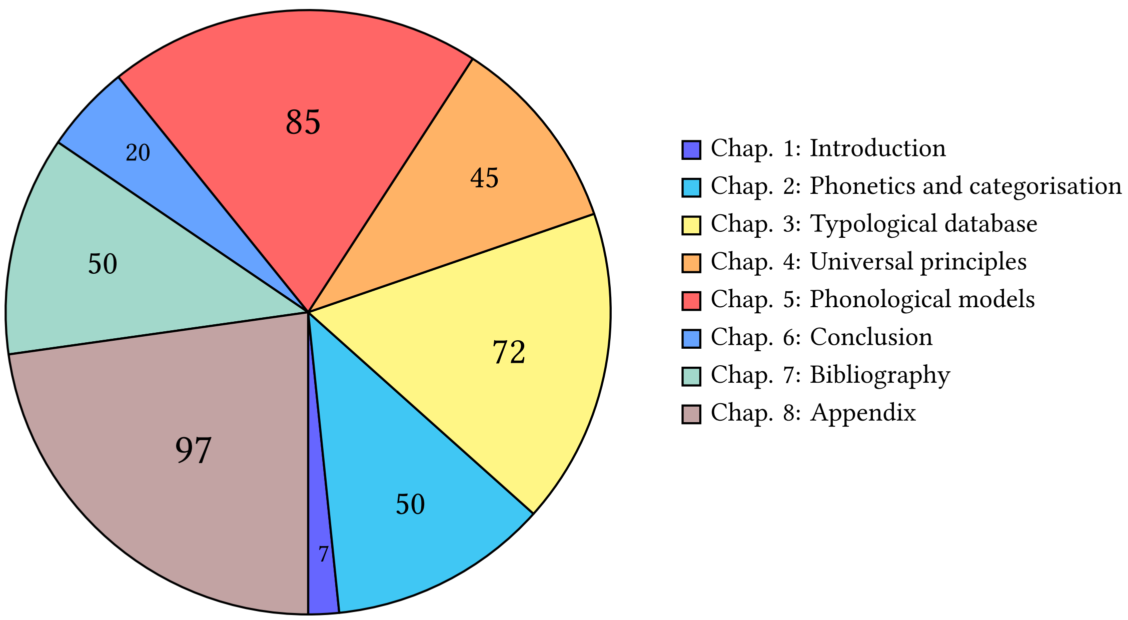
Chapter 4, based on the recurrent patterns observed in the preceding chapter, posits the existence of universal principles that restrict the possible structures of sibilant inventories in natural languages. It is first shown with the help of statistical analysis that the recurrent patterns observed are not likely to be the product of mere chance, and the 8 principles that shape sibilant inventories are then addressed. For each principle, evidence of varying nature is brought forward (e.g. language acquisition, statistical frequency or phonological processes such as assimilation and reduplication).

Chapter 5 then integrates several of these principles into predictive typological models based on phonetic-phonological frameworks. The existing accounts of sibilant inventories are first described together with their stronger and weaker points. Then, I propose a sibilant typology in Optimality Theory (Prince and Smolensky 1993/2004) and analyse it in detail. It results that the most frequent sibilant inventory is not predicted by the typology, which brings forward the need for Neural Networks in Bidirectional Phonology and Phonetics (Boersma et al. 2020). I then model sibilant inventories in a combination of OT and NN, using the software Praat (Boersma and Weenink 2019) to simulate the acquisition and transmission of each predicted or unpredicted inventory across generations. The chapter is concluded with an alternative typology in Distinctive Feature Theory (Jakobson et al. 1952 et seq.).

The conclusion chapter summarises the main findings and sheds an additional light on peripheral findings of this dissertation. It is followed by the bibliography and an appendix, which itself contains all sibilant inventories of the SibInv database.

To summarise the main structure of the dissertation in one sentence: sibilants are described comprehensively (chapter 2) before turning to the description of the ways in which they pattern in natural languages (chapter 3), from which the universal principles that shape sibilant inventories are deduced (chapter 4), which are then incorporated into several frameworks to yield a predictive typology of sibilant inventories (chapter 5).

Pie chart of the number of pages per chapter:



Total number of pages without the bibliography and appendix: **279**.

2. Phonetic and phonological categorisation of sibilants

This chapter addresses the following 4 introductory questions about sibilants, where each question loosely coincides with the corresponding section of this chapter:

1. *What* precisely is a sibilant?
2. *Why* are sibilants considered to behave as a distinct natural class?
3. *How* can sibilants be categorised based on phonetic or phonological criteria?
4. *What* role do sibilants play in society and sociolinguistics?

Section 2.3 contributes to the literature by bringing so to say 'new information', while the others summarise what is already known in the literature about the phonetics, phonology and sociolinguistics of sibilants, as an introduction to the next chapters.

2.1 Sibilants in phonetics

Each subpart of this section focuses first on all sibilants as a separate class, then on the within-class differences between sibilants. The interested reader is referred to Toda (2009) for a detailed description of the production and perception of sibilants (in French).

2.1.1 Articulation

From a phonetic articulatory perspective, sibilants are sounds produced by letting the air escape from the lungs through a narrow opening between the tongue and the dental-alveolar-palatal zone. What distinguishes them from other consonants is the narrowness of this opening and the high-frequency turbulence that results from the air subsequently striking the teeth at high speed and resonating in the cavity between teeth and tongue (Stevens 1971; Shadle 1985: 149; Ladefoged and Maddieson 1996: 145; Nozaki et al. 2011: 298). Trask (2004: 322) provides the following definition of SIBILANT:

> 1. A fricative or affricate produced with a concentration of energy at high frequencies, usually by means of a groove in the tongue, and having a characteristic hissing sound (...).
> 2. A binary distinctive feature sometimes invoked to distinguish obstacle fricatives ([+sib]) from non-obstacle fricatives ([-sib]): essentially the same feature as strident (...).

Crystal (2008: 435f.) defines a sibilant as a fricative, although it is uncontroversial that sibilant affricates are also sibilants (behaving phonologically as sibilants and consisting, phonetically, of a stop-like occlusion followed by a sibilant fricative release). He defines sibilants as sounds "made by producing a narrow, groove-like stricture between the blade of the tongue and the back part of the alveolar ridge". This definition applies to well-known sibilants such as [s] but not precisely to all sibilants, since not only the blade but also the tongue tip is primarily involved for at least some sibilants (e.g. apical [s̠̺] or subapical [s̢̠]), some sibilants are not grooved (e.g. [ɕ]; cfr. Pétursson 1971) and not only the back part of the alveolar ridge but a wide zone ranging from the teeth to the palate can in fact be in contact with the tongue. One can thus summarise what characterises sibilants as follows:

- Fricative or affricate obstruent consonants
- A narrow constriction on the range from dental to palatal, typically grooved
- Air striking the obstacle of the teeth and resonating in the anterior cavity, between teeth and tongue
- High-frequency energy due to turbulence when striking the teeth and resonating in the anterior cavity

The articulatory difference between sibilant affricates and sibilant fricatives is that for the affricates, the tongue first makes a full occlusion along the dental-alveolar-palatal range, a full occlusion that is characteristic of a coronal stop (e.g. [t], [d̠̺]), before releasing the air through the narrow opening that characterises sibilants (Berns 2014: 369). Sibilant fricatives have no such preliminary full occlusion. The production of e.g. the sibilant fricative [s] thus involves raising the tongue from its normal resting position to the alveolar zone, almost obstructing the free passage of air entirely but leaving a narrow opening in the centre, and expelling air from the lungs. Because of the constriction, the air pressure behind the opening (i.e. back in the throat) will rise, as much as – to give a concrete comparison – a crowd trying to exit a building quickly through a single door ends up pushing a lot before passing the door. The production of the sibilant affricate [t͡s] involves raising the tongue to completely obstruct the channel, building up pressure behind the tongue by ordering the air to exit the lungs, and then releasing the pressure through the narrow central opening characteristic of sibilants. Figure II provides sketches of the position of the tongue during the production of sibilants with four different places of articulation; fig. CXXI in the appendix supplements it with MRI pictures. One observes as a constant the presence of a small opening made with the tongue

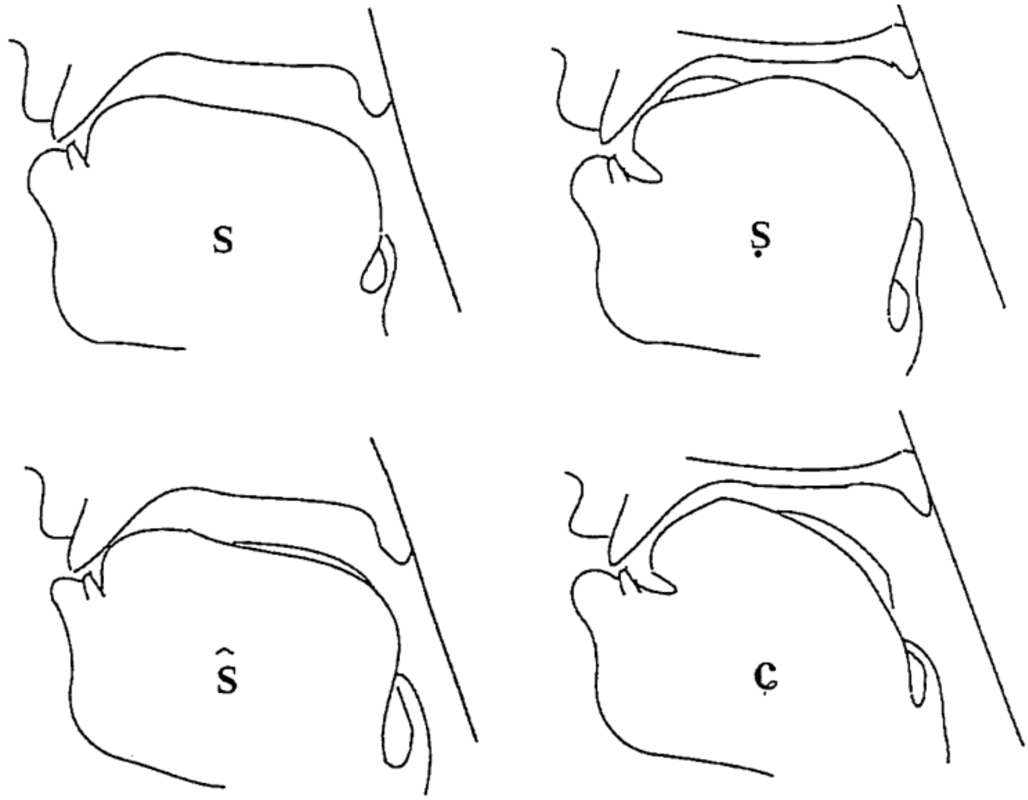


Figure II: *Midsagittal views of the articulation of four sibilant fricatives in Bzyp Abkhaz, taken from Ladefoged and Maddieson (1996: 161). The lips are on the left for each picture.*

at the dental-alveolar-palatal range, although the position and the shape of the tongue themselves are varying from a sibilant to another. Please note that the opening is central, which means that the images in fig. II are only accurate if one understands that they represent the position of the tongue at the central line rather than at its left or right sides. The sides of the tongue, illustrated with two different lines in fig. II, are in fact largely in contact with the roof of the mouth (Gafos 1999: 147), as shown in fig. III. The tongue position in fig. III is more anterior (i.e. closer to the teeth) for [s] compared to the two other sounds, alveopalatal [ɕ] and retroflex [ʂ]. The tongue shape of the sibilants differs in the sense that alveopalatal [ɕ] has a more raised blade of the tongue with respect to [s] and [ʂ], resulting in a larger contact area. Even if all sibilants have the characteristic central narrowing, differences in place of articulation and tongue shape, combined with the possibility of having a preliminary full occlusion (affricate) or none (fricative), allow

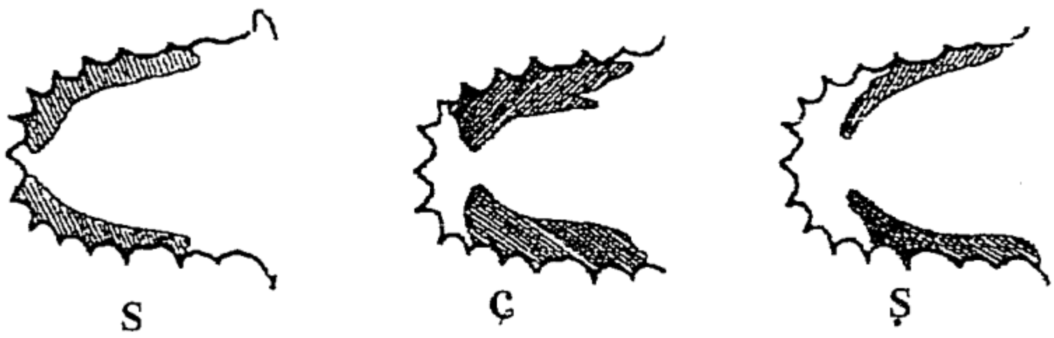


Figure III: *Sketches of the palatograms of three sibilant fricatives in Mandarin, taken and rearranged from Ladefoged and Maddieson (1996: 152). The contact area is drawn in black. Here also, the left part of the pictures is closer to the lips and the right to the throat.*

to distinguish many phonetic realisations within the class (Clements 2004: 4). Table 1 details differences in PoA and tongue shape typically used in the phonetic literature.

	dental	alveolar	retracted alveolar	alveopalatal	palatoalveolar	subapical
PoA	s̪	s	s̠	ɕ	ʃ	s̢̠
shape	s̪̻ s̪̺	s̻ s̺	s̠̻ s̠̺	ɕ	ʃ̻ ʃ̺	s̢̠̻ s̢̠̺

Table 1: Place of articulation and tongue shape in the articulation of sibilants.

The first row in table 1 corresponds only to the place of articulation, i.e. the zone that is approached by the tongue, while the second row includes the distinction between laminal segments (written with a square subscript [s̻]), which have the tongue tip typically bent downwards and most importantly a larger constriction area, and apical segments (written with a u-shaped subscript [s̺]), which have the opposite (Gafos 1999: 141f.). In his dissertation, Gafos (1999: 141-170) shows that the relevant parameter in distributedness contrasts (distributedness being the fact of being apical or laminal, i.e. what I hitherto called 'tongue shape') is not the shape of the tongue itself, but rather the area of the contact zone between the tongue and the dental-alveolar-palatal range. This new feature, called Tongue-Tip Constriction Area (TTCA), distinguishes [s] from [s̠] and [ʃ] to the extent that the area of contact is larger for the two latter with respect to the former. He does not however provide evidence for other sibilant contrasts (e.g. [s̠] vs. [ʃ]), which makes it difficult to use the TTCA parameter in this dissertation. In what follows, the term 'laminal' will be used to designate sibilants with large contact areas, even if the tongue tip is actively used, and the term 'apical' will designate sibilants with small contact areas, regardless of the implication of the tongue blade.

One observes in table 1 that the alveopalatal [ɕ] has no apical counterpart; since it has a raised tongue body that maximises the contact area with the palate, it is inherently laminal (Hall 1997*b*: 60). Note that table 1 contains a simplification, in the sense that there is a more complex interplay between place of articulation and distributedness:

[ɕ] has a zone of contact that is not very different from that of [ʃ] or [ʂ], as can be seen in fig. III, but the fact that it has the tongue tip much less curved upwards than the others leads to a smaller sublingual resonance cavity, which makes it perceptually more ‘front-sounding’, i.e. acoustically closer to [s] than [ʃ] or [ʂ] are (Bukmaier and Harrington 2016: 317; Catford 1977: 157; Dart 1991: 29). The actual place of articulation for [ɕ] and [ʃ] or [ʂ] in production is thus not very different in terms of contact zone, but because of the interaction with the tongue shape, the acoustic end result will be more front for [ɕ] (Żygis et al. 2015) as detailed in table 1, which takes both considerations into account when ordering the sibilants from front to back. The same holds for e.g. the retracted alveolar sibilants, where [s̠̺] results as an acoustically more posterior sound than its laminal counterpart [s̠̻] (cfr. Catford 1977 and Dart 1991 cited above).

Another type of sibilant is reported in the literature on Caucasian languages, but not integrated separately in the typological model of this dissertation. This sibilant is named *hissing-hushing* as well as *closed postalveolar* [ŝ] (Ladefoged and Maddieson 1996: 138, 161ff.). The term ‘closed’ refers to the absence of “the sublingual cavity that often characterizes postalveolar fricatives cross-linguistically” (Applebaum and Gordon 2013: 9), making this sibilant acoustically similar to [s]. A manuscript by J. C. Catford mentioned in Ladefoged and Maddieson (1996: 161) describes it as “acoustically and physiologically between a typical s and a typical ʃ”, something which matches a commonly encountered description of retracted alveolars. Both in the Ubykh spectrograms provided in Ladefoged and Maddieson (1996: 162) and those provided in Applebaum and Gordon (2013), the ‘hissing-hushing’ sibilant visibly has the second highest CoG (after /s/), and my DCT coefficient analysis of J.C. Catford’s Ubykh audio extracts provided in Catford and UC Berkeley (2020) confirms this in fig. IV. It thus seems that the ‘hissing-hushing’ sibilant

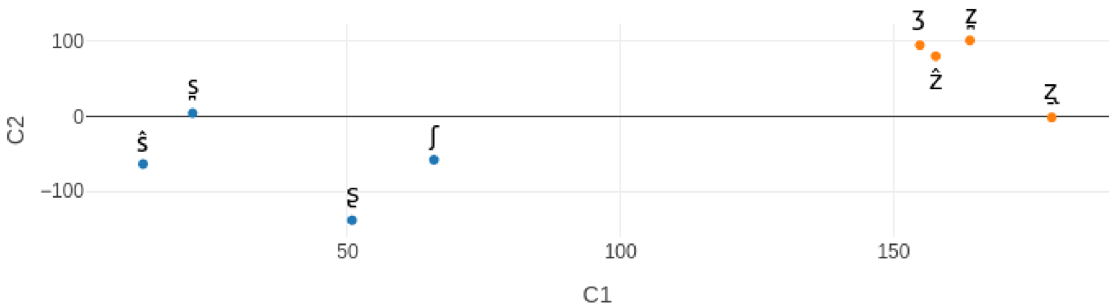


Figure IV: *DCT coefficients 1 and 2 (see 2.1.2.1) for voiceless and voiced sibilants in the recordings of Catford and UC Berkeley (2020). The quality of the voiced sibilant recording explains their undifferentiation.*

and retracted alveolar sibilants are one and the same type of sibilant (considering especially that the literature mentioned does not use the term ‘retracted alveolar’; cfr. Hall 1997*b*: 95f., who arrives at the same conclusion as I do despite using the different IPA sign [s̪] for what he describes as [+anterior, +distributed]). This becomes clear when considering the facts that they do not contrast in any language and that [ŝ] is only used specifically in the literature on Northwest Caucasian languages. In this dissertation, I will thus consider [ŝ] to be [s̠] by lack of any evidence for a different phonetic quality and by presence of evidence in favour of the opposite.

Yet another type of sibilant that is not included in this model needs to be mentioned. Adell (2016, 2018, 2019) presents new evidence in favour of the existence of a phoneme /s͡ʃ/, which behaves as a unitary segment in Chajul Ixil and corresponds to the phoneme /ʃ/ in related languages. His spectrograms (Adell 2019: 63) clearly show a strong, narrow concentration of energy in the upper frequencies, coinciding with the acoustic characteristics of apical sibilants, and a steep linear decline in frequency. It is hard to establish whether this rare pattern bears significant implications for the sibilant typology presented in this dissertation. This language with /s͡ʃ/ should count as one more attested sibilant inventory, but then a very large number of other logically possible sibilant inventories should be predicted (e.g. with /ʃ͡s/, /ɕ͡s̺/...), and it is unclear how adding these to the typology contributes to understanding the structure of sibilant inventories better. Instead, it boils down to the question whether unitary sibilants consisting of a transition between two PoAs are possible in the typology; even if we answer 'yes' to this question, the only insight gained by doing so is that any possible combination of two PoAs in a sibilant inventory should constitute a possible unitary segment. Considering this limited insight and since alveo-postalveolars are a recent discovery described by one source, I will leave the question open and not include them in this dissertation.

Sibilants are acquired late by children (Hardcastle 1976; Koenig et al. 2013: 1175). This is because the coordination of the tongue muscles and aerotactile acuity of the tongue (cfr. Francis et al. 2012) required to produce the gestures yielding a sibilant is difficult (Iskarous et al. 2011: 944). Producing a sound with a large opening (e.g. a vowel) or a complete occlusion (e.g. a stop) is easier, so that phonemes like /a/ or /t/ are acquired earlier; sibilants require the tongue to be maintained at a precise height with a specific shape, otherwise the resulting sound becomes a stop (opening too narrow) or another non-sibilant (not narrow enough) (Iskarous et al. 2011: 944).

To summarise, sibilants are characterised by a narrow stricture between the tongue and the dental-alveolar-palatal zone through which the air is projected against the teeth and resonates at high velocities. The tongue can be raised against the roof of the mouth with a large or small constriction area, which results in a laminal or apical sound respectively. It is to a certain extent possible to be articulatorily 'more' or 'less' sibilant by having the tongue more or less close to the position required for sibilantness (Trask 2004: 322; cfr. e.g. the fact that one can make a fluid transition between [ç] and [ɕ] by narrowing the opening and raising the tongue and jaw progressively), but the languages of the world only contrast binarily between sibilants and non-sibilants, rather than allowing such a thing as a 'semi-sibilant'. This is probably due to the fact that the acoustic difference between such a 'semi-sibilant' and the others is hardly perceptible.

2.1.2 Acoustics and perception

Perceptual cues used by humans to identify sibilants in speech are: Centre of Gravity, Intensity, Distributedness, Formant transitions, Affricatedness and Voicing. The first three parameters distinguish sibilants from non-sibilants as well as different sibilants within the class, while the others mainly serve to distinguish sibilants within the class.

2.1.2.1 Centre of Gravity

As mentioned above, during the production of a sibilant, the airflow passes through the narrow opening created by the constriction and hits the teeth at high speeds. Unlike non-sibilant fricatives like "/ɸ, f, θ/, which have sound generation at a surface nearly parallel to the jet (...) and /ç, x/, which have sound generation along the wall of the tube" in Shadle (1985: 149)'s speech-like model, sibilants are characterised by the fact that "the predominant sources are generated by an obstacle at right angles to the jet". The source of the noise is the direct collision of the airstream against the teeth (the obstacle), which generates high-frequency turbulence and resonance in the cavity between the tongue and the teeth (Shadle 1985, 1991; Stevens et al. 1986; Toda et al. 2010; Pont et al. 2018). Catford (1977) has shown that in the absence of teeth in subjects with false teeth, "/s/ and /š/ change drastically, losing much high-frequency energy" (Shadle 1985: 149). The collision against the teeth and the turbulence are illustrated in fig. V. The

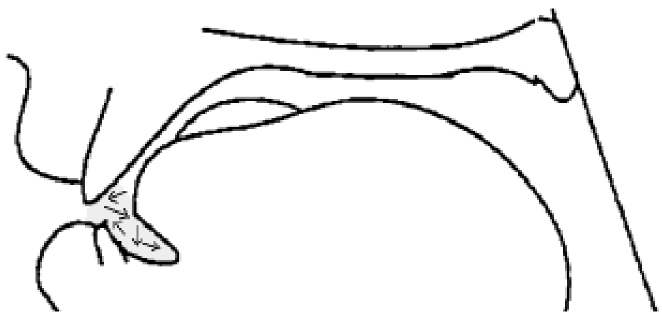


Figure V: *Detail of the resonance cavity (in grey) between the tongue and the lower teeth during the production of a sibilant, with arrows indicating the direction of airflow, taken and modified from Ladefoged and Maddieson (1996: 161).*

more anterior the constriction (closer to the lips), the higher the average *frequency* of the sound (measured in Hertz, abbreviated Hz) (Stevens 1971: 1185). This average frequency is the major acoustic cue that allows to distinguish sibilants from non-sibilants, since sibilants "such as [s] and [ʃ], have a high-frequency hiss characteristic (sibilance)" (Crystal 2008: 436), while most other sounds (though not all, see below) have energy at lower frequencies (Maddieson 1984: 42). The average frequency is *also* the main acoustic cue which distinguishes more anterior sibilants like [s], which have a smaller front cavity and thus higher frequencies, from more posterior sibilants like [ʃ], which have a larger front cavity and thus lower frequencies (Boersma and Hamann 2008: 231). The parameter of average frequency is called the Centre of Gravity (abbreviated CoG), or also 'centroid', '1st moment' and 'spectral mean' (Wrench 1995; Koenig et al. 2013). It is calculated "by weighing the frequencies in the spectrum by their power densities (Forrest et al. 1988; Jongman et al. 2000)" (Boersma and Hamann 2008: 229). This calculation is the standard setting in the phonetics/phonology software Praat (Boersma and Weenink 2019) and will be the one used in this dissertation, with a sampling frequency of 44100

Hz. Typical CoG values for the sibilant [s], for children and women or any individual with a rather high-pitched voice, exceed 6500 Hz while CoG values for adult men and individuals with lower-pitched voices are around 500∼1000 Hz lower (Jongman et al. 2000; Fox and Nissen 2005; Stuart-Smith 2020). To give an example, the upper threshold in terms of CoG reached by the author of this dissertation (25yo, male, baritone) when recording the tokens for the sibilants of the experiment in section 2.3 is 5959.68 Hz (corresponding to [s̪]); the lower threshold was 2286.98 Hz, corresponding to an extremely retracted subapical [ʂ̠]. Boersma and Hamann (2008: 229) use a range from 2000 Hz to 9000 Hz to represent the possible CoG values of sibilants, irrespectively of age and pitch.

The importance of the front resonance cavity is such that if the upper and lower teeth are too far from each other, the cavity becomes too large to generate energy at high frequencies and the sound becomes a non-sibilant such as [ç] or [h] (a process called debuccalisation, see C.4). This is witnessed by the upward motion of the jaw in sibilant production, which aims to make the front cavity small enough for high-frequency resonance (Toda et al. 2010: 343; Iskarous et al. 2011). Sibilants thus have a specific upper threshold in maxillary aperture.

The sound that is generated can be visualised on a spectrogram, which represents frequency as the vertical axis, time as the horizontal axis and energy density as darkness (Boersma 2011*a*). A spectrogram showing the sound sequence [sɔs] in the Dutch word *sociale* [sɔsja:lə] 'social' is displayed in fig. VI, right below the oscillogram. If the

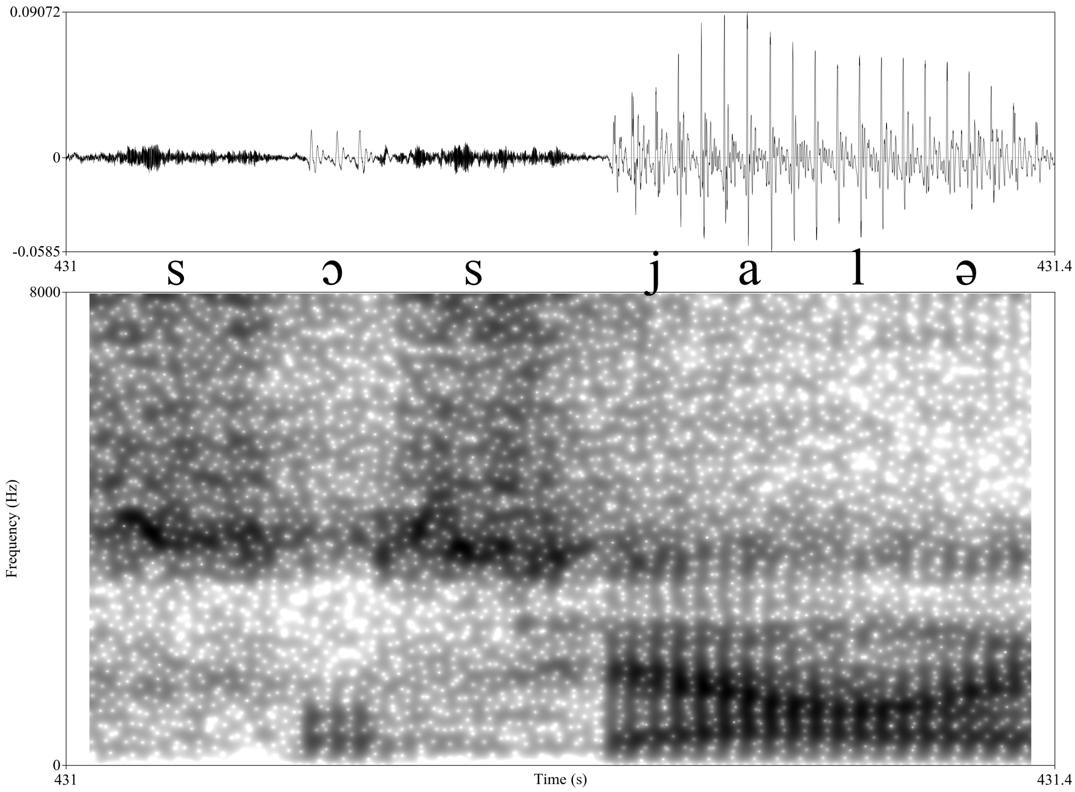


Figure VI: *Spectrogram of the sibilant [s] in the Dutch word [sɔsja:lə], pronounced by a native speaker of Flemish Dutch, taken from Kokkelmans (2017). The spectrogram is located in the lower half of the picture, below the oscillogram.*

spectrogram is very dark at a certain time and frequency, it means that a lot of the energy of the sound is concentrated at that precise time and frequency. One can see in fig. VI that most of the energy of the sibilants is concentrated at much higher frequencies than that of the vowels. Also, contrarily to e.g. vowels, sibilants are not periodic (see the well-ordered vertical slices in the oscillogram corresponding with the vowels) but made of unpredictable noise instead. This is a property of fricatives in general.

The time dimension is not always necessarily relevant, for example if one wants to know the average frequency at which energy is concentrated for a sibilant that keeps a stable energy distribution across time anyway. The time dimension can be eliminated if one averages the frequency distribution across the entire time span of a sibilant. The averaged spectral information can then be represented two-dimensionally, as the spectrum in fig. VII, which displays frequency (in Hz) as a function of intensity (in dB). Note that frequency is no longer the vertical axis but the horizontal one instead. This pho-

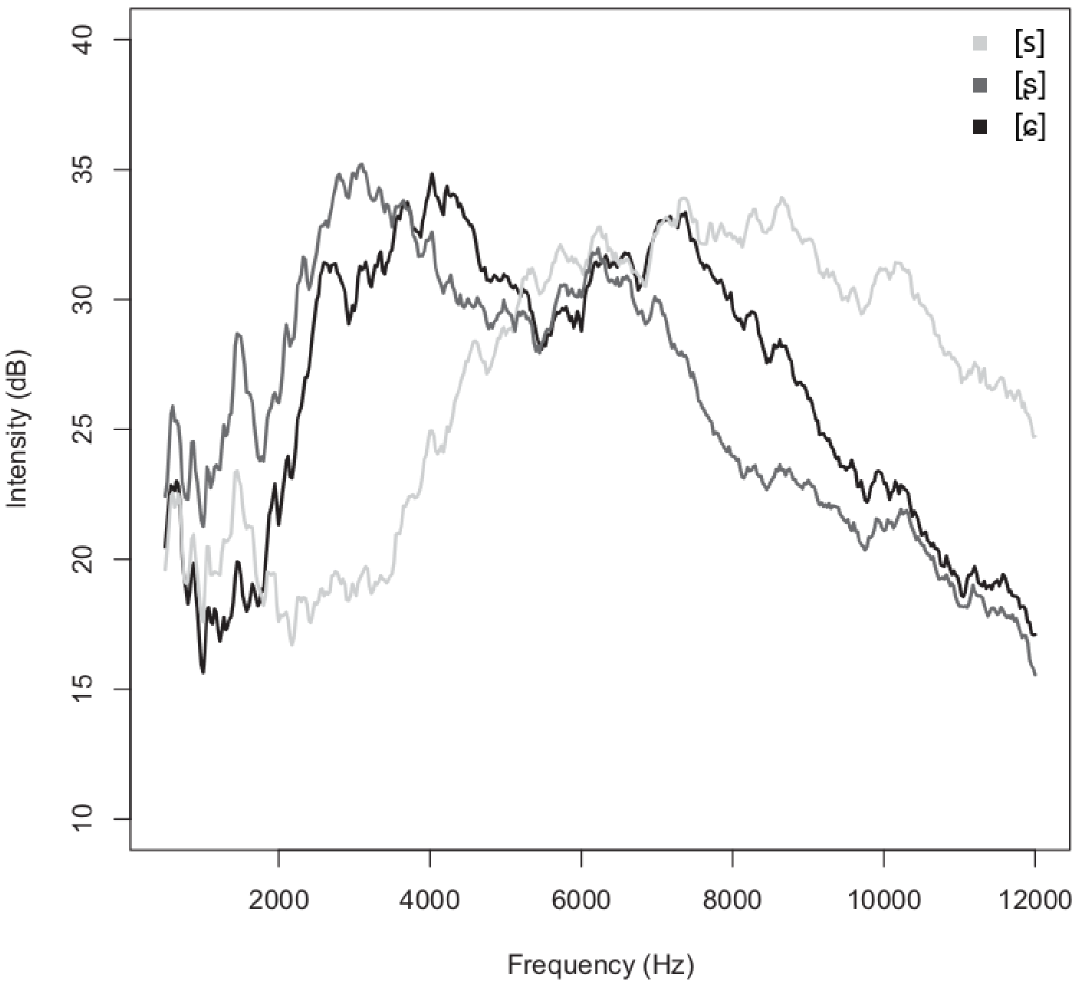


Figure VII: *Spectral representation of the frequency (Hz) as a function of intensity (dB) for the three sibilants of standard Polish, taken from Bukmaier and Harrington (2016: 319).*

netic object is called *spectrum* and is used to represent either the frequency information at a precise moment in time, or averaged across a period of time. A high value for a certain frequency on a spectrum will correspond to a dark zone at the same frequency on a spectrogram. As can be seen in fig. VII, the anterior sibilant [s] has energy at much higher frequencies than [ɕ] and [ṣ], while [ɕ] is only a bit higher in CoG than [ṣ] (recall the explanation in 2.1.1 on page 11 for this acoustic difference).

Sibilants are discernible from non-sibilants because of their higher CoG in most cases, but not always: to distinguish reliably all sibilants from non-sibilant fricatives and affricates, the shape of the spectrum is also a determining parameter (Jongman et al. 2000: 1257f.; see also 2.1.2.3). For example, a labiodental fricative [f] or interdental fricative [θ] might have a CoG similar to that of a retracted alveolar [s̠] or [ʃ] (as measured in fact by e.g. Themistocleous et al. 2016: 1027 in standard Greek, Jannedy and Weirich 2016: 72 in German), but the non-sibilants have a consistently flatter spectrum, as can be seen in fig. VIII. The spectra of sibilants typically show distinct peaks, whereas non-

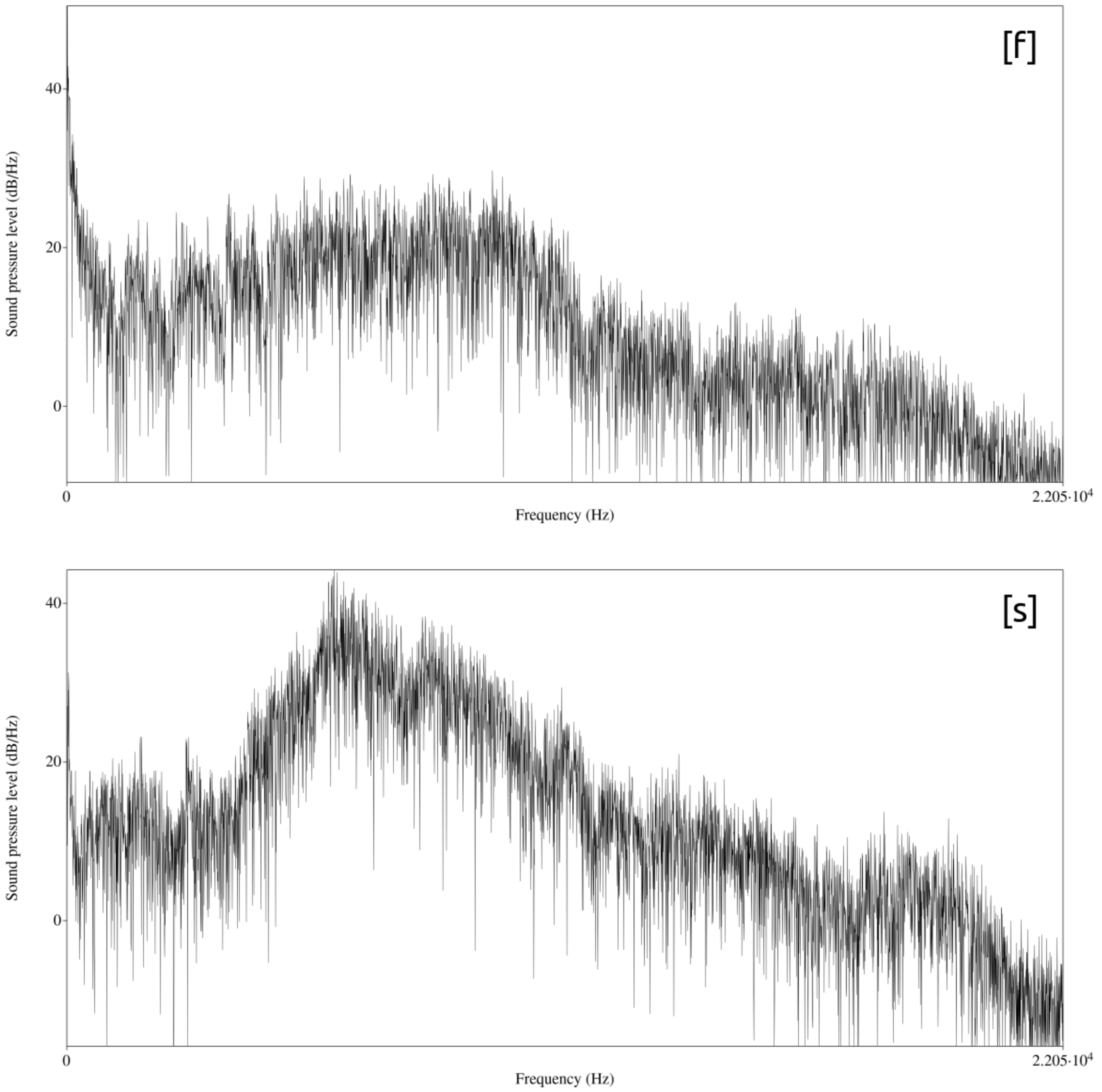


Figure VIII: *Spectrum of the non-sibilant [f] and the sibilant [s], recorded by the author of this dissertation in Praat (Boersma and Weenink 2019).*

sibilant fricatives and affricates have less abrupt differences between frequencies (Tabain 1998: 107). On a spectrogram, the 'flatness' of non-sibilant fricatives translates as a very smooth transition in blackness from low to high frequencies, while the sibilants have an abruptly prominent and intense black zone, as illustrated in fig. IX. Despite the small

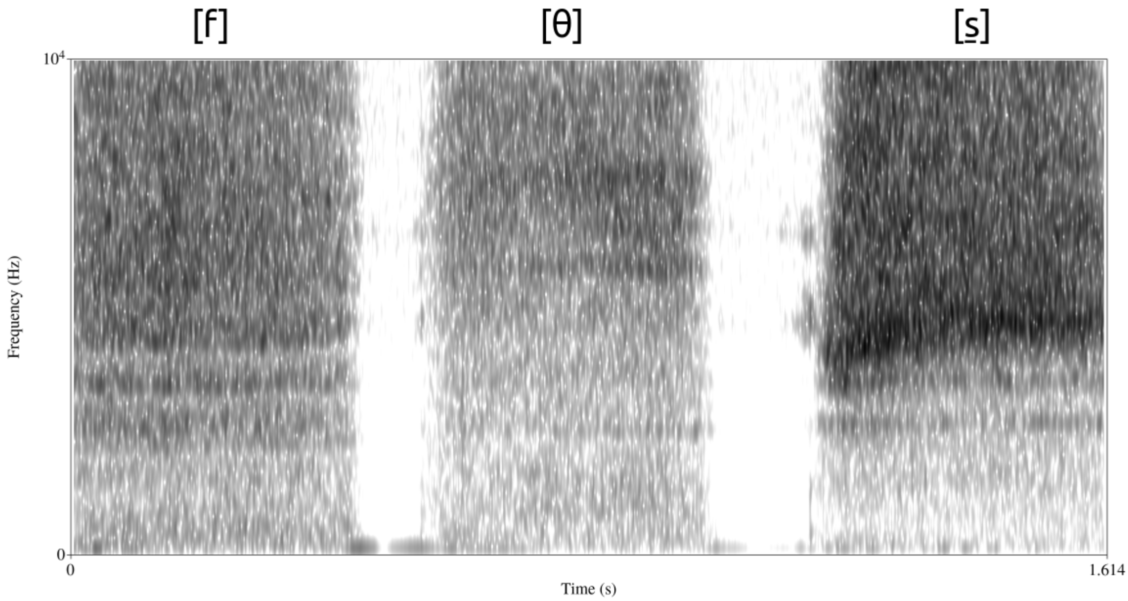


Figure IX: *Spectrogram of the non-sibilants [f] and [θ] vs. the sibilant [s̪], recorded by the author of this dissertation in Praat (Boersma and Weenink 2019).*

front cavities of [θ] and especially [f], the non-sibilants do not have large concentrations of energy in the higher frequencies as much as sibilants do. An explanation is that the airflow strikes the obstructing teeth more in sibilants, generating more noise at all frequencies and thus also the higher ones (Iskarous et al. 2011: 951ff.). A compatible explanation is provided by Flemming (2002: 20):

> In labial and non-sibilant dental fricatives the front cavity is so small that its resonant frequencies are too high to be significantly represented in the speech signal – there is little energy in the noise source at such high frequencies, and radiation losses are greater at higher frequencies. As a result the spectrum is relatively flat, but lower frequencies tend to dominate because the amplitude of the noise source is greater at lower frequencies and radiation losses are greater at higher frequencies (Fant 1960) (...).

Tabain (1998) confirms that spectral peaks and concentrations of energy are usually absent in the spectra of [f] and [θ], despite the front place of articulation of these segments. In sum, sibilants are thus best distinguished from any other sound by the *combination* of a high CoG and a peaked spectral shape (i.e. a concentration of energy at a particular frequency zone).

Over the course of an utterance containing a sibilant, the CoG rises at the beginning of the sibilant and decreases at its end, as a transition between the sibilant and the surrounding non-sibilants, which are extremely likely to have lower CoGs. It rises and decreases to a large extent (Koenig et al. 2013), with a slower and larger increase with respect to the decrease (Iskarous et al. 2011: 951). The authors last cited report up to a 2000 Hz difference between the CoG at the beginning of the sibilant /s/ and at its peak. The reported average evolution of the CoG in /s/ through time is shown in fig. X.

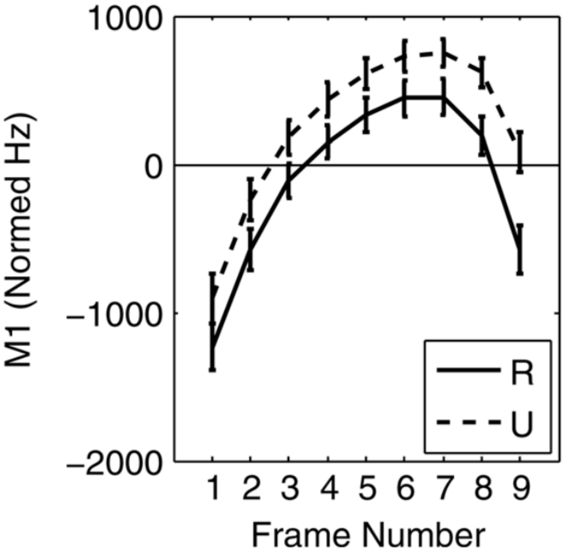


Figure X: *Average temporal trajectory of the CoG of /s/ in Hertz, divided into nine intervals from the onset to the end of the sibilant, taken from Iskarous et al. (2011: 951). 'R' stands for the context of a following rounded vowel, 'U' for an unrounded vowel (see below).*

Up to this point, the CoG parameter of sibilants has been explained in terms of *acoustics* rather than actual *perception* by human beings. Indeed, the values given in Hertz correspond to what computers are able to measure and quantify, but this does not necessarily mean that the CoG scale is an accurate representation of what humans hear (e.g. when evaluating the perceptual distance between two sibilants; Pintér 2014: 73f.). In fact, although humans can normally perceive frequencies from around 16∼20 Hz to 16000∼20000 Hz, their ears are tuned to perceive more information and fine-grained differences in the lower frequencies (from around 250 Hz to 10000 Hz), since these are the frequencies most frequently used in human speech (Evers et al. 1998; Møller and Pedersen 2004). Consequently, the ability to perceive the difference between two frequencies progressively diminishes when approaching higher frequencies, so that the auditory distance between 2000 Hz and 4000 Hz is slightly larger than that between 4000 Hz and 8000 Hz, even if the distance in Hertz of the former is precisely one half of that of the latter (Boersma and Hamann 2008: 229). To account for actual auditory distances in human perception, the ERB scale is used (Moore and Glasberg 1983), a scale that compensates this 'imbalance in favour of fine-grained lower frequencies' by ensuring that a change of 1 ERB will be equally noticed by human ears irrespectively of whether it is a change in the lower or higher frequencies. The ERB scale is derived mathematically from the Hertz scale (see Boersma 2017 for the formula). It has been shown to be better motivated than other perceptual scales such as Bark or mel (Rosner and Pickering 1994; Evers et al. 1998: 366). The relation between ERB and CoG is illustrated in fig. XI. The distinction between a CoG in Hertz or ERB is crucial in explaining both phonetic variability and

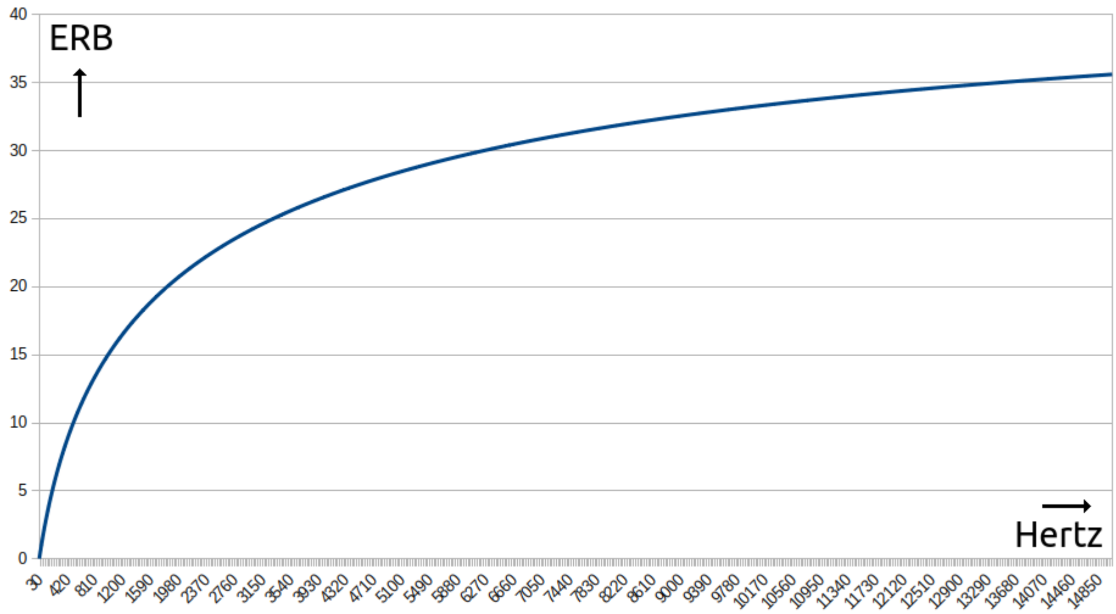


Figure XI: *Values on the ERB scale (vertically) as a function of the Hertz scale (horizontally).*

phonological contrast: for example, a [s] will be able to assume values in Hertz as diverse as +/- 5500 to 9000, all of which are perceived unambiguously as a /s/, while the variability in Hertz of a [ʂ] will be much more restricted in order to prevent phonemic confusion, namely around +/- 2500 to 4000 Hz (as visible in e.g. Howson 2015: 2671). With respect to phonological contrast, the spacing in between e.g. three well-dispersed sibilant categories is logically supposed to be *auditory* rather than acoustic space, and thus approximatively equal in terms of ERB rather than in Hertz. Figure XII shows that contrast between different sibilant phoneme categories is best represented on the ERB scale, where the two categories of a language have approximatively equal standard deviations in ERB, something which would not be the case in Hertz.

As for other segments in general (e.g. back vowels), lip rounding contributes to lowering the CoG of a sibilant (Ladefoged and Maddieson 1996: 163; Iskarous et al. 2011: 951; Koenig et al. 2013: 1177). This secondary articulation is cross-linguistically used to lower the CoG of posterior sibilants, which enhances the contrast with anterior sibilants (Flemming 1995: 8):

> Non-anterior fricatives are rounded. Palato-alveolar sibilants in French and English, and retroflex sibilants in Polish, are produced with some lip-rounding. These non-anterior sibilants are differentiated from anterior sibilants by the frequency at which energy is concentrated in the fricative spectrum. This frequency is lower in non-anterior fricatives because there is a larger resonating cavity in front of the noise source than in an anterior fricative (e.g. [s]). Lip-rounding further lowers the resonant frequency of this front cavity, and so makes non-anterior sibilants more distinct from their anterior counterparts.

In the domain of spectral analysis, there is another parameter that correlates with

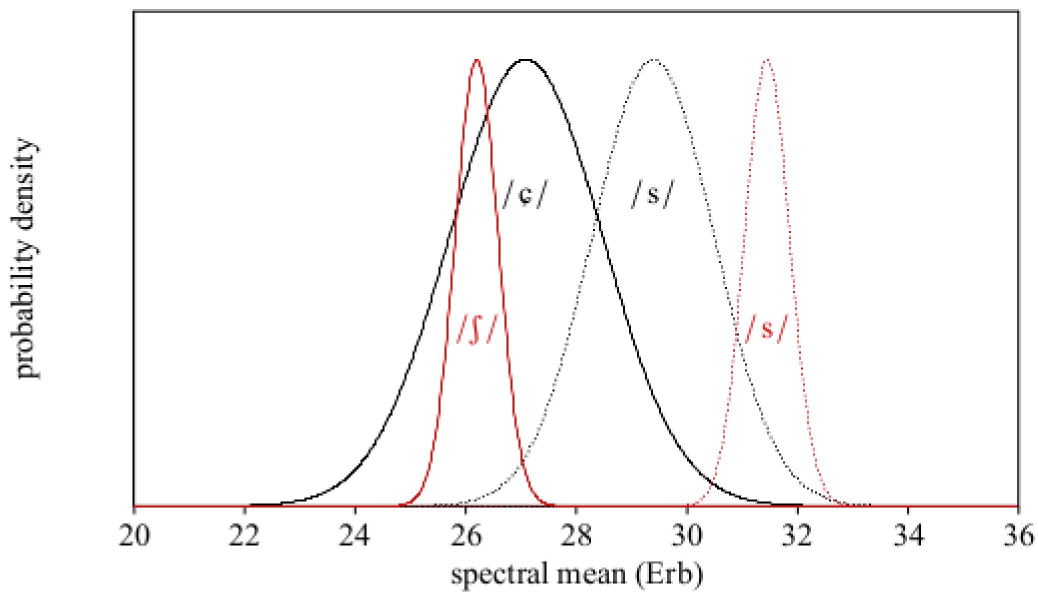


Figure XII: *Figure taken from Seinhorst (2012: 55), based on Seinhorst and Ooijevaar (2011), showing the averaged values on the ERB scale of the CoG of four sibilants in French and Hollandic Dutch: [s - ʃ] (red, French) and [s̺ - ɕ] (dark, Hollandic Dutch), respectively.*

the Centre of Gravity, called skewness or 'third moment' (of a random probability distribution). Skewness is a measure of how much a frequency distribution 'leans' towards one or the other side, i.e. whether the left or right tail of a distribution is higher than the other. Fig. XIII illustrates a distribution with negative and positive skewness, respectively. If the curve were perfectly symmetrical, its skewness would be 0. It becomes

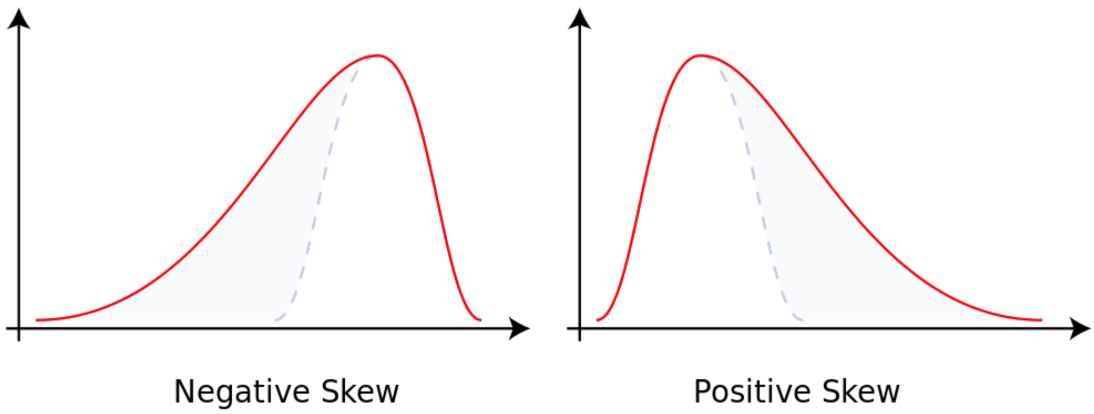


Figure XIII: *Skewness in a normal distribution. Picture taken from Wikimedia Commons (Wikimedia: Rodolfo Hermans 2008).*

apparent in fig. XIII that the spectrum for anterior sibilants, with more energy in the higher frequencies, corresponds to negative skewness while posterior sibilants and non-sibilants rather have a spectral shape with a positive skewness. Superimposing the spectra of Polish [s] and [ʂ] from fig. VII and the curves from fig. XIII, one obtains fig. XIV, in which the curve with negative skewness accurately matches the spectrum of [s] and the one with positive skewness that of [ʂ]. As Koenig et al. (2013: 1177) report, the CoG

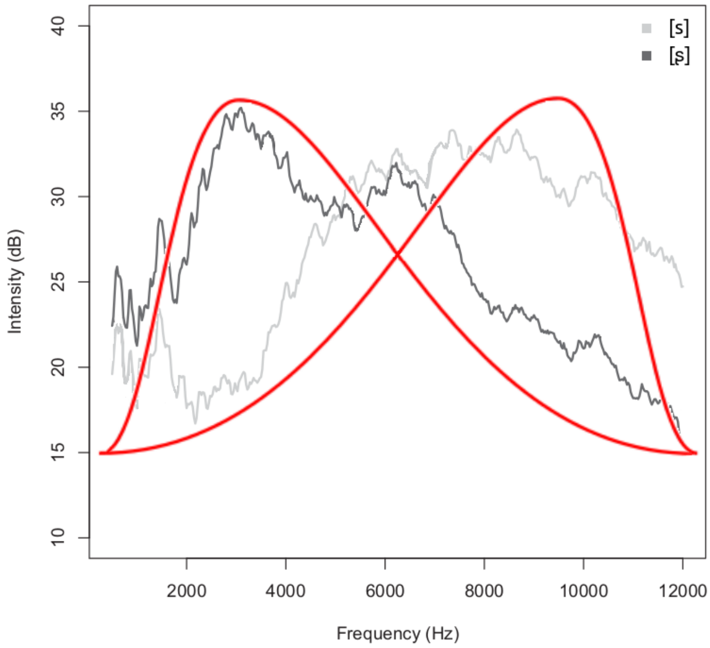


Figure XIV: *Positive and negative skewness shapes of a normal distribution (modified from fig. XIII), superimposed on the spectrum of [s] and [ʂ] (modified from fig. VII).*

of a sibilant and its skewness "can be strongly correlated", more precisely as an inverse correlation (higher CoG = lower skewness; lower CoG = higher skewness).

To calculate the spectral moments (CoG and skewness, as already described, but also Standard Deviation and kurtosis, as described further below), a Fourier transform is applied to the waveform of the relevant sound. There exists another relevant method to analyse the properties of sibilants, named *Discrete Cosine Transform*. Applied to the power spectrum, a DCT analysis allows to obtain measurements (called DCT coefficients, similar to cepstral coefficients) that have been used in the literature to distinguish fricative spectra (Harrington 2010; Pintér 2014; Spinu et al. 2018). The DCT coefficients have been shown to distinguish certain (if not all) fricatives better than the 'classical' spectral moment method, for example the [ç] - [ʃ] contrast in German (Jannedy and Weirich 2016, 2017). The DCT coefficient k_0 corresponds to the mean amplitude, k_1 to the slope, k_2 to the curvature and k_3 to the amplitude at higher frequencies of the relevant sound (Harrington 2010: 207). The four DCT coefficients $k_{0\text{-}3}$ calculated on a sibilant spectrum correspond to the similarity between the shape of the spectrum and the shape of the curve $y = cos(0.5 \cdot {}_{0\text{-}3} \cdot x)$ over the frequency domain. For example, the coefficient k_1 is proportional to the similarity between the spectrum and the curve of $y = cos(0.5 \cdot 1 \cdot x)$ (and $y = cos(0.5 \cdot 2 \cdot x)$ for k_2 etc.), as represented in fig. XV (Jannedy and Weirich 2017: 398). Since [ʃ] has most of its energy in the mid-lower frequencies, its spectrum fits the curve of the coefficient k_1 (in blue in fig. XV) quite accurately, resulting in a large positive number. The spectrum of [s], on the other hand, would fit the curve inversely, in the sense that it fits the inverse of the k_1 curve (growing from $y = -1$ to 1), resulting in a low or negative number. The value of the coefficient k_1 will thus be significantly higher

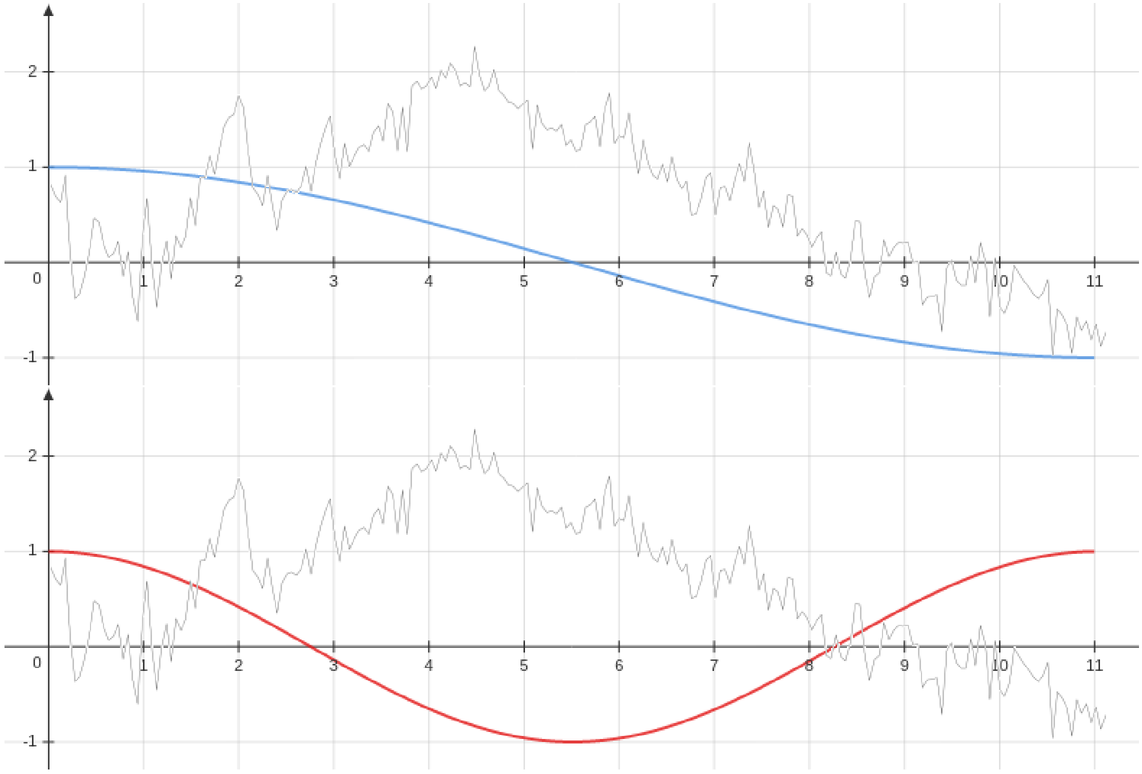


Figure XV: *Curves for the DCT coefficients k_1 (above) and k_2 (below), superimposed with the spectrum of a southern German [ʃ] in the domain 0 - 11000 Hz.*

for [ʃ] than for [s]. The curve of the coefficient k_2 (in red in fig. XV) fits the spectrum of [ʃ] very strongly but inversely, since it has most of its energy precisely in the middle frequency range while [s] has in common with the curve of k_2 that energy is high in the high frequencies. Its resulting value will thus be significantly higher than that for [ʃ]. The advantage of DCT coefficients over spectral moments is that the CoG, as a weighted average, is a single number for a vast range of possible shapes (Toda 2009: 97) while DCT coefficients evaluate the *entire* shape of the sibilant from *several* perspectives, thus yielding more precision.

2.1.2.2 Intensity

Intensity, which is correlated with amplitude and loudness in the domain of human perception (hence their interchangeable use in the literature), is relevant to distinguish sibilants as a class as well as within the class. Sibilants are distinguished from non-sibilant fricatives and affricates by their higher intensity (measured as relative intensity with respect to the following vowel, in dB), both in the noise and in the overall intensity of the sound (Maddieson 1991; Flemming 2002: 20f.; Berns 2013: 4). Intensity also serves to distinguish sibilants specifically from [f] and [θ], which “have a very low intensity” (Tabain 1998: 107f.). As she notes (1998: 109):

> Shadle (1985) showed that the greater amplitude in [s] and [ʃ] is due to the presence of an obstacle, the lower teeth, some 3 cm downstream from the

noise source at the constriction. This obstacle serves to increase the turbulence of the airflow and thereby its amplitude. This is one of the characteristic features of sibilant fricatives.

This was evidenced e.g. in English (Jongman et al. 2000: 1259) as well as in Greek, where the mean normalised amplitude of the sibilants reported by Nirgianaki (2014: 2972) is 7.84 dB higher than that of non-sibilant fricatives. On the intensity scale of segment classes, which corresponds somewhat loosely to the sonority hierarchy (Flemming 2002: 21, 24), sibilants take in the position between non-sibilant fricatives and nasal sonorant consonants (see 2.2 for the place of sibilants in the sonority hierarchy).

Within the class of sibilants, Jongman et al. (2000: 1259) report higher noise amplitudes for posterior [ʃ], [ʒ] in English compared to [s], [z], in line with the findings of Hedrick and Ohde (1993). In Mandarin Chinese, Li and Zhang (2017) report the same pattern for alveopalatal [ɕ], which is louder than [s̺]. It thus seems safe to assume that posterior sibilants are louder than anterior ones, and much louder than non-sibilant fricatives and affricates, something which contributes to distinctions as a class (sibilants > non-sibilants) and within the class (posterior > anterior sibilants).

2.1.2.3 Distributedness

In 2.1.2.1, it has been shown that sibilants are distinguished from non-sibilants based on the peaked vs. flat shape of their spectrum, respectively (Flemming 2002: 20). This has already been illustrated for [f] and [θ] vs. [s̺] in fig. IX, and is further illustrated for [ç] vs. [ʃ] (measured in German by Jannedy and Weirich 2017: 396) in fig. XVI. One can

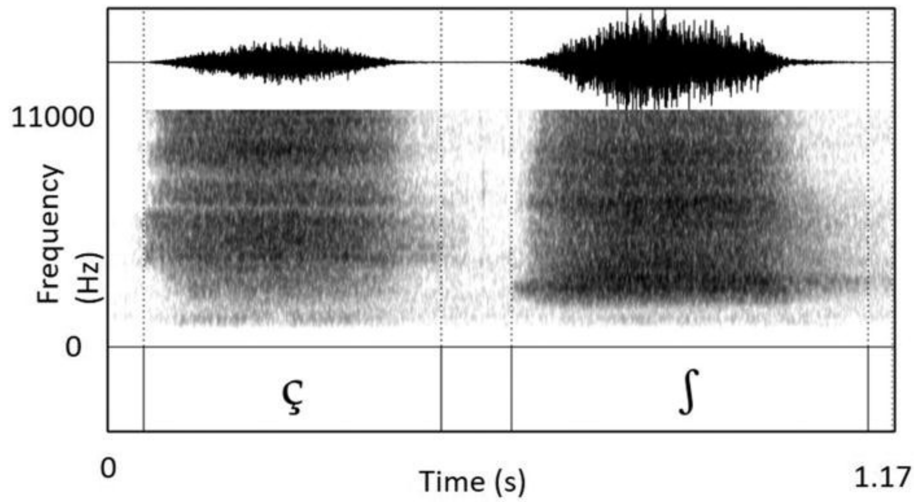


Figure XVI: *Oscillogram and spectrogram of [ç] and [ʃ] in German, taken from Jannedy and Weirich (2017: 396).*

also notice on the oscillogram that [ʃ] has a larger intensity than [ç], as predicted in the previous subsubsection. To measure peakedness or flatness, the second spectral moment (Standard Deviation) and the fourth (kurtosis) provide a means of quantification (Li et al. 2009: 115). Standard deviation (calculated as the square root of the variance), which is a measure of how much the values deviate from their mean in a normal probability distribution, allows to distinguish peaked distributions (with a smaller SD) and flatter distributions (with a larger SD). In Brazilian Portuguese for example, "the variance values for the labiodental fricatives are higher than for the alveolar and postalveolar, that is, the variance is capable of differentiating the non-sibilant fricatives (labiodental) from the sibilants (alveolar and postalveolar)" (Ferreira-Silva et al. 2015: 376). This corroborates the findings of Jongman et al. (2000) for English and of Nirgianaki (2014: 2969) for Greek, in which sibilants have the lowest SD of all Greek fricatives ([f v θ ð s̠ z̠ ç ʝ x ɣ]). The SD or variance is higher in the non-sibilant fricatives and affricates (Lee-Kim et al. 2014: 54), because their spectral shape is more diffuse and a large portion of the distribution can be found further away from the mean.

The fourth spectral moment, kurtosis, is defined as follows by Jongman et al. (2000: 1253):

> [K]urtosis is an indicator of the peakedness of the distribution. Positive kurtosis values indicate a relatively high peakedness (the higher the value, the more peaked the distribution), while negative values indicate a relatively flat distribution. Positive kurtosis thus suggests a clearly defined spectrum with well-resolved peaks, while negative kurtosis indicates a flat spectrum without clearly defined peaks.

Similarly to the relation between CoG and skewness, there is an inverse correlation between SD and kurtosis: the former is expressed in Hertz in both pairs, and coincides with a negative value of the latter if it is high or inversely, in both cases. However, SD and kurtosis only correlate within the class of sibilants; in non-sibilant fricatives, larger kurtosis values are found at posterior places of articulation (Jassem 1995: 252; Flemming 2002: 20; Nirgianaki 2014: 2969). Only SD thus reliably distinguishes sibilants from non-sibilant fricatives and affricates, while both SD and kurtosis distinguish laminal from apical sibilants, as the next paragraph will show.

Within the class of sibilants, one observes that the two acoustic parameters of distributedness (SD and kurtosis) correlate with distributedness contrasts, although distributedness is cross-linguistically used as a secondary contrast parameter wrt. Centre of Gravity (e.g. if there is a distributedness contrast, there is one in CoG, but not necessarily the other way around; Toda et al. 2010: 356). For example, Ferreira-Silva et al. (2015: 376) report a slightly higher SD for Brazilian Portuguese [s] wrt. [ʃ], although this difference is not consistent across speakers. This was expectable for a language with only two already well-dispersed PoAs, in which distributedness is not the primarily distinguishing parameter for phonemic contrast (cfr. Dart 1991, 1993): Portuguese [s] and [ʃ] are distinguished by their CoG rather than by distributedness, which can thus freely vary according to the speaker. Acoustic studies on English fricatives (Nittrouer 1995: 525; McFarland et al. 1996: 1097; Jongman et al. 2000: 1254) report lower SD and higher kurtosis values for English [s] than [ʃ], which indicates an apical peaked distribution for [s], as confirmed articulatorily by Dart (1991). Unlike Brazilian Portuguese [s], English [s] contrasts with a flat laminal [θ], which explains the more consistently apical [s]. The SD of alveopalatals in Polish is significantly higher than that of retroflexes (Żygis et al.

2015; t-values between 3.71 and 8.412) and the kurtosis of the former is significantly lower than that of the latter (Żygis et al. 2015; t-values between 1.897 and 6.234), both of which indicate that Polish [ɕ] is consistently laminal where [ṣ] is apical. This is expected in a language that contrasts not only CoG (anterior vs. posterior) but also distributedness ([ɕ] vs. [ṣ]). In O'odham, which phonemically contrasts an apical retracted alveolar [s̺̠] with a dental [s̪] (Dart 1991: 36), the peakedness in the apical sibilant is especially marked, as much as the flatness in the dental sibilant, as shown in fig. XVII. Interest-

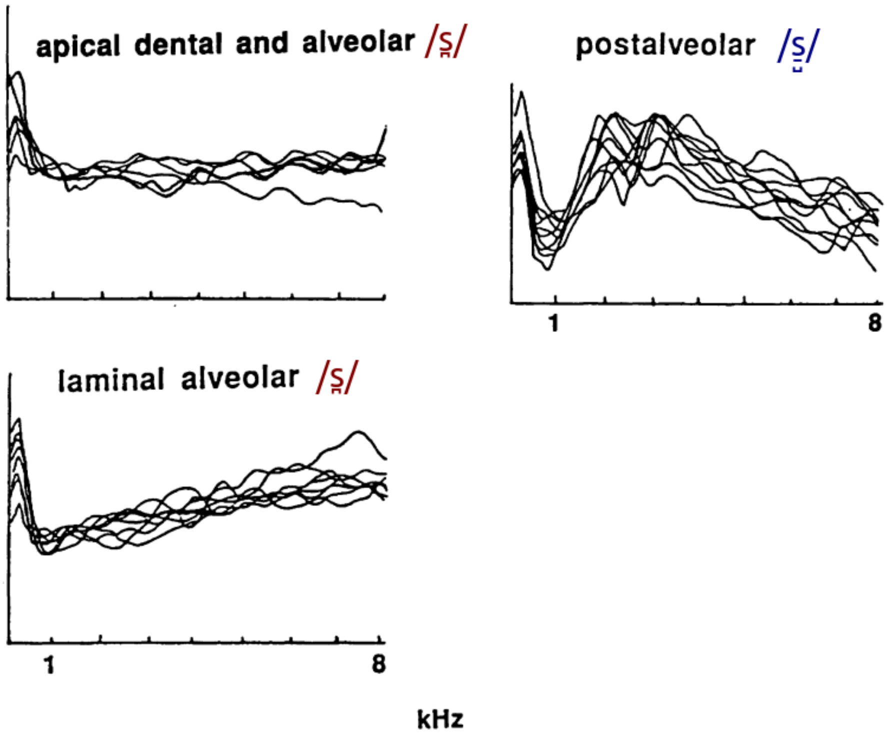


Figure XVII: *Averaged spectra of the two sibilant fricatives of O'odham, taken from Dart (1993: 36). Note that she labels the retracted alveolar sibilant as postalveolar, although its description corresponds to the retracted alveolar of this dissertation (Dart 1993: 31f.).*

ingly, the acoustic difference between apical and laminal shapes is less evident at the dental or dentialveolar place of articulation, as revealed by the spectra on the left (red) of fig. XVII. Whether a speaker produces the dental sibilant with an apical or laminal articulation will change little to the acoustic properties of the resulting sound, unlike for (retracted) alveolar and palatoalveolar sibilants. Dart (1991: 29), together with Catford (1977: 157), states that "the acoustic and aerodynamic differences between apical and laminal /s/ are more evident if they are alveolar[1] than if they are dental". In her study on the O'odham distributedness contrast, it turns out to be "unimportant whether a speaker normally makes an apical or laminal /s/, as long as the contrasting /ṣ/ is articulated farther back, and apically". This points towards an important aspect of distributedness contrasts: the configuration of the larger cavity and freer tongue movement possibilities around [s̪∼ṣ] leaves more slack for better distributedness distinctions in this range,

[1] Dart (1991) seems to use the term 'alveolar' both for alveolar and retracted alveolar.

whereas the very front dental sibilants ([s̪]) and very back subapical retracted retroflexes ([ʂ̠]) offer little room for these contrasts. Acoustic evidence for distributedness contrasts is thus more salient in the middle range of possible PoAs.

A type of sibilant that is typically absent from IPA charts and rare in the world's languages is whistled sibilants (Bladon et al. 1987; Shosted 2006, 2011; Lee-Kim et al. 2014). Whistled sibilants are universally known from non-linguistic recreative whistling, but are almost exclusively found as linguistic phonemes in southern Bantu languages, and written in the extended IPA with an upward pointing arrow ([s͎], [z͎]; Ball et al. 2017). While they have long been considered to be labialised in some way, Shosted (2006) showed that their articulation does not primarily rest on any particular lip protusion; in fact, the regulation of lip rounding can optionally serve to raise or lower the resonance frequency (i.e. to make the 'whistled note' more grave or acute). Whereas the primary whistle component in recreative *lip* whistling is made by rounding the lips and the tongue serves to regulate the tonality of the sound, in whistled sibilants the primary whistle component is the narrowing made by the tongue against the roof of the mouth and lip rounding serves to regulate tonality (i.e., the inverse mechanism of lip whistling). The reason for which whistled sibilants are addressed here is that spectral measures have indicated that they are in fact an extreme case of apicality (Shadle 1985: 165–177): "they are accompanied by a spectral peak narrower than that of non-whistled sibilants" (Shosted 2006: 2). This results in a significantly lower Standard Deviation for e.g. XiTsonga [s͎] wrt. [ʃ] (Lee-Kim et al. 2014: 70), and a narrower spectral peak bandwidth for Shona [s͎] wrt. [s] (Shosted 2006: 5). An illustration of the narrow peak in the spectrum of a whistled sibilant is given in fig. XVIII. The spectrum of this whistled

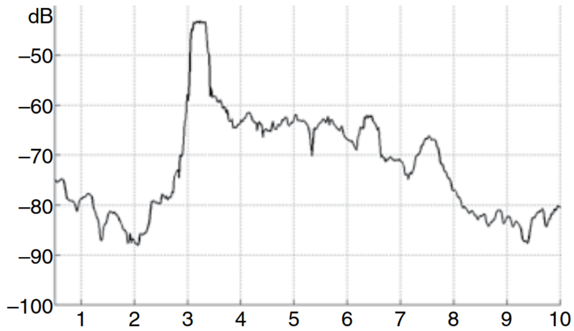


Figure XVIII: *Spectrum of the XiTsonga whistled [s͎], taken from Lee-Kim et al. (2014: 65).*

sibilant is very similar to that of recreative sibilant whistling as produced by the author of this dissertation in the right part of fig. XIX. It becomes apparent when inspecting the spectrogram of the whistled voiceless retracted alveolar sibilant that lip rounding and unrounding is used to lower and raise the peak iteratively. On the left side of the figure, a laminal retracted alveolar pronounced at approximatively the same place but with a flat tongue configuration is illustrated. It can be observed that its CoG is higher despite the similar PoA, a consequence of the larger sublingual cavity of apical sibilants. The acoustic correlates of distributedness are well represented by the second and fourth spectral moments in fig. XIX: the extremely apical sibilant has a Standard Deviation almost half that of the laminal one (639 vs. 1122 Hz), and its kurtosis is much larger.

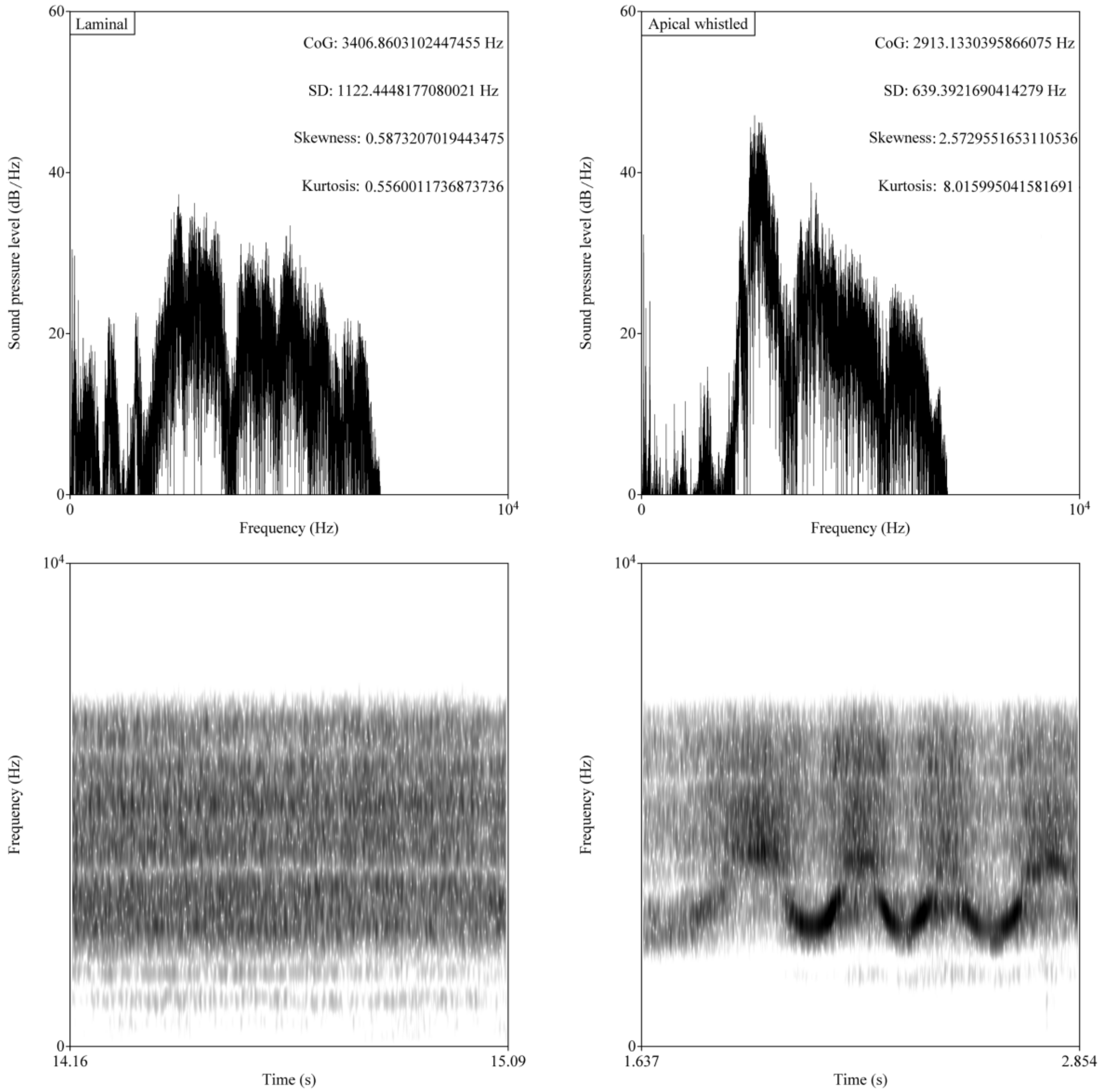


Figure XIX: *Comparison (averaged spectrum, spectrogram and spectral moments) of a laminal vs. apical whistled retracted alveolar sibilant. Note: weaker energies have been filtered out in Praat (Boersma and Weenink 2019), explaining the white zones.*

To summarise, the literature has shown that sibilants can be distinguished from non-sibilant fricatives and affricates based on a combination of parameters (Ladefoged and Maddieson 1996: 139): higher CoG and lower skewness, higher intensity and lower Standard Deviation/variance. Progress has thus been made since 1987, when Bladon et al. (1987: 49) were writing:

> Ladefoged (1971) found it phonetically useful – as indeed we do – to identify a class of sibilant (versus nonsibilant) fricatives, for reasons of how they pattern in language use. But, apart from noting that sibilant fricatives have a 'comparatively large amount of energy at high frequencies' (1971: 57), Ladefoged admitted to not knowing what actual acoustic measurements could be made that would correlate with the perceptual characteristics of sibilance.

Furthermore, the same parameters allow to distinguish sibilants within the class: anterior sibilants are characterised by a higher CoG, lower skewness and lower intensity; apical sibilants are distinguished by a lower Standard Deviation/variance and higher kurtosis. An approximate representation of sibilant and non-sibilant fricatives along the three continua is provided in fig. XX.

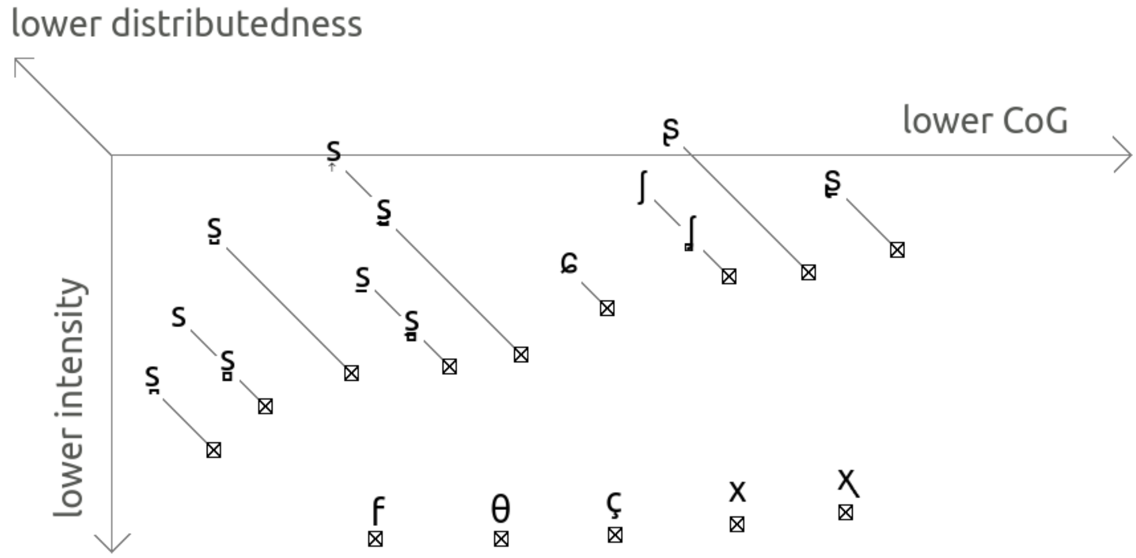


Figure XX: *Approximate acoustic map of sibilants and non-sibilants, along the dimensions of CoG, intensity and distributedness. Boxed crosses indicate the placement of the segments on the CoG and intensity continuum, while the grey bars indicate the height of the segments on the distributedness axis. Note that the posterior non-sibilants [x, χ] have a high SD but also a very high kurtosis, something which could not be represented simultaneously here.*

2.1.2.4 Formant transitions

An additional parameter useful in distinguishing alveopalatal [ɕ] or retroflex [ʂ] from other sibilants is formant transitions at the onset of a following vowel, in particular the second formant, F2 (Cristià et al. 2011: 392f.). In Polish for example, Bukmaier and Harrington (2016: 312) mention that formant transitions contribute to discerning [ɕ] and [ʂ]: the former makes the following vowel closer to [i] due to the (alveo-)palatal articulation, while the latter has an opposite effect of retracting and lowering the vowels (Nowak 2006: 140; Hamann 2003: 111-129). More precisely, the alveopalatal [ɕ] is consistently accompanied with rising F2 transitions (Nowak 2006: 144), as well as a low F1 (both of which characterise [i] and palatal segments in general; Bukmaier and Harrington 2016: 323). The F2 parameter also serves to distinguish the non-sibilant palatal [ç] from the posterior sibilant [ʃ] in standard German, where the former is palatal and therefore coincides with a high F2 (Wottawa et al. 2016). Since (denti-)alveolar sibilants are very resistant to coarticulation (Iskarous et al. 2011) and articulated further forwards than the frontmost vowel (i.e. outside of the vowel diagram), formant transition cues are relevant in particular for the range from retracted alveolar to more posterior sibilants.

2.1.2.5 Affricatedness

Articulatorily, it has been explained on page 9 that a sibilant affricate consists of the full occlusion characteristic of a stop followed by a sibilant fricative release. This is expectedly reflected in the acoustic properties of a sibilant affricate: a first phase of silence on the spectrogram corresponds to the moment in which the stop occlusion is made and pressure is building up behind the full closure; a burst corresponds to the moment in which the tongue moves from the full occlusion to letting the air escape rapidly through the small constriction; and the spectral shape of a sibilant fricative follows, which has already been detailed above. The three successive phases are illustrated in fig. XXI. Some languages, like e.g. Polish, distinguish sibilant affricates from sequences of a stop

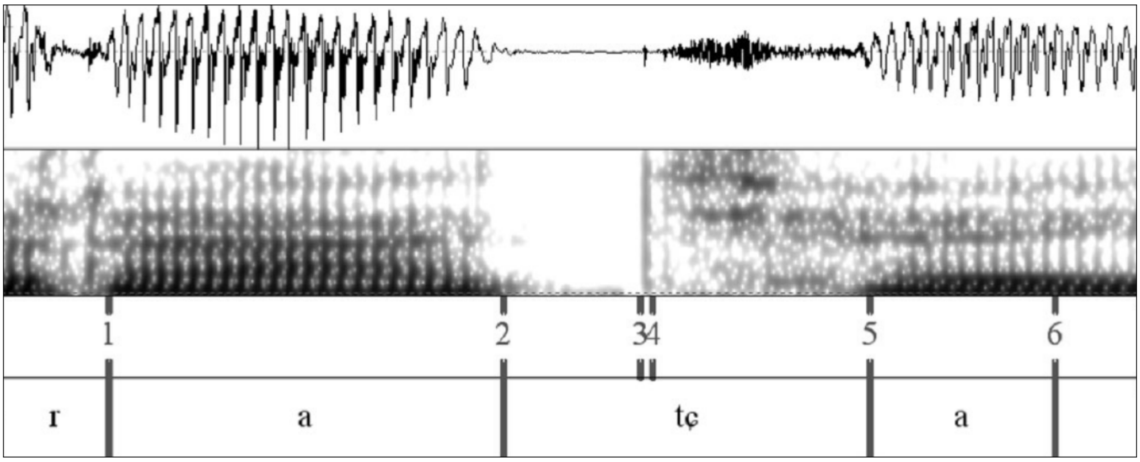


Figure XXI: *Spectrogram and oscillogram of the Polish alveopalatal affricate [t͡ɕ], adapted from Żygis, Pape and Jesus (2012: 284). The occlusion/silence phase spans from 2 to 3, the burst from 3 to 4 and the sibilant release from 4 to 5.*

followed by a sibilant fricative (Repp et al. 1978). For example, the word *czysta* ([ṭ͡ṣɨsta],[2] 'clean.FEM/vodka') with an affricate is distinguished from *trzysta* ([tṣɨsta], 'three hundred') with a stop and fricative. Fig. XXII displays the spectrograms and oscillograms of both words as pronounced by a female native speaker. The difference resides mainly in the temporal cue of duration: the sibilant fricative part in fig. XXII is four times as long in the stop + fricative sequence compared to the affricate. Sibilant fricatives are distinguished from sibilant affricates by being longer and having a slower rise in amplitude from the onset to the peak (Mitani et al. 2006; Berns 2013: 4), which means that *trzysta* will have the longest duration, as visible in fig. XXII. A sibilant affricate is thus not only characterised by the occlusion, burst and sibilant fricative phases described above, but also by the fact that the sibilant fricative phase is shorter than if it were phonologically an underlying fricative rather than an affricate.

[2] The subscript dot under the /ṭ/ in this particular case indicates the assimilation of its PoA to that of the following fricative, which is an indicator of being an affricate (Recasens and Espinosa 2007: 146). This place assimilation in affricates is rendered here but not elsewhere in this dissertation because the affricate vs. consonant sequence contrast is directly relevant here, unlike elsewhere. The spectrograms in fig. XXII illustrate the CoG-lowering transition from the alveolar [t] burst to the retroflex [ṣ] only in the case of the consonant sequence.

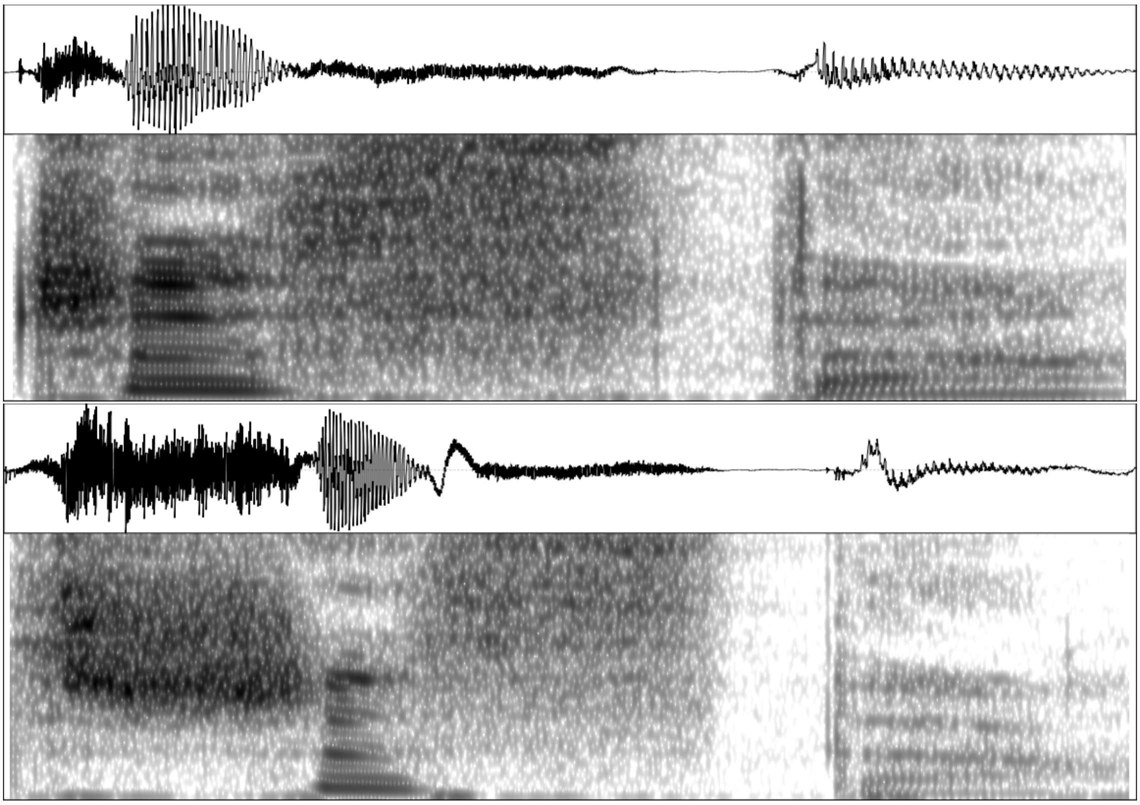


Figure XXII: *Spectrograms and oscillograms of the Polish words* czysta *(above) and* trzysta *(below), pronounced by a female native speaker (Wikimedia: Rovdyr 2007*a,b*).*

2.1.2.6 Voicing

Voiced sibilants are produced by vibrating the vocal cords, while maintaining the tongue configuration of their voiceless counterpart. This is reflected on a spectrogram as the

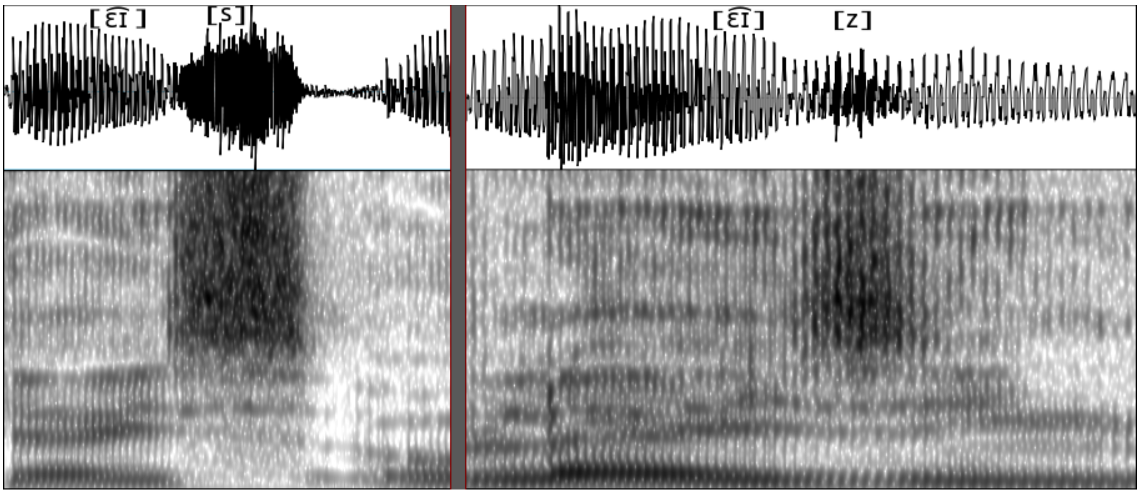


Figure XXIII: *Spectrograms and oscillograms of [ɛ͡ɪs] vs. [ɛ͡ɪz] in the nonce Flemish words* Geyshoorn *and* Vlijzingen*, pronounced and recorded by the author of this dissertation in Praat (Boersma and Weenink 2019).*

shape of the voiceless counterpart, overlayered with dark periodic bars in the lowest frequencies due to vocal cord vibration, as illustrated in fig. XXIII. The figure shows how the vocalic formants pervade through the intervocalic [z], in particular the formant bars in the lowest frequencies. Voicing is also visible as periodicity, rather than random noise, on the oscillogram. Because of the voicing energy at the bottom of the spectrogram, the Centre of Gravity of voiced sibilants is lower than that of their voiceless counterparts.

2.2 Sibilants in phonology

The citation on page 29 already evidenced that sibilants are to be considered a class because "of how they pattern in language use" (Bladon et al. 1987: 49). Reasons to believe that sibilants pattern as an independent natural class in phonology (Gussenhoven and Jacobs 2017: 65) are multiple, out of which three are provided here.

A classical example revealing how sibilants pattern as a class is the English *-s* suffixation in plurals, possessives and 3rd person verbs (cfr. among others Yip 1988: 86; Crystal 2008: 44, 337; Hall and Żygis 2010: 11): if a noun ends in a sibilant, it takes the allomorph -[ɪz] rather than -[z], as shown in (7).

(7)

a love -[v] → two lo[vz]
I bathe -[ð] → she ba[ðz]
a bus -[s] → two bu[sɪz] (*buss)
a match -[t͡ʃ] → two ma[t͡ʃɪz] (*matchs)

A similar rule exists in the southern Bavarian *Mòcheno* dialect, where the neuter marker *-s* surfaces as *-es* after sibilants such as /s̺, s̩, t͡s, ʃ/ (though apparently also after /x/; Rowley 2017: 127). Catalan also has a similar plural rule, whereby *amic* 'friend' becomes *amics* 'friends' but *cas* 'case' becomes *casos* 'cases' rather than **cass*. The epenthesis with [o/u] (or rather, the preservation of historical vowels) clearly serves to avoid adjacent *sibilants*, since the consonants that trigger epenthesis are specifically /s̺/, /ʃ/, /t͡ʃ/ etc. (Dols and Mansell 2017: 23). All these examples show that the Obligatory Contour Principle (Leben 1973; McCarthy 1981), which prohibits adjacent identical elements, 'sees' sibilants as one class by prohibiting their adjacency irrespectively of differences in manner (fricative vs. affricate) and/or PoA (anterior vs. posterior). This explains why not only words like *pas+s* are illicit in languages that satisfy the OCP with epenthesis, but also words like *paʃ+s* or *pat͡ʃ+s*.

Other evidence for sibilants as a natural class is the fact that several languages exhibit sibilant harmony patterns (Beeler 1970; Yip 1988; Hansson 2001*a*; Walker et al. 2006; McCarthy 2007; Oberly 2008; Hansson 2010; Kosa 2010; Berkson 2013; Zellou 2013; Bennett 2015). Sibilant harmony is described in detail in section B.12 of chapter 3. In e.g. Chumash, Kinyarwanda or Navajo, a sibilant agrees in place of articulation with a sibilant (much) further in the word despite its original PoA specification. For example, a Chumash word like *kiškín* 'I save it', when suffixed with *-us* 'for him', becomes *kiskínus* (i.e.

[+anterior] is spread leftwards to the /š/; Yip 1988: 79). Such a process targets sibilants specifically as a class, irrespectively of manner (fricative vs. affricate). This behaviour of sibilants as a class is evidenced by the fact that in *kiskínus*, the /k/, /i/, /n/ and /u/ in between the two sibilants are simply unaffected and totally 'transparent' with respect to sibilant harmony. Any sibilant in the word, however (e.g. /t͡ʃ/), will be targeted by sibilant harmony so that e.g. the hypothetical word *tiškítš + us* should become *tiskitsus*.

Additional evidence comes from tongue twisters, which aim to be difficult to pronounce by accumulating slightly different segments from an identical phonological class. English [ʃisɛlzsiʃɛlzba͡ɪðəsiʃɔː] (*She sells seashells by the seashore*) or Polish [vt͡s̪aɕes̪us̪ɨs̪os̪as̪uxa] (*W czasie suszy szosa sucha*) are made difficult by their recurrent changes in voicing (English example), manner (Polish example) and PoA (both English and Polish), alternations which always occur within the class of sibilants. A good (i.e. difficult) tongue twister does not accumulate precisely identical segments (e.g. French [sisasəsesasəsoʀa] *Si ça se sait ça se saura*) nor radically different segments (e.g. French [vwasiynfʀaːzpakɔ̃plikeːadiːʀ] *Voici une phrase pas compliquée à dire*); it takes one phonological class (such as sibilants) and alternates between its elements (e.g. French [t͡ʃɛksisisɑ̃t͡ʃɛkʃɛʀʃsaʃasɑ̃sɛʃosyːʀ] *Checke si six cents tchèques cherchent Sacha sans ses chaussures*). The similarity between the different sibilants tricks the pronouncer into assimilating them to identical realisations (e.g. everywhere /s/ or /ʃ/), i.e. they are hard to keep apart from each other because of their common phonetic characteristics (sibilance) and at the same time, they *shall* be kept apart from each other because they contrast phonemically in the relevant language.

Evidence for sibilant *fricatives* patterning as a natural class is provided by Polish *e*-insertion after the preposition *z* ('with'), which occurs before all Polish sibilant fricatives (/s̪, z̪, ɕ, ʑ, ṣ, ẓ/) followed by a consonant. This means that e.g. *z + sztuka* 'play, piece' is realised as *ze sztuką* rather than **z sztuką* (Hall 1997*b*: 13). However, the same does not apply to sibilant affricates, so that *z + cztery* 'four' is realised as *z czterema*.

2.2.1 The representation and categorisation of sibilants

Sibilants are universally classified as the following: *consonants* (as opposed to (semi-)vowels), *obstruents* (as opposed to sonorants), *fricatives* or *affricates* (as opposed to stops; Hall and Żygis 2010: 1), *coronals* (as opposed to labials, dorsals, radicals and laryngeals) and *strident* (as opposed to non-strident segments). This means that in any language, sibilants will behave phonologically as a consonant, obstruent etc. regardless of the language-specific phonemic contrasts. No language has thus been reported to have sibilants behaving as e.g. vowels, stops or laryngeals.

In the literature, sibilants are represented as strident coronals, i.e. [coronal, +strident] in distinctive features (Yip 1988: 79f.; Hall 1997*b*; Poser 2004: 2; McCarthy 2007). They are not the only strident segments (also e.g. /f/ or /χ/ are strident), but they are the only stridents that are coronal at the same time (Hall and Żygis 2010: 9f.). Within the class of sibilants, single segments are represented with the specification for sibilants ([coronal, +strident]), plus additional contrast-dependent specifications:

- [anterior]: anterior vs. posterior (/s̪, s, s̠/ vs. /ʃ, ṣ, s̪̠/)
- [distributed]: apical vs. laminal (/s̺, ṣ/ vs. /s̻, ʃ̻/)
- [continuant]: fricative vs. affricate (/s, ʃ/ vs. /t͡s, t͡ʃ/)
- [voice]: voiceless vs. voiced (/s, ʃ/ vs. /z, ʒ/)
- secondary articulations not addressed in this dissertation (e.g. /sʰ, sʲ, sʷ, s̃, s', sː/)

This means that for example, a /z/ is minimally specified as [coronal, +strident, +anterior, +voice], a /ṣ/ as [coronal, +strident, -anterior, -distributed]. These additional specifications are determined by what is contrastive in the sibilant inventory: for example, a language with only 1 PoA for sibilants, affricates and no voicing distinction will have a /s̠/ ([coronal, +strident, +continuant]) and a /t͡s̠/ ([coronal, +strident, -continuant]) but no specification for [anterior], [distributed] or [voice]. A detailed typology of sibilants with distinctive features can be found in 5.5.

Most of the categorisations of sibilants reported in the list above (e.g. that /ṣ/ is apical, that /s/ is anterior etc.) are undisputed in the literature. However, some sibilants are less straightforwardly categorised, being situated at the border between two specifications. This is the case of the alveopalatal sibilants /ɕ, ʑ, t͡ɕ, d͡ʑ/. In what follows, I argue for their behaviour as anterior rather than posterior segments.

Alveopalatal sibilants are considered to behave as [-anterior] in a majority of analyses (e.g. Hall 1997*b*: 43, 46; de Lacy 2006: 178). At the same time, going past the [±anterior] dichotomy, it is also uncontroversial that alveopalatals are perceived as more anterior than any other posterior segment such as /ʃ, ṣ/ (Hall 1997*b*: 51; Padgett and Żygis 2007: 160), and more posterior than any alveolar segment (Crystal 2008: 22). Alveopalatals are thus situated right in between [s] and [ʃ] on the front - back axis, behind retracted alveolars such as [s̠] (Seinhorst and Ooijevaar 2011; Pétursson 1971: 210; recall also table 1). They "sound 'high pitched', as much hissing as hushing", very much like retracted alveolars (Padgett and Żygis 2007: 160). Furthermore, it will be shown in 3.2.2.2 and 3.3.1.2 that natural languages with /s̠/ and /ɕ/ distinguish them only very little phonetically (Seinhorst and Ooijevaar 2011; Ooijevaar 2011) or phonologically, something which shows how close they are on the continuum [s - s̠ - ɕ - ʃ]. The literature thus sets the 'border' between [+anterior] and [-anterior] in between retracted alveolar and alveopalatal, and recognises that they are neighbouring realisations. Nevertheless, I will show below that retracted alveolar and alveopalatal in fact belong together, and propose that the 'border' is rather to be set between alveopalatal and palatoalveolar.

1. A preliminary observation is that the distinction between [+anterior] and [-anterior] itself is controversial. Although it is a straightforward distinction for the /s - ʃ/ contrast, it has been much less clearly and more arbitrarily applied to the border cases /s̠ - ɕ/ (compare e.g. Adams 1975, who considers even retracted alveolars to be [-anterior]). The controversiality of [anterior] has been noted by Trask (2004: 25):

> In the ***SPE*** **feature system**, [JK: [±anterior] is] a **distinctive feature** defined as 'produced with an obstruction located in front of the palato-alveolar re-

> gion'. Thus, labial, dental and alveolar consonants are [+anterior], while palato-alveolars, palatals, velars, uvulars and glottals are [-anterior]. Vowels are assumed to be universally [-ant]. Though arguably derived from the **diffuse/compact** contrast of the **Jakobson-Halle feature system**, this is the most unsatisfactory and controversial of all the *SPE* features, since neither [+ant] nor [-ant] segments ever appear to behave as a natural class, and the definition appears to be little more than an *ad hoc* device for distinguishing alveolars from palato-alveolars.

Even if many categorise alveopalatals as [-anterior], there is thus little evidence for its behaviour as such rather than e.g. as [+anterior, +high].

2. In languages with a single sibilant fricative /s̠/, stop palatalisation in /tj/ yields an alveopalatal /t͡ɕ/ which has been observed to depalatalise to /t͡s̠/. This happened e.g. in Ancient Greek (Teodorsson 1979; Vijūnas 2010: 41), Latin (Pope 1934: 129) and Late Proto-Finnic (Pajusalu 2012; Koivulehto 1986: 293). If the alveopalatal affricate necessarily behaved as a posterior segment, its deaffrication would be likely to create a posterior segment contrasting with /s̠/, thus creating an anterior - posterior contrast. Cross-linguistic observations in 3.3.1.2 and A.1 reveal that alveopalatals do not in fact behave as posterior segments if there is one PoA (retracted alveolar) but rather assimilate to that PoA, and that in languages with an anterior - posterior contrast, they are *variably* reanalysed as anterior or posterior sibilants (whereas they would always remain posterior if they were inherently posterior).

3. Another cross-linguistic argument for grouping alveopalatals with retracted alveolars is that in the two attested types of 3s inventory (the 'Polish' type, /s̪ - ɕ - ṣ/ and the 'Basque' type, /s̪ - s̠ - ʃ/), they both function as the middle sibilant (see 3.2.2.4). There is no reason to assume that Polish has two posterior sibilants and Basque only one, when the alveopalatals and retracted alveolars yield the same contrast from the point of view of auditory dispersion.

4. Children acquire anterior sibilants earlier than posterior sibilants (see 4.4). Alveopalatals are acquired earliest by children in Japanese (Li et al. 2009: 121), Mandarin Chinese (Li and Munson 2016) as well as Polish (Żygis et al. 2019). If the alveopalatals were [-anterior], they would be acquired latest rather than earliest.

5. Although in phonetic terms, there is no clear articulatory border between anterior and posterior sibilants, but a continuum instead (see 2.3.1), Lindblad and Lundqvist (1994) show that Swedish /ɕ/ is an *anterior* segment and only rarely fully palatal, in their own terms.

To summarise, alveopalatals do not seem to pattern as purely posterior sibilants, but do pattern with retracted alveolars instead. In this dissertation, they will be grouped right behind retracted alveolars on the PoA continuum (see table 2).

2.2.2 The representation of sibilant affricates

Phonetically, as shown in 2.1.2.5, sibilant affricates consist of a stop followed by a sibilant fricative release. At the same time, several languages (like Polish with the *czysta*

- *trzysta* example of 2.1.2.5) reveal that sibilant affricates can contrast phonologically with stop + fricative combinations. Although the articulation and acoustics of affricates are thus well described, their underlying representation in phonology is the object of much debate. Berns (2013: 17-47) summarises the different hypotheses with respect to the representation of affricates as follows, starting from Jakobson et al. (1952) up to the most recent literature.

- Jakobson et al. (1952) describe affricates as **strident stops**, i.e. [-continuant, +strident] where stops are [-continuant, -strident] and sibilant fricatives are [+continuant, +strident].
- Chomsky and Halle (1968) describe affricates as **stops with a delayed release**, i.e. [-continuant, +delayed release] where stops are [-continuant, -delayed release], sibilant fricatives are [+continuant, +strident] and affricates further distinguish between [+strident] (e.g. /t͡ɕ/) and [-strident] (e.g. /p͡f/).
- Hoard (1975) describes affricates as **[-continuant] and [+continuant] segments** at the same time, to account for the way they pattern with both stops and fricatives.
- Building further on this combination of [±continuant], Sagey (1986) describes affricates as **linearily ordered** [-continuant] then [+continuant] segments, exhibiting stop-like behaviour at their left edge and fricative-like behaviour at their right edge (Hall and Żygis 2010: 26).
- Building further on this combination of [±continuant], Hualde (1988), Lombardi (1990) and van de Weijer (1996) argue against this linear ordering and in favour of **linearly unordered** [+continuant] and [-continuant].
- Building further on this combination of [±continuant], Steriade (1994) describes affricates as segments with an **intermediate aperture**, where stops have a minimal and fricatives a maximal aperture.
- Schafer (1995) describes affricates as stops with a **secondary [+continuant] stricture**, where stops have a primary [stop] stricture and fricatives a primary [+continuant] stricture.
- Shaw (1991) describes affricates as **underlying stops** which get additional features based on language-particular contrasts.
- Rubach (1994) and Clements (1999) describe affricates as **strident stops**, moving back towards Jakobson et al. (1952).
- LaCharité (1993) and Kehrein (2002) describe affricates as **underlying stops** indistinct from other stops, implying that affricates do not exist as a distinct natural class.

It is not necessary here to fully take position in favour of one analysis in particular, since several of their implications lie outside of the scope of this dissertation (e.g. the status of non-coronal affricates). Nevertheless, the sibilant typology I propose incorporates the principle of affricate markedness (i.e. that affricates are more marked than stops and fricatives, cfr. Fletcher 1989: 746), which requires affricates to constitute a natural class and thus means that the analyses seeing affricates as simple stops (LaCharité 1993; Kehrein 2002) are incompatible with it. Sibilant affricates pattern with sibilant fricatives because of their [+strident] specification, with stops because of their stop specification (e.g. in terms of sonority, cfr. Parker 2002: 262), and constitute a natural class distinct from stop + fricative combinations, stops and fricatives. A phonological rule can target both sibilant fricatives and affricates alike, as in English *-s* suffixation, as much as it can target only sibilant fricatives, like in Polish *z* + noun constructions, or only sibilant affricates, as in deaffrication (see B.9 in chapter 3) or stop intrusion (see B.7 in chapter 3). The strident stop approach and the [-continuant][+continuant] approach are thus both compatible with the claims of this dissertation, since they allow the [+strident] feature of sibilant affricates to interact with that of sibilant fricatives and they allow for the expression of affricate markedness as *[-continuant, +strident] or *[-continuant, +continuant] respectively.

2.2.3 On the sonority hierarchy

Sibilants (in particular sibilant fricatives, cfr. Henke et al. 2012: 70f.) have been observed to occupy a particular place on the sonority hierarchy (Parker 2002; Jany et al. 2007). The sonority hierarchy, which describes what segments are more likely to constitute a syllable nucleus or be nearer the nucleus (= more sonorous) and what segments are rather likely to be in the onset or coda (= less sonorous), is illustrated in (8) (based on Parker 2002: 262 and Henke et al. 2012: 71). Note that some analyses such as e.g. Clements (1990) merge obstruents altogether on the hierarchy rather than distinguishing between fricatives and stops/affricates.

(8)

(more sonorous > less sonorous)

vowels > semivowels > liquids > nasals > voiced fricatives > voiceless fricatives > voiced stops and affricates > voiceless stops and affricates

The sonority hierarchy has been correlated with intensity in a large portion of the literature (Clements 2006: 2), whereby more sonorous elements are louder. However, on a scale from loudest to least loud, sibilants then rather outrank the other obstruents (Baroni 2014*a*: 95), and /ʃ/ even outranks the nasals in e.g. German (Wolf 1871: 71) and English (Clements 2006: 2). There is thus a discrepancy in the sense that sibilants, “though rather noisy sounds, do not pattern as high-sonority sounds in most languages” (Clements 2006: 2). In fact, although they pattern as less sonorous than nasals, their ranking with respect to the other obstruents is not universally fixed, varying flexibly across and within languages. One can thus encounter onsets like e.g. /sk/ in French *scaphandre* or /sp/ in Italian *Spagna* ‘Spain’, as much as the reverse onset /ks/ in French *xylophone* or /ps/ in Italian *psicologia* ‘psychology’. This is remarkable to the extent that both French and Italian otherwise obey the sonority hierarchy (Maïonchi-Pino et al. 2014; Davis 1990), with sibilants constituting the only exception to the rule. An example with another fricative could be Albanian /sf/ in *sfera* 'sphere' vs. /fʃ/ in *fshat* 'village'; one with an affricate could be [t͡svant͡sk] 'twenty' in Swiss German, with the sibilant preceding an obstruent both in onset and coda. Even within one and the same onset or coda, sibilants can defy the sonority sequencing principle by occurring both before and after a stop (e.g. Tyrolean [pʃt]*elln* 'to order', Meneguzzo 2017: 98; English *gra*[sps]), inversely between two stops (e.g. English *mi*[tst] 'midst') or before and after a fricative (e.g. careful English pronunciation *si*[ksθs] 'sixths'). They thus benefit from a certain freedom of movement on the hierarchy, whereas other segments remain fixed.

This particularity of sibilants wrt. sonority can be explained by their perceptual robustness: sibilants can be distinguished robustly enough without the help of vocalic formants (Harris 1958; Jongman et al. 2003; Wright 2004: 51f.; Goad 2012; Li et al. 2012), while other segments rely on acoustic information contained in formant transitions.

Several analyses posit furthermore that sibilants can be extrasyllabic, i.e. that they are not parsed into the syllable but can be freely appended in front of or behind the syllable. This has been posited among others for German (Wiese 1986: 12ff.) or English (although refuted by Hall 2002), Sanskrit (Kobayashi 2004: 41f.) and Latin (Kostakis 2017). This constitutes further evidence for the fact that sibilants, and sibilant fricatives in particular, behave more flexibly and independently than other segments with respect to their placement in or next to the syllable.

2.3 Learnability experiment: overcrowded inventories

This section addresses the problem that there is an infinity of possible phonetic realisations for sibilants on the place of articulation continuum (e.g. 3756.0290 Hertz, 3756.0291 Hertz etc.), and at the same time, that natural languages restrict these possibilities to a maximum of 4 categories. A learnability experiment quantifies the ability of individuals to distinguish 3, 5 and 7 places of articulation for sibilants, shedding light on the physical limitations within a population that cause the natural languages to be as they are.

2.3.1 Finite categories on an infinite continuum

Human beings are able to produce sounds on a finite range of values (e.g. sibilants with a CoG ranging from X to Y Hertz; for each individual, X and Y might be lower or higher in function of physical characteristics such as one's vocal tract dimensions). However, such a range can in principle be split into an infinite number of phonemic categories (Goswami 2012: 2625): a range from 0% to 100% can be split in 4 parts of 25% width each, but it can even well be split in 537 parts of 0.186219739% width each, and so forth. Hall (1997*b*: 86) observes about phonemes and acoustic space that "[g]eneralizations pertaining to the maximum number of entities that can 'fit' into any given space are traditionally ignored in segmental phonology". Nevertheless, the question of 'how many categories can fit' is not trivial, in the sense that the literature recognises the mismatch between the continuous (phonetics) and the discrete (phonology). For example, the phonetic continuum of sibilant CoG, ranging from high to low CoG, is illustrated with 8 categories in Boersma and Hamann (2008: 229):

(9) s̠̪ s sʲ s̠ ɕ ʃ ṣ ʂ

Clements (2004: 4) in turn mentions 7 phonetic distinctions on the same coronal continuum. At the same time, both Boersma and Hamann and Clements recognise that even if it is possible to make (at least) these many distinctions from a *phonetic* point of view, natural languages never make use of as many *phonemic* distinctions: instead, they limit the number of contrasting PoAs for sibilants to 4. This upper limit is also stated as a universal by Hall (1997*b*: 93), who says that "no language can have more than four places of articulation among sibilants of any given series". Boersma (1997*b*: 68) furthermore declares about the mismatch between finite categories and infinite phonetic values:

> Communicatively, the notion of an originally continuous vowel-height [JK: or PoA] feature is not problematic at all: because of innate capacities of human perception, the learner will divide it into a finite number of categories. This number is language-specific; the fact that many features are binary is caused by nothing more than the inability of the listener to distinguish faithfully more than two values of those perceptual features.

Boersma is right in saying that learners divide a continuum in a finite number of categories because of innate capacities, even if it must be observed that these innate capacities, together with experience-influenced capacities (e.g. whether one perfected or damaged one's hearing by attending piano lessons or loud rave parties), are different for each individual. As the experiment below will show, there is a wide variation in the ability to distinguish sibilants even for speakers of the same mother tongue: whereas some are able to reliably discriminate between 7 categories, others struggle with 3 categories. This brings us to the following doubt:

(10) Does the fact that maximum 4 PoAs for sibilants are attested in natural languages mean that languages with 5 PoAs are impossible, or is it just that by chance, no language with 5 PoAs exists despite being possible?

An important fuel for this doubt is the fact that, as will be shown in 3.2.1.1, the more PoAs a sibilant inventory contrasts, the rarer such an inventory is. If a system with 5 PoAs is possible, it is thus expectedly very rare and likely to occur perhaps once in every, let's say, 12000 languages. If 5 PoAs is possible, it is thus very much possible that it is unattested simply because of its expected rarity.

The following subsections describe an experiment that aims at providing a necessarily imperfect yet insightful answer to the doubt formulated in (10). The answer has to be imperfect in the sense that one cannot obtain a categorical answer like "4 is perfectly fine, but 5 is completely impossible"; the average ability to distinguish PoAs rather decreases gradually as the number of PoAs increases. Nevertheless, the experiment allows to quantify the average perceptual acuity of a population sample, thus estimating reliably how empirically sound it is to assume that 4 PoAs is the maximum contrastable.

2.3.2 Methods

The experiment, which relied on a trip to Poland in 2020 to gather data, has only been performed by 19 participants up to now because of the COVID crisis impacting the possibility for participants to all perform one task on the same computer with the same headphones. The results reported are thus only temporary.

2.3.2.1 Structure

The experiment is split in two parts, one of which appeals to absolute, non-linguistic perceptual acuity and the other of which appeals to linguistic categorical perception. The first part is named A (phonetic perception) and the second B (phonemic perception). Part A aims to quantify how well participants perceive differences in CoG between two sibilants in the form of a 'same-or-different' (i.e. AX) task, and part B aims to quantify how accurately participants distinguish 3, 5 and then 7 PoAs for sibilants phonemically, inside nonce minimal pairs of the shape /aSa/. All participants either have Brussels French or standard Polish as a native language: this allows to test whether native speakers of a sibilant inventory with 3 PoAs like Polish perform better than those of a sibilant inventory with 2 PoAs like French.

The entire experiment was scripted in Praat (Boersma and Weenink 2019). It was run on one and the same computer, with the same mouse and the same mousepad, so that there would not be any external influence on the reaction times of the participants (e.g. a slower computer or a less easily usable mouse for some participants). All participants used the same headphones (Audio Technica ATH-SR5BT) throughout the experiment. The experiment window was maximised in fullscreen for all participants. The researcher launched the first frame of the experiment each time, and then allowed every participant to choose the luminosity of the screen. The native speaker of Polish who was perfectly fluent in French was told to do the experiment in Polish. All the text written in French was written by the researcher (native speaker), and all the text written in Polish was checked and corrected by a native speaker of Polish with university background. An overview of the structure of the experiment (also shown to participants in French or Polish shortly after starting the experiment) is provided in fig. XXIV. After the experiment is launched, participants read the experiment title "The Mysterious Indigenous Language" (in French or Polish), with jungle-related pictures taken from the Marsupilami comics series by André Franquin and calm jungle noise in the background. After clicking anywhere to reach the next frame, a character named David Aliverock presents himself saying that he discovered a tribe speaking a strange language, and he asks for the participant's help in learning 12 words of the language. David Aliverock declares that before the participant can learn the words, he must test whether the participant hears well enough. The participant is then led to a frame that looks as in fig. XXV. In that frame, the participant must click on the "quiet", "average" and "loud" buttons to hear the same recording of a Sandawe speaker uttering one and the same sentence, with intensities scaled in Praat at 58, 64 and 70 dB respectively. The participant is told that the "quiet" sound must be comfortably perceptible, that the "loud" sound must not be

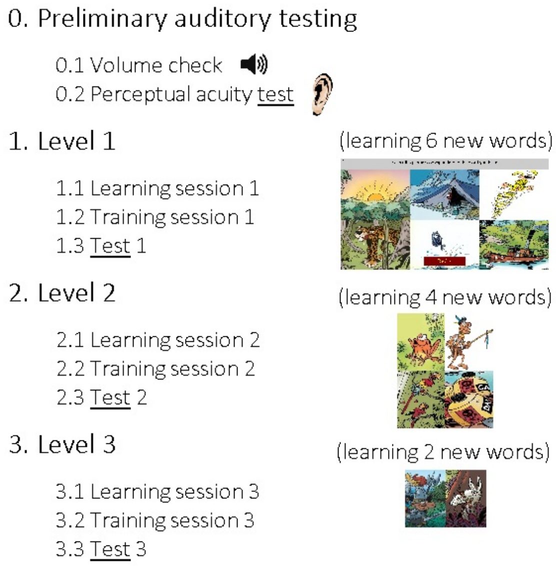


Figure XXIV: *Structure of the experiment.*

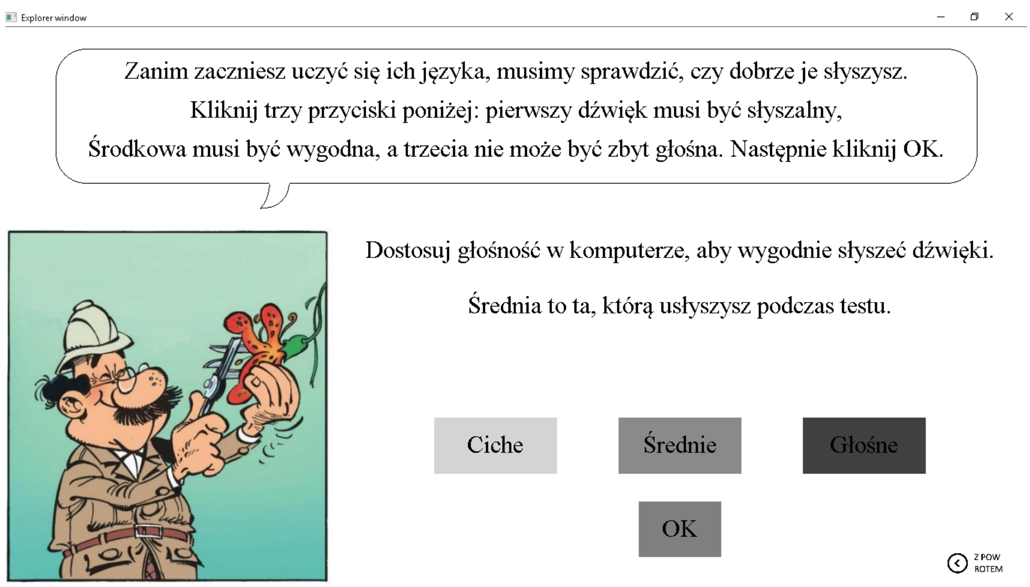


Figure XXV: *Experiment frame asking the user to regulate the volume of the computer.*

uncomfortable, and that after regulating the volume on the computer with the help of the researcher, (s)he will hear all sounds at the "average" level chosen (i.e. all sound files will be normalised to 64 dB in Praat before being heard). The purpose of volume regulation is to attenuate the person-specific effect of auditory sensibility: participants who perceive sounds at lower intensity thresholds than others are not advantaged, since the volume is calibrated to what each participant hears best.

2.3.2.1.1 Part A: phonetic perception

Participants are told by David Aliverock that they will hear two sounds successively and must determine whether they were identical or different. He insists on the fact that the sounds are either *perfectly* identical or different, so that they must be considered different even if there is the slightest acoustic difference between them. The tokens (see 2.3.2.2 below for a precise description of their acoustic properties) consist of identical or different pairs of 7 voiceless sibilant fricatives, equally spaced in ERB (1.275 ERB) along the Centre of Gravity continuum from [ʂ] to [s̪], embedded between two [a]'s produced by the author of this dissertation. This means that all sounds sound like /aSa/ and consist of always the same [a]-parts, with the sibilant CoG as only variable (as shown in fig. XXVIII). The formants are preserved as they are in the original audio parts, without any modification of formant transitions in the concatenation of the parts. The sibilant with the highest CoG (5959 Hertz) is a natural sibilant recorded by the author of this dissertation, while the other sibilants were modified with a formula in Praat to shift their CoG down to the minimum of 2287 Hz. The formula is $self = self + (x - self)$, where x is the CoG in Hertz that needs to be reached. Part A was divided into 6 levels: the first level confronts all sibilants that were spaced precisely 6/6[ths] away from each other on the continuum (that is, only the one with 5959 Hz and the one with 2287 Hz) and these same sibilants with themselves; level 2 confronts the ones spaced 5/6[ths] away from each other and these sibilants with themselves, and so further. This means that the auditory promiscuity between the different pairs i.e. the difficulty of the task increases at each level. The different pairs for each level are illustrated in fig. XXVI. It can be seen in

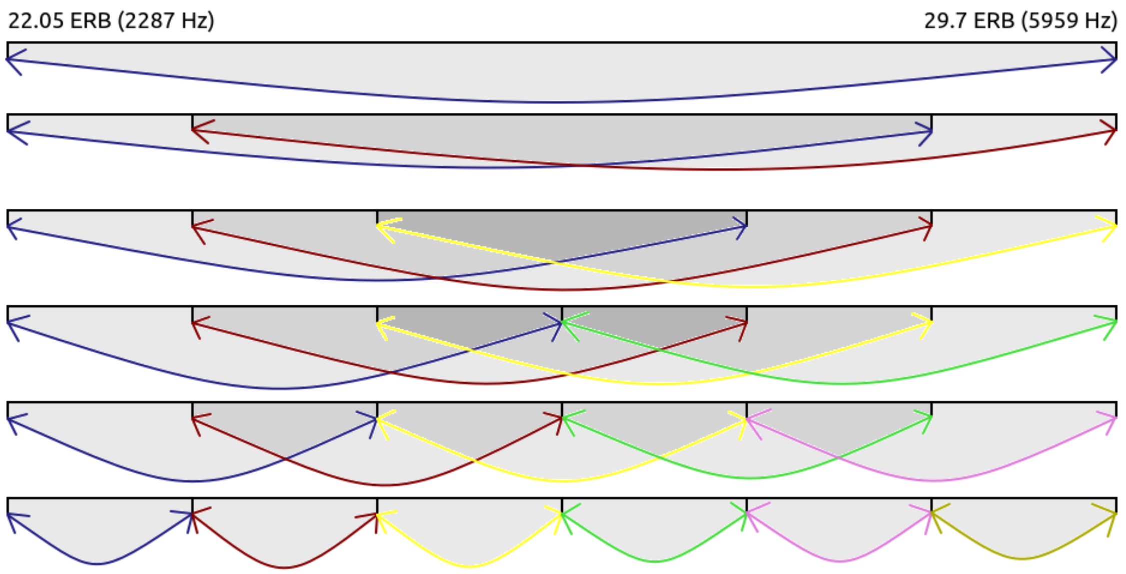


Figure XXVI: *Auditory space between the sibilants for each of the 6 levels, from 6/6 to 1/6.*

fig. XXVI that the sixth level basically corresponds to discriminating between 7 evenly spaced sibilants: either the token is an identical pair, or it is a different pair consisting of a sibilant and its nearest neighbour on the scale (up- or downwards). Each possible different pair is tested once in each direction (e.g. 2287 followed by 2666 Hz, as well as 2666 followed by 2287 Hz in the last level). To avoid that some identical pairs (namely the ones at the extrema of the continuum) would be tested more often than others, levels 3 and 5 do not test identical pairs at 5959 and 2287 Hz, level 4 does not test identical pairs at 2666 and 5030 Hz, and level 5 tests the identical pair at 3640 Hz twice. This ensures that every identical pair is tested precisely 4 times across the entire experiment, and that each CoG value is heard precisely 20 times across the entire experiment. No participant can thus be favoured by the frequency of occurrence of certain tokens if (s)he perceives tokens at certain frequencies better, because each frequency from 2287 to 5959 Hz is equally represented in terms of frequency of occurrence.

The number of tokens to evaluate increases at each level, as illustrated in fig. XXVII. Crucially, fig. XXVII shows that the numerical proportion between identical and differ-

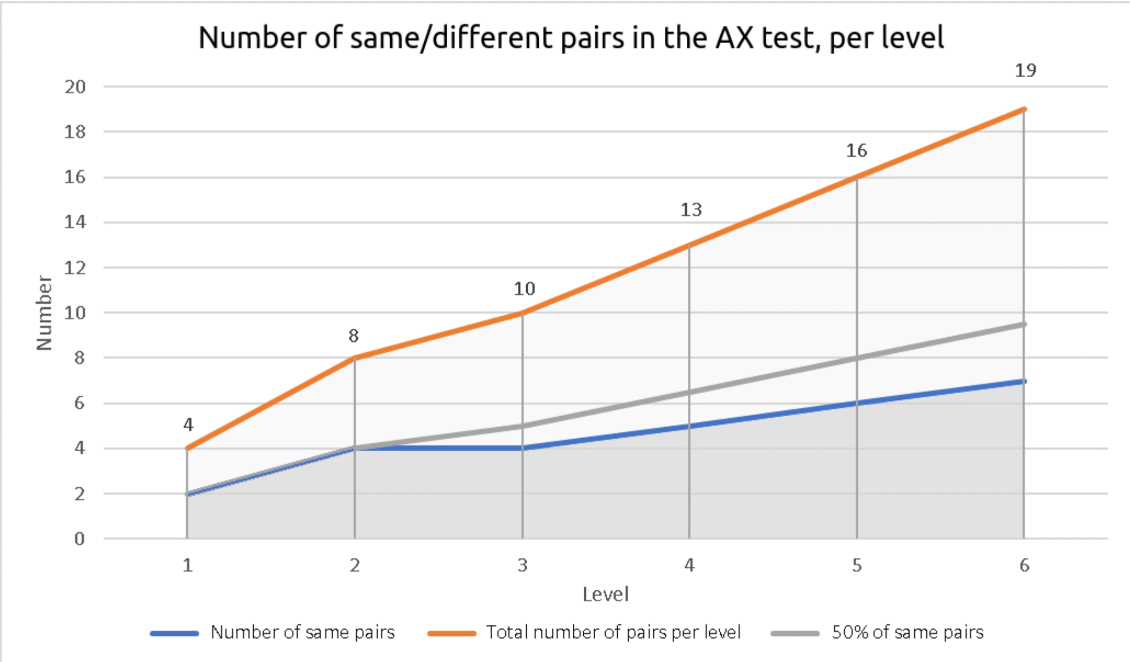


Figure XXVII: *Evolution of the number of tokens in the AX test.*

ent pairs tested starts at 50%, and slowly moves away from it to reach 36.8% identical pairs at the 6^{th} level. Participants in AX experiments typically have the expectation that identical pairs should occur about half of the time, and it is important that this impression be maintained throughout the experiment. This allows to avoid that some participants then notice and start clicking all the time on "identical" or all the time on "different", while others do not even notice or react by still clicking around 50% of the time on each because they are convinced that it *should* be 50% even if it does not seem to be the case. Having the number of identical pairs move slowly and discretely away from 50% allows to make sure that the illusion of the 50% is maintained, since participants who must concentrate on very subtle acoustic differences in level 6 are unlikely to notice that they should have clicked more frequently on "different" (12 times out of 19).

For both this part and part B, the tokens were randomised by Praat. Each pair is however unique, so that no effect of token order is to be expected. Each stimulus was preceded by 0.24 seconds of silence, had a short silence interval between A and X of 0.44 seconds (compare the similar 0.5 second interval in Boersma and Chládková 2013), and was followed by 0.24 seconds of silence. The interstimulus interval of 0.44 seconds ensures that non-linguistic perception is triggered: immediate acoustic memory typically lasts up to 4 seconds (Thaut 2014) and if an interval between two sounds is of such a short duration, the sounds are more readily processed phonetically. If the interval is longer, acoustic memory no longer helps and perceivers then resort to categorising strategies, i.e. they make use of phonemic classification (Nygaard et al. 1995; Gerrits and Schouten 2004, though see also Gerrits 2001 for counterevidence against influence of interstimulus intervals). The literature presents mixed results wrt. the importance of interstimulus interval duration, but the short interval used here at least favours phonetic perception *if* there is such a difference. Participants were asked to take a break after the first 35 tokens (i.e. after completing level 4), before doing the remaining 35 tokens.

2.3.2.1.2 Part B: phonemic perception

After finishing part A, the participant is told that (s)he will now learn a few words of the mysterious language. David Aliverock insists on the fact that "all words are different and mean different things" (cit.). At each of the 3 levels, the participant takes part in a learning session, a training session and then the actual test. During the learning sessions, the participant is presented (triggered by a mouseclick whenever choosing to proceed, i.e. at self-paced rhythm) with 6 (level 1), 4 (level 2) or 2 (level 3) images of an animal or object related to the jungle, simultaneously with an acoustic input. This acoustic input is a nonce word of the shape /a_a/, with either one out of seven sibilants or a filler (/p, p^h, b, k, q/). For example, the image of the sun appears on the screen as the participant hears [as̺a] with a sibilant CoG of 5959 Hz. The ERB values for the sibilants are the same as in part A, except that they were extracted from natural tokens produced by the author of this dissertation rather than modified with a formula. After learning the new words, the participant is led to a training phase in which a token already learnt is heard, and (s)he must select the corresponding picture (e.g. when hearing [as̺a], (s)he must click on the sun). The tokens are randomly drawn from the pool of words already learnt and only eliminated from this pool when the participant answers correctly; this means that if a participant deliberately answered wrongly each time a token is heard, (s)he would be stuck eternally in the training session. Of course, since participants are trying to answer correctly, they eventually finish the training session sooner or later (out of a logical minimum of 36 trials for the 3 training sessions, the soonest was 56 trials and the latest was 161, with an average of 76). Each time an incorrect answer is given, the picture that was the correct answer appears on the screen. Importantly, these training sessions allow to balance out the effect of person-specific memory capacity: a participant with a less efficient memory will have been trained proportionally more, thereby compensating for the memory disadvantage. An indicator of less efficient memory is the number of incorrect answers in which a participant answered with the picture of a different phonological class than that of the correct answer (e.g. answering /aqa/ for /asa/): it is extremely unlikely that such an incorrect answer is due to a difficulty in distinguishing a uvular stop from a sibilant acoustically. It turns out indeed that the participants who made more errors across phonological classes (i.e. who had more difficulty remembering the image-sound associations) were more likely to remain longer inside the training sessions ($r = 0.875$, $p < 0.001$). The percentage of errors across phonological classes wrt. the total number of trials in the training sessions ranged from 0% (for 10 out of 19 participants) to 21.7% (a clear outlier). The average percentage for all participants was 3.3%. In the actual tests, the errors across phonological classes ranged from 0% (for 14 out of 19 participants) to 11.1% (the same outlier), with an average of 1.6%. It could not be established whether the 10 participants who made 0 errors across phonological classes in the training sessions were significantly more likely to perform better or worse than the others in identifying sibilants in part B ($p = 0.422$ *t*-test),[3] which does not allow to determine whether sibilant identification was significantly impacted by memory capacity. The drastic reduction of errors across phonological classes from the training sessions to the actual tests (-46.4% on average, where relevant), combined with the drastic reduction of any type of error from the training sessions to the actual tests (-19.6% on average), point in any case towards the interpretation that the training session loop proved helpful to reduce the effect of memory capacity.

After completing the training session, participants enter the actual test. It works precisely as in the training session, except that participants are allowed to click on the "Replay" button at most once to hear the word again, that incorrect answers of course do not come back to make the test last longer and that participants are asked to confirm their selection by pressing the space bar. A break is offered to participants after the test of level 1.

At the end of the test, Polish participants are asked how they pronounce the word 'frog' (*żaba*) and whether most people they know pronounce it the same way. This serves to identify any potential influence of *mazurzenie*, a dialectal phenomenon in which a series of dental sibilants can correspond both to dental and retroflex sibilants in standard Polish. The word *żaba* is expected to be realised with an initial retroflex in standard Polish and an initial dental /z/ in *Mazurzenie* varieties. *Mazurzenie* is a frequent phenomenon in dialects of Eastern Poland (Stankiewicz 1986; Gogolewski 2001: 128). Nevertheless, the vast majority of varieties with *mazurzenie* has [ẓ] from historical /r^j/ (and its voiceless counterpart in voiceless contexts) and therefore, to some extent, a sibilant inventory parallel to the standard Polish three-way distinction /s̺ - ɕ - ṣ/. All of the 3 Polish participants declared having been taught standard Polish with the pronunciation [ẓaba], and not having relatives who pronounce otherwise.

2.3.2.2 Tokens

For the purposes of this experiment, the following tokens were used: 7 recordings of the shape /aSa/ for part A, 3 times 7 recordings of the shape /aSa/ for part B and 3 recordings each of /apa/, /ap^{h}a/, /aba/, /aka/ and /aqa/ as fillers for part B. The sibilants had CoGs ranging from 22.05 ERB (2287 Hz) to 29.7 ERB (5959 Hz), with 5 intermediate steps spaced precisely 1.275 ERB each time. To generate these sibilants, the sibilant with the highest CoG was recorded by the author of this dissertation and then modified with a formula

[3] The 5 participants who made errors across phonological classes in the actual tests, however, made more errors in identifying sibilants in the actual tests ($t = 2.675$, $p = 0.016$), but that is because 64% of errors across phonological classes were precisely a sibilant identified as a non-sibilant or inversely (i.e. 64% of it is simply the same data counted twice and it therefore correlates).

in Praat to lower its spectral shape, as shown in fig. XXVIII. The sibilants of part B

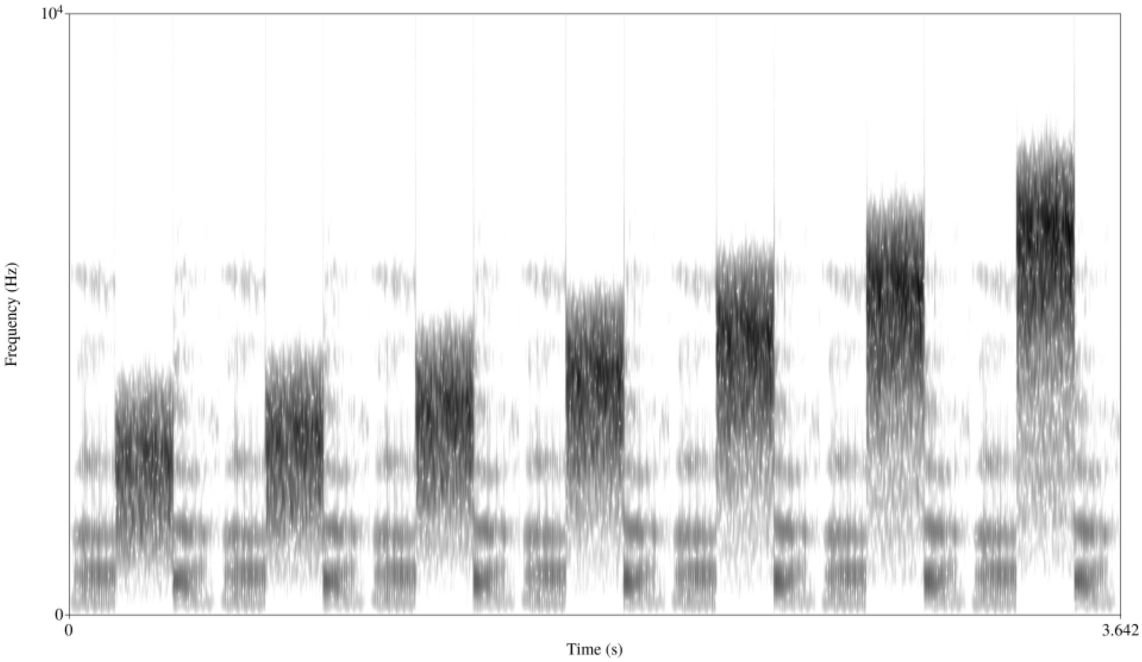


Figure XXVIII: *Spectrogram of the 7 tokens of part A, from [aṣ̠a] to [as̪̣a].*

were identical in terms of CoG, but were all natural sibilants recorded by the author of this dissertation rather than modified ones. Also, each token was recorded thrice. In the learning sessions of part B, participants hear each time one of the 3 recordings (e.g. recording 1 of sibilant X in learning session 1, recording 2 of sibilant X in learning session 2 etc.), something which ensures that they will have heard each recording for each sibilant before starting the third test. In the training sessions, the recording to be played is selected at random.

All tokens were recorded in a silent room (with doors and windows closed, at around 03:00 in the night in a fully silent building and street) by the author of this dissertation, with a Sanken COS-11D PT Lavalier microphone. The sampling frequency was 44100 Hz for all sounds used in this experiment, i.e. frequency values ranged from 0 to 22050 Hz. All sounds were also normalised in terms of intensity: the sibilant segments in the tokens of part A did not deviate for more than 0.055 dB from 65.276 dB, and the sibilant segments in the tokens of part B did not deviate for more than 0.00002 dB from 64 dB; the intensity of the full /a_a/ tokens did not deviate from 64.53 dB (for part A) or from 64 dB (for part B) for more than 0.030 dB. In terms of duration, the largest deviation from a duration of precisely 0.2 seconds was 0.0002 seconds for sibilants and 0.01 for the non-sibilants. In the non-sibilants, the intensity and timing of closure and stop burst was strictly aligned where possible (e.g. not more than 0.005 seconds deviation from the average burst moment; for /p^h/, the duration had to be longer than that of the others because of the aspiration phase). All of this ensured that participants could not rely on differences in intensity or duration to recognise a sound but rather have to rely exclusively on information in the spectral shape. Every time a segment was cut and extracted to be inserted in the [a_a] mould, zero crossings were selected on both sides in Praat to avoid unnaturally sounding transitions. In terms of Centre of Gravity, the sibilants of part A did not deviate from the intended CoG for more than 2.97 Hertz; this largest deviation was 5.47 Hz for the sibilants of part B. Both are illustrated in fig. XXIX, which shows how precisely the sibilants matched their intended CoG. At level 1, participants

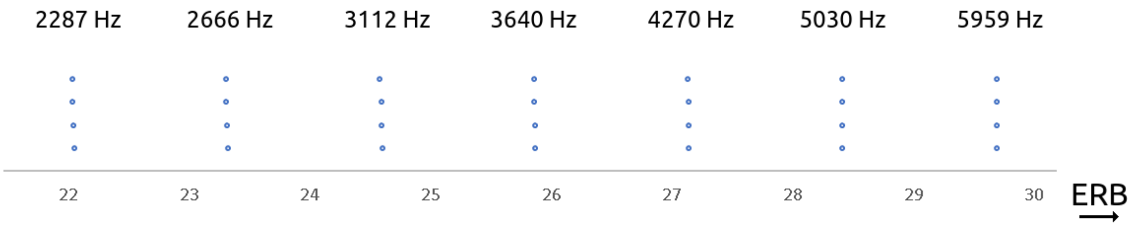


Figure XXIX: *ERB values of the 7 tokens of part A and the 3x7 tokens of part B, from [as̠̹a] to [as̪a].*

learn words with 3 sibilants and 3 fillers: the sibilant with the highest CoG at 5959 Hz (corresponding to the picture of the sun), that with the lowest one at 2287 Hz (picture of a boat), that with the CoG perfectly in the middle at 3640 Hz (picture of a jaguar) and 3 pictures for /apa/ (Marsupilami), /apʰa/ (tent) and /aba/ (Piranha fish). This ensures that participants do not immediately understand that the task actually is about sibilants, because they must also concentrate on the non-native contrast /p - pʰ - b/ in half of the tokens. At level 2, two words with sibilants at 4270 Hz (parrot) and 3112 Hz (native inhabitant of the tribe) are added, i.e. with the two sibilants immediately bordering the middle sibilant, and the 2 fillers with dorsal stops (toxic waste and frog). At level 3, the 2 remaining sibilants are added (conquistador and llama). Participants are thus initially presented with 3 sibilants maximally dispersed in acoustic space, but the acoustic space in between is gradually filled with additional sibilants to distinguish.

2.3.2.3 Participants

Participants were split in two groups: native speakers of Brussels French and native speakers of Polish. All participants signed a consent form that described the experiment as an experiment about perceptual acuity and categorisation, without mentioning sibilants. The consent form and experiment were verified and approved with modifications by the ethical commitee at UniVR. Variables that were controlled are:

1. Age
2. Sex
3. Mother tongue
4. Languages spoken
5. Higher education background
6. Any particularity to be reported wrt. language skills, sight or hearing (e.g. absolute pitch or hearing impairment)

7. The answers and reaction time for each token, subsequently averaged across levels and across parts A and B

1. With respect to age, all participants were between 18 and 28 years old, with an average of 21.2 years old. The sixteen participants of the Brussels French group were aged 20.3 years old on average, the three participants of the Polish group 26 years old on average. No statistically significant correlation was found between the age of the participants and their ability to discriminate sibilants in part A ($p = 0.887$), part B ($p = 0.376$) nor any level of part B. No correlation was found with reaction times, nor with the number of languages spoken. Age only interacted with participant group (because the Polish participants were significantly older) and with sex, to the extent that female participants were by coincidence younger on average than the male participants. Age differences were thus small (at most 10 years) and predominantly non-significant.

2. 12 participants were male, out of which 2 in the Polish group; 7 were female, out of which 1 in the Polish group. The female participants scored slightly better in distinguishing sibilants than the males, with non-significantly better scores in phonetic perception ($p = 0.279$) and in phonemic perception ($p = 0.227$). This must not necessarily be linked to a better perceptual acuity, but could also be explained by a higher level of attention and willingness to perform the task precisely. Nevertheless, no significant interaction was found between reaction times and sex, which was only shown to interact significantly with age.

3. The mother tongue (Brussels French or Polish) was not found to interact significantly with anything but age (since the three participants of the Polish group are older on average). The Polish participants had approximatively equal scores both in phonetic and phonemic perception with respect to the Brussels French group.

4. The number of languages spoken was not found to interact significantly with any parameter. It was strictly controlled for participants of the Brussels French group that they were not exposed to or fluent in any language that does not have a sibilant inventory with and only with [s] and [ʃ] as PoAs. This means that speaking Castilian Spanish, Danish, Polish, Russian etc. was a reason sufficient to exclude participants. The languages in which participants of the Brussels French group reported being fluent were: English, Flemish Dutch, Latin American Spanish, Portuguese, Italian, German. All of these languages contrast (denti-)alveolar with palatoalveolar sibilants, like Brussels French. On average, they spoke 3.375 languages including Brussels French. The speakers of Polish reported being fluent in languages with a /s - ʃ/ contrast (something which is inevitable for plurilinguals in a world with English as *lingua franca*): English, Brussels French and Flemish Dutch. Besides this, one participant spoke the Serbian variety of standard Serbo-Croatian, which has 3 PoAs like Polish. The Polish participants spoke on average 3 languages including Polish.

5. All participants were about to start university, or studying at university, or had finished studying at university. Differences in education background were limited.

6. All participants reported having normal linguistic skills (e.g. no dyslexia or language impairment). One participant in the Brussels French group reported having had an artificial eardrum set in his ear at a young age (around 5 years old). The participant

did not report having any particular difficulty or ease to hear with respect to the average. In part A, the participant performed 9% better than the others (81.4% vs. 72.4% correct discrimination) and in part B, (s)he distinguished sibilants 10% better than the others (57.1% vs. 46.8% correct sibilant identifications). None of these results turned out to be statistically significant.

Two participants, one in each group, reported having absolute pitch. Both had studied music and piano or singing since an early age. Their scores were only 4% higher than those of the others for part A (non-significant $p = 0.398$), but 32% higher for part B (76.2% vs. 44%; $t =$ -3.088, $p = 0.007$). Within the levels of part B, they were better in all levels but only significantly so in level 2 (i.e. with 5 sibilant PoAs; $t =$ -3.157, $p = 0.006$). Their reaction times were significantly longer than the average for part A ($t =$ -2.86, $p = 0.005$), but not for part B. They thus clearly performed better in part B but since they are evenly distributed in the Brussels French and Polish groups, and since their results were similar, they were not excluded from the analysis.

2.3.3 Results

2.3.3.1 Part A

In part A, the correct identification rates (i.e. the percentage of correct "same" or "different" answers) per participant for all levels ranged from 61.4% to 82.9%; the average was 72.9%. The perceptual acuity in distinguishing differences in sibilant CoG was thus above chance for all participants, as shown in fig. XXX. To provide a base of comparison, young

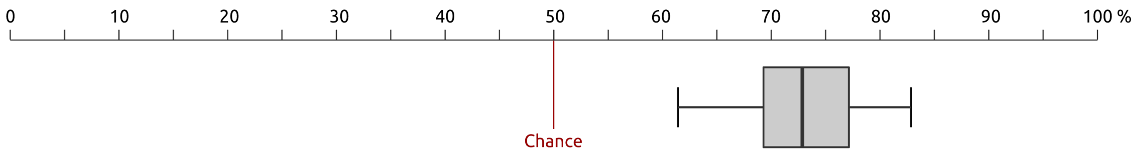


Figure XXX: *Correct identification rates for all participants at all levels (JASP Team 2021).*

participants (18-35 years old) with normal hearing scored on average 87% for the recognition of native phonemes in 'home' conditions (i.e. with a little background noise, +8 dB SNR) in Billings et al. (2016: 80)'s study on age- and environment-dependent phoneme perception. The average is clearly lower here, although the fact that the sounds were phonetic rather than native as well as the fact that participants were in an almost silent room rather than 'home' conditions are expected to change the average scores in both directions. The average correct identification rate for the same participants in Billings et al. (2016)'s study was 80% in 'store/hospital' conditions (+4 dB SNR), which can be considered to represent a sound illustrative threshold for normal identification rates at ages 18-28. In what follows, correct identification rates approaching or surpassing 80% will be considered to correspond to what is stably learnable. Thus, if a participant scores e.g. 64% at level X, it will be considered that the participant would not be able to speak a language that has the number of contrasts at level X satisfactorily.

When separating per level, as could have been expected, it becomes clear that correct identification rates correlate inversely with the level number, so that higher levels have lower scores (r = -0.584, $p < 0.001$; also significant when including participant as a random effect). This is illustrated in fig. XXXI. Nevertheless, it can also be seen in

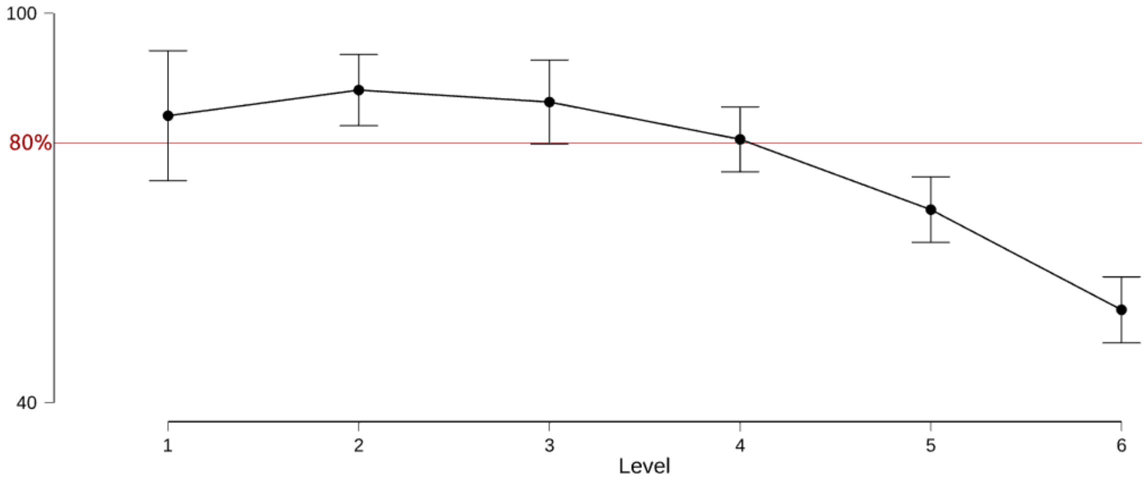


Figure XXXI: *Correct identification rates from level 1 to level 6 (JASP Team 2021).*

fig. XXXI that level 1 constitutes an exception to this trend, in the sense that correct identification scores for its 4 sibilants were lower than at levels 2 and 3. An observation of the reaction times for each level provides an explanation: level 1 had significantly longer reaction times than any other level ($p < 0.001$ for all comparisons of level 1 with the other levels, whereas all other levels have no significantly different reaction time among themselves). The average reaction time per token was 2.3 seconds at level 1, and 1.1 seconds at the other levels, i.e. more than twice as long. There is thus a *first-test-effect*: right after starting part A, participants discover the tokens for the first time and try to understand what precisely was meant with “identical or different” in acoustic terms. Error rates and reaction times are thus higher because the participants need a bit of practice before executing the task well. In part B, since there was a training session before each test, this *first-test-effect* should be (and was) absent.

Judging from fig. XXXI and its correct identification rate dropping below the 80% bar at level 5 (i.e. when the acoustic distance between two sibilants becomes 2.55 ERB), participants should be able to fit at most 4 distinct although partially overlapping categories on the phonetic continuum. This can be seen in fig. XXVI, where level 4 has space for 4 categories. An abrupt step is taken from level 4 to level 5, where the number of participants with scores above 80% go from 9 down to 3. At level 6, which corresponds to distinguishing 7 sibilant PoAs, the average correct identification rate reaches chance level (54.3%) with only 2 participants who are near 80% by having an average of 73.7%.

A small inverse correlation between reaction times and correct identification rates was found (r = -0.155, $p < 0.001$). Incorrect answers were significantly preceded by more time, as participants were hesitating strongly between “identical” and “different”. Rather than finding speed-accuracy tradeoffs, one thus finds that answers were more likely to be correct if they were easy enough as to be answerable quickly. At the participant level however, speed did not correlate with accuracy ($p = 0.426$).

The sibilant with a CoG of 5959 Hz had a higher correct identification average than the other sibilants ($t = 3.283$, $p = 0.001$), with an average of 82.6% versus 71.2% for the other sibilants. The explanation for this result could be that the sibilant at 5959 Hz is the only token that was recorded naturally, and thus that participants are sensitive to the naturalness of the recorded vs. modified tokens. This explanation is supported by the fact that when removing the sibilant at 5959 Hz from the data, no correlation between CoG and correct identification can be found ($p = 0.848$).

2.3.3.2 Part B

Whereas the chance level was situated at 50% in part A, participants who would answer randomly in the three levels of part B would only have 1 chance out of 6, then 1 out of 10 and then 1 out of 12 to answer correctly, respectively. The average participant answered well above chance at level 1 (78.9% vs. 16.7% chance), at level 2 (59.4% vs. 10%) and at level 3 (52.6% vs. 8.3%). At any level, the participant with the lowest correct identification rate at that particular level still had a score at least twice that of chance rates. This shows that all participants were aiming at performing well.

As could be expected, correct identification rates were much lower on average for words with sibilants (47.4%) than for words with fillers (76.8%; $t = 8.098$, $p < 0.001$). Reaction times were also much longer for words with sibilants, since they were more likely to be difficult to identify. The following results focus exclusively on the words with sibilants.

Fig. XXXII shows that the correct identification rate for sibilants drops abruptly from level 1 (3 sibilant PoAs) to level 2 (5 sibilant PoAs) and 3 (7 sibilant PoAs). The average

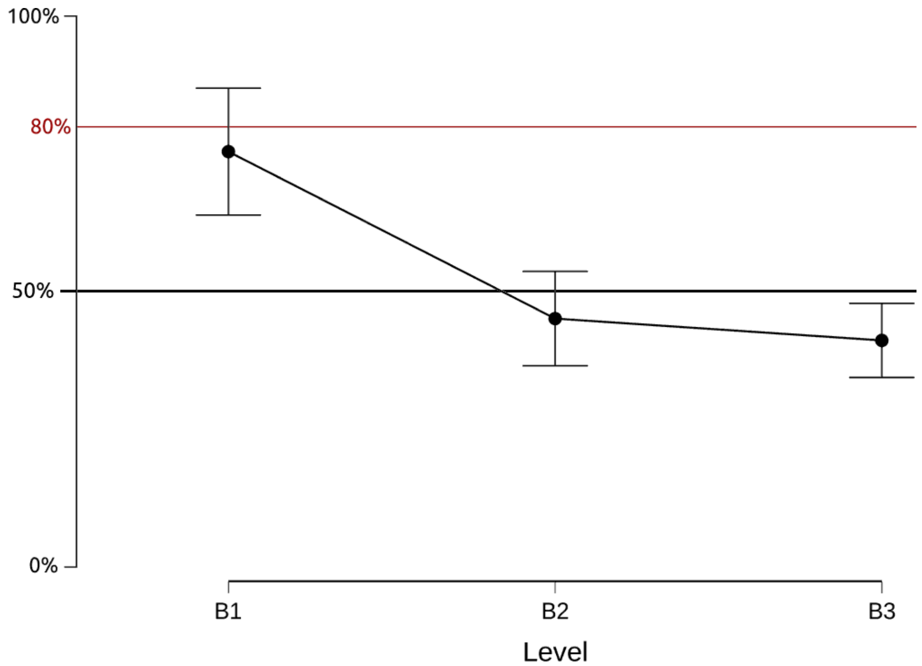


Figure XXXII: *Correct identification rates from level 1 to level 3 (JASP Team 2021).*

correct identification rate indeed goes from 75.4% at level 1 to 45.1% at level 2 and 41.1% at level 3. It is still surprising that the 16 native French speakers performed well at level 1 (72.9% on average) despite not having 3 sibilant PoAs in their mother tongue; it must however be observed that the most anterior and the most posterior sibilant were slightly more dispersed away from each other than in natural languages with 3 sibilant PoAs (7.65 ERB, versus 6.89 ERB reported for Polish in Żygis and Hamann 2003), something which could explain this good performance. The different ERB values for the Polish sibilants reported in Żygis and Hamann (2003) are compared to the most anterior, most posterior and middle sibilants of this experiment in fig. XXXIII. The correct identification rates

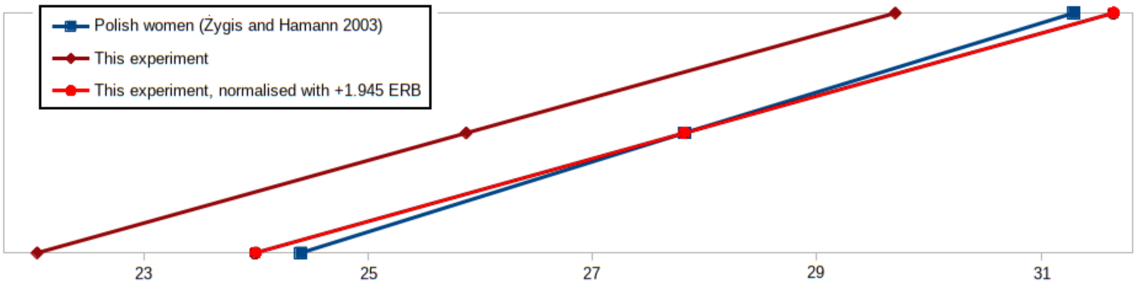


Figure XXXIII: *Spacing in ERB of the sibilants of Polish women in Żygis and Hamann (2003), versus sibilants at 2287, 3640 and 5959 Hz in this experiment.*

of Brussels French participants at level 1 indicate that 8 out of 16 were able to discriminate the 3 sibilants perfectly, 3 made one error and 5 made two errors. Interestingly, it is the anterior sibilant at 5959 Hz that was best distinguished by the Brussels French participants (81.3% vs. 68.8% average correct identification rate for the middle and posterior sibilants of level 1). This coincides with the difficulty of native French speakers (Shoemaker 2016: 8; similar to that of native English speakers, cfr. Lisker 2001, Żygis and Padgett 2010: 208, or even that of the automated classification of Polish sibilants, cfr. Bukmaier and Harrington 2016: 319f.) to distinguish the middle and most posterior sibilants from each other in languages with 3 sibilant PoAs, whereas /s/ is easily identified. This observation is not verified in the results of the Polish participants: they had 100% correct identification rates for the two frontmost sibilants and would also have had 100% for the posterior one if not for 1 error (mistaking it for the anterior sibilant). This difference in sibilant accuracy at level 1 between the native Brussels French and the Polish speakers was however not significant ($p = 0.316$), due to the insufficient number of Polish participants. At level 2, the difference in correct identification rates in favour of the Polish group was smaller and still non-significant while at level 3, it even reversed in favour of the Brussels French group despite not being significant ($p = 0.078$).

Fig. XXXII shows the results for 3, 5 and 7 sibilant PoAs. Since there were only 3 levels, the data in between these PoAs can only be simulated. This is what I undertook in fig. XXXIV: the answers of each participant were considered across all levels and the percentage of correct identifications for sibilants was calculated. Then, the sibilant pair that was most frequently confused by the participant was merged into one, and the percentage of correct identifications recalculated (6 PoAs). This is done again and again until all sibilants are logically merged to a single sibilant (1 PoA), where correct identi-

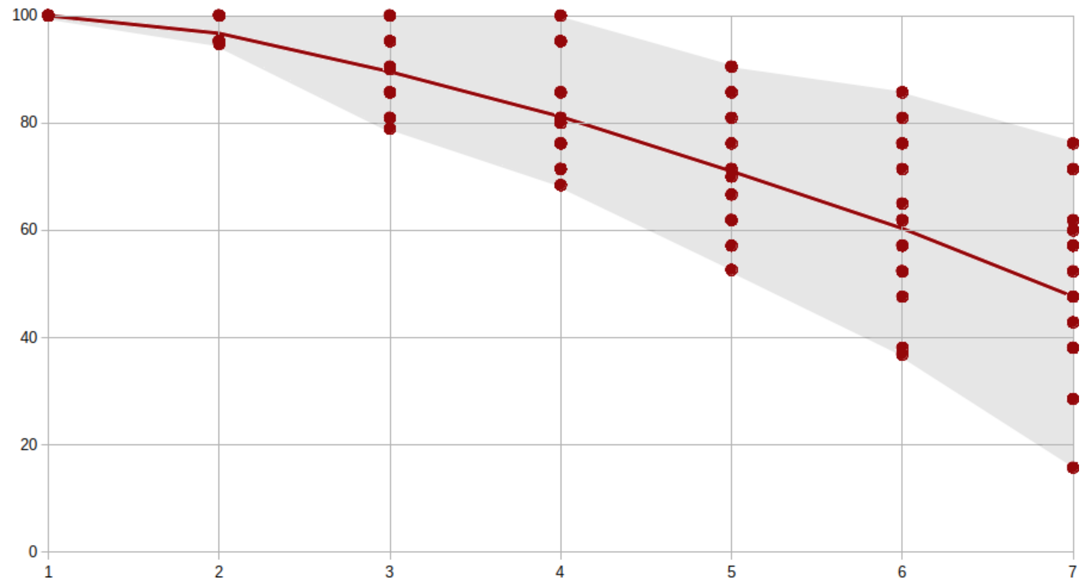


Figure XXXIV: *Merger projection of the correct identification rate for the 19 participants, from 1 to 7 sibilant PoAs.*

fications amount to 100%. Note that the results when merging the pair most confused by everybody rather than merging the pair most confused for each participant individually expectably yields lower scores (e.g. 66.8% rather than 71% correct identification at 5 PoAs). Such a graph allows to represent the curve of the average decline in accuracy as a function of the number of PoA distinctions, together with its widest deviation (in grey). Here also, the average score drops below 80% between level 4 and level 5.

The detailed confusion matrix for all participants and all levels is provided in fig. XXXV. In the matrix, the rows correspond to what was the correct answer and the

	aca	aza	asa	aja	axa	aha	ara	apha	apa	aba	aka	aqa
aca	37	7	4	4	1	0	4	0	0	0	0	0
aza	16	30	5	0	1	2	1	0	0	0	2	0
asa	3	3	17	11	21	2	0	0	0	0	0	0
aja	4	0	8	28	7	6	4	0	0	0	0	0
axa	2	2	9	19	18	5	2	0	0	0	0	0
aha	0	5	10	5	12	20	4	1	0	0	0	0
ara	3	0	4	5	2	3	39	0	0	0	1	0
apha	0	0	0	0	0	0	0	50	7	0	0	0
apa	0	0	0	0	0	0	0	7	43	6	1	0
aba	0	0	0	0	0	0	0	1	1	54	1	0
aka	0	0	0	0	0	0	0	0	2	0	36	19
aqa	0	0	1	0	1	0	1	0	0	0	18	36

Figure XXXV: *Confusion matrix (horizontal: correct answer; vertical: answer given) for the tokens of part B. Sibilants are ordered gradually from 5959 Hz* (aca) *to 2287 Hz* (ara)*.*

columns to what was actually answered (e.g. [aka] was perceived twice as [apa] and [apa] was perceived once as [aka]). Each token was asked precisely thrice in part B and there were 19 participants, so the sum of any row is always 57 (i.e. 19 x 3). The sibilants are ordered at the upper left of the matrix, from highest to lowest CoG. It can be observed in fig. XXXV that participants identified the extreme sibilants (5959 Hz and 2287 Hz) more reliably than the others. There was a significant correlation (r = -0.163, p = 0.001) between the level at which a word with a sibilant was learnt and its correct identification rate, in the sense that the sibilants of level 1 especially were more often correctly identified. This can be because they have been learnt across a larger time span, or because 2 of them have extreme CoGs. It can also be observed that the retracted alveolar sibilant (*asa*) was more frequently perceived as a palatoalveolar sibilant (*axa*) than as itself, something which corroborates the observations made in section 2.1.1 on page 11 wrt. retracted alveolars being misperceived as more posterior.

Precisely as for part A, reaction times correlated inversely with correct identification rates ($p < 0.001$). Participants thus hesitated longer before answering when the question was difficult, and were expectedly more likely to make errors when the question was difficult. At the participant level, participants who gave more correct answers had slightly shorter reaction times, but this trend was not statistically significant ($p = 0.286$).

2.3.3.3 Discussion

The experiment yielded results relatable to both language-particular and language-independent effects. On the language-particular side, the Polish participants were expectedly better at discriminating the 3 sibilants at level 1 of part B, which approximatively correspond to their native sibilants. However, they did not perform better than the Brussels French speakers in discriminating sibilants at the subsequent levels of part B, nor in discriminating differences in sibilant CoG in part A or any of its levels. Results for 5 sibilant PoAs in part B and for auditory distances corresponding to 4 sibilant PoAs (level 5) in part A indicate that a vast majority of participants would be unable to distinguish them reliably (all but 2 and all but 3 participants, respectively). If a language with 5 PoAs for sibilants existed, it would then see its sibilant contrasts reduce as soon as L2 speakers try to learn it: *more* than half of the time, L2 speakers would misidentify the sibilant they perceived, without even talking about the difficulty with which they would attempt to produce the 5 sibilant PoAs themselves. It thus seems safe to assume that 4 PoAs is the maximum learnable.

2.4 The sociolinguistics of sibilants

From a sociolinguistic perspective, sibilants are important social markers for several reasons. A first, obvious reason is that they are indicators of geographical, ethnic or class provenience (e.g. having a retracted alveolar /s/ in English ≃ being from Glasgow, cfr. Stuart-Smith 2020). This 'Shibboleth'-aspect, i.e. that the pronunciation reveals the social group of a person, is not specific to sibilants but to any aspect of social variation in phonetics or phonology, where also e.g. realising /r/ as alveolar [r] or uvular [ʀ] or doing final devoicing can indicate belonging to a social class, ethnicity and so forth. Summarising the literature on this topic, Themistocleous et al. (2016: 1027) mention that "a number of earlier studies showed that social factors, such as gender and age (e.g., see Jongman et al. 2000; Fox and Nissen 2005; Li et al. 2016), education, social identity, social networks (e.g., Baran 2014) and the place of origin, urban vs. rural (Dubois and Horvath 1998; Kochetov 2006; Stuart-Smith 2007; Mazzaro 2011) have significant effects on fricatives". The effects of belonging to a social group on the pronunciation of individuals is language-particular, in the sense that e.g. a more anterior /s/ might be associated with a high social status in one language and with a low one in another.

In what follows, I will concentrate instead on the universal sociolinguistics of sibilants, i.e. aspects of their pronunciation that carry the same mental associations in any language. These aspects are mainly:

1. Social stigma around misarticulation
2. Infantility
3. Sex and gender
4. Sexual orientation
5. Airflow sound symbolism

These five points correspond to the observations, respectively, that sibilants can lead to social stigma when misarticulated, that alveopalatal articulations are associated with infantility, that higher CoGs are associated with infantility, femininity as well as male homosexuality, and that sibilants are associated with the semantic fields of air and airflow in sound symbolic patterns.

1. Recall from page 11 that sibilants are among the latest speech sounds to be mastered correctly by children (Koenig et al. 2013: 1175; Sebregts 2014: 134). Their articulatory complexity provokes the fact that some individuals in a population do not necessarily manage to copy the articulation of one or several sibilants of their mother tongue accurately as pronounced by their environment, leading to a sibilant realisation considered 'deviant' from the norm. Such misarticulations are frequent in sibilants to such an extent that there is a specific word for sibilant misarticulation: *lisp*. Deviant pronunciations of sibilants are typically associated with infancy because children are overall less accurate than adults at pronouncing target-like native phonemes, but deviant pronunciations can also persist into adulthood. Nozaki et al. (2011) and Anjos et al. (2018) report that deviant sibilant realisations affect the quality of life of both children and adults negatively, e.g. by causing embarrassment or facilitating bullying. Deviant articulations include the interdental lisp, in which sibilants are fronted to interdentals, and the lateral lisp, in which their pronunciation approximates that of a lateral obstruent such as [ɬ]. Anterior sibilants (e.g. /s/) are typically affected by these lisps, but it also occurs that both anterior and posterior sibilants are affected. The social stigma around deviant sibilant realisations is quite unique, because sibilant misarticulations are very frequent

compared to other sound classes (especially in adulthood) and rarely go unnoticed because of their acoustic saliency (whereas e.g. uvular realisations of alveolar /r/ are more discrete in terms of perception).

2. Not only sibilant misarticulations, but also high CoGs, fronting (see 4.4.1.3) and alveopalatal realisations (see 4.3.1.1) are associated with infancy. As mentioned earlier in 2.1.2.1, children and women are more likely to produce sibilants with high CoGs than men because of their smaller resonance cavity. Children also typically front posterior sibilants (e.g. pronouncing [sasa] for *Sasha*; Stemberger and Bernhardt 1997: 299f.; Flynn 2012: 106) until the age of +/- 4 years, so that adults who want to imitate child speech are likely to front posterior sibilants. Even before the fronting phase, there is a phase in which children pronounce sibilants with "greater tongue-palate contact in their speech than adults" (Richtsmeier 2010: 7), which yields alveopalatal-like realisations. This explains why adults who want to express through babytalk how cute a baby is are more likely to pronounce utterances like [put͡ɕiput͡ɕi] rather than e.g. [rombaromba] (Kochetov and Alderete 2011: 363ff.). The association between infancy, smallness and palatality has yielded many patterns of expressive palatalisation cross-linguistically (Kochetov and Alderete 2011; Czaplicki et al. 2016; Alderete and Kochetov 2017), in which babytalk and/or diminutives are formed with surfacing alveopalatal affricates where the base/adult form has other segments.

3. Since female human beings are more likely to pronounce sibilants with a high CoG (see Fuchs and Toda 2010: 282 for an overview of the literature on this topic), sibilants with high CoGs have come to be associated cross-linguistically with femininity. Stuart-Smith (2007, 2020: 4) found for example that working-class girls from Glasgow lower the CoG of their /s/ to assimilate with the boys and men, compensating for the physiological constraints that lead them to produce higher CoGs. The association between 'high CoG-female' and 'low CoG-male' is thus not only a universally observed phenomenon based on *sex*; it is also actively used (be it consciously or not) by individuals to identify with a *gender*. For example, men who imitate women and take over the role of a woman are likely to produce an intentionally dental /s/ (Norton 1997: 20). Czaplicki et al. (2016) report the emergence of a new variant realisation of the Polish alveopalatal sibilants, in which specifically young women front them to (approximatively) [sʲ] instead of [ɕ]. They mention that this higher frequency variant has come to the consciousness of many Poles, one of which e.g. states in a radio program that this pronunciation "is supposed to make them sound kind, sweet and womanly", and that young women who use this variant "speak like small children" (Czaplicki et al. 2016: 3). One notices here again the association of fronted sibilants with both infantility and femininity.

4. As shown above, differences in the CoG of /s/ are not only indicators of sex, but also of gender. The cross-linguistic conceptual association of male homosexuality with higher CoGs (as with femininity in general) has been verified in several studies (e.g. Linville 1998; Munson et al. 2006; Mack and Munson 2012; Rácz and Shepácz 2013; Pharao et al. 2014; Tracy et al. 2015; Liem 2019), whereby most of the literature finds that higher CoGs are *perceptually* associated with being gay by listeners, and simultaneously that gay men do *not* pronounce sibilants with significantly higher CoGs in actual speech.

It is thus more likely to be a stereotype, linked with the stereotypical (and not necessarily true, in reality) association of gayness with femininity. The association was found not only in English, but also e.g. in Hollandic Dutch (Liem 2019), Hungarian (Rácz and Shepácz 2013) and Danish (Pharao et al. 2014).

5. Kawahara et al. (2020: 386), whose paper starts with the atypical question "Do sibilants fly?", mention a long history of sound symbolism in sibilants:

> Socrates in *Cratylus* suggests that [s] (=σ) and [z] (=ζ) are suited for words that represent wind and vibration, because the production of these sounds accompanies strong breath (427). Likewise, the Upanishads, ancient Sanskrit texts, suggests that "[t]he mute consonants represent the earth, **the sibilants the sky**, the vowels heaven. The mute consonants represent fire, **the sibilants air**, the vowels the sun" (...). [their emphasis]

This association of sibilants with wind, sky, air etc. was shown by Kawahara et al. (2020) to hold for the names given to Pokémons. Both [s̺] and [ɕ] were significantly ($p < 0.001$) associated with flying Pokémon types in Japanese, whereby the anterior sibilant was even more strongly so than the alveopalatal sibilant. The sound-symbolic association of sibilants with air and flight is furthermore supported by the cross-linguistic use of sibilants to imitate sounds like e.g. the whistling of very strong winds (typically voiceless apical anterior sibilants) or the flight of bees, flies and other flying insects (typically [zzzzz]). As to the origin of this association, Kawahara et al. (2020: 394f.) say:

> [T]his sound symbolic association is likely to have its roots in the fact that the production of sibilants involves a large amount of oral airflow to create frication noise, compared to the other types of sounds (Mielke 2012) – we can "hear" the air blowing/moving in a sibilant sound (...). We note, however, that it is not impossible to imagine that since sibilants have energy concentration in the high-frequency region because of their very small resonance cavities (Johnson 2003), this "highness" is iconically mapped onto the notion of sky, and by extension, to the notion of flying. See Parise et al. (2014) and references cited therein for the possible iconic connections between high-frequency sounds and the notion of elevation, which may also be relevant for the sound symbolic pattern that is identified in the current experiment.

3. A synchronic and diachronic typological database of sibilant inventories

This chapter consists of three sections. The first section, 3.1, explains according to what principles I have created a typological database of sibilant inventories. Section 3.2 then gives a quantitative (statistical) and qualitative overview of the different types of attested sibilant inventories, from a **synchronic** perspective. In section 3.3, attested transitions from one inventory to another are described from a **diachronic** perspective (3.3.1) followed by an overview of all the sound shifts that provoke these transitions and/or concern sibilants in general (3.3.2).

3.1 The database and its structure

3.1.1 Why a new database?

Several phonological databases of segment inventories do exist already, for example the UPSID (*UCLA Phonological Segment Inventory Database*, Maddieson 1984), the PHOIBLE (Moran and McCloy 2019) and the P-Base (Mielke 2007) database. These involve a large number of languages, with descriptions of their entire segment inventory aiming at giving an overview of most of their phonological traits. As a consequence of the rather generally-oriented (rather than sibilant-specific) aims and the resulting large workload required to fully describe the phonology of a language (especially when the literature is contradictory or imprecise), these databases are not free of error when it comes to the sibilant inventories described, as much as the literature on which they rest (Piñeros 2016). To give one example, Qualla Cherokee is mentioned in the P-Base as having a process of preconsonantal s-retraction from [s] to [ʃ], but it results from the thorough consultation of the original source cited there (King 1975) that the language actually only has a single retracted alveolar sibilant [s̠]. Such errors in language descriptions are much more likely to be avoided if one focuses on one precise aspect of the language (in this case: the sibinv) rather than the entire inventory and its processes. This is one of the reasons for which I undertook to set up a sibilant inventory database, documenting the largest possible typological variation while simultaneously being based on thorough investigation and comparison of sources. Doing so should reduce the error margin, since being intensely interested in one precise aspect of an inventory contributes to accuracy more than attempts that aim at capturing most if not all aspects of the phonology of a language.

Another important reason for the constitution of the SibInv database is the fact that the existing databases either do not allow retracted alveolar to exist as a category *at all*, simply classifying retracted alveolar sibilants as alveolar (e.g. UPSID, LAPSyD and P-Base), or that they allow the possibility of retracted alveolars but then still include many retracted alveolars as plain alveolars (e.g. Greek and Basque with /s̠/, but not Castilian Spanish, Hollandic Dutch, Danish, Faroese, Estonian, Finnish etc., in both the PHOIBLE and the database of Eurasian phonological inventories; cfr. Nikolaev et al. 2015). The fact that much of the literature (and thus, most segment databases) treat retracted alveolars like simple alveolars, despite their radically different distribution in the languages of the world, is in itself sufficient to motivate the creation of a new database. This misapprehension is, in fact, one among several recurrent inaccurate descriptions of sibinvs in the literature. These misconceptions are detailed in 3.1.5 with the aim of summarising the aspects of sibinv description that appeal particularly for caution.

3.1.2 The problem of genetical and geographical diversity

I tried to have a large genetic and geographic diversity in the sample, but did not necessarily succeed, since I gave the highest priority to accurately documented languages whose sibilants have been analysed by several researchers (rather than, stereotypically speaking, language isolates spoken in the heart of the jungle that have been described by one or two Westerners). This lead to e.g. 54.2% of the languages in the sample to be Indo-European, since IE languages are both widespread in the world and better documented than many others. Nevertheless, I do not consider this to be too problematic, in the sense that IE languages exhibit an enormous diversity and widely diverging diachronic developments of sibilant inventories, even inside one and the same language family (e.g. from Latin with 1s across 3s to French 2s, or from Old East Norse with 1s to the modern Swedish 3s). This relativises the potential impact of genetic biases on the distribution of sibilants: even if they are closely related, northern Italian dialects for example exhibit inventories as varied as the Lombard /s - ʃ/ or the Veneto/Trentino /s̺ - s̪/ (Quaresima 1957; Adams 1975; Alber and Kokkelmans in prep.). Areal phenomena inside Europe play a considerable role that might influence the proportions within the sample, but they still do not prevent Europe from having seen all kinds of sibinvs (except 0s) over the last 2000 years, something which is well represented in the sample. There might thus be an (Indo-)European bias in the SibInv, illustrated among others by a scarcity of 0s, but exactly the same bias does in fact exist in natural languages, with speakers of 0s being a minority precisely because of the contact-facilitated areal spread of sibilants (see 3.2.2.1). Considering that sibinvs change rapidly, genetic biases thus have a limited effect. This is demonstrated by the still very similar repartitions between 0s, 1s, 2s, 3s and 4s in the non-IE languages and the IE languages of the SibInv database, respectively, as shown in figure XXXVI. A t-test shows no significant difference ($t = 1.478$, $p = 0.141$), with

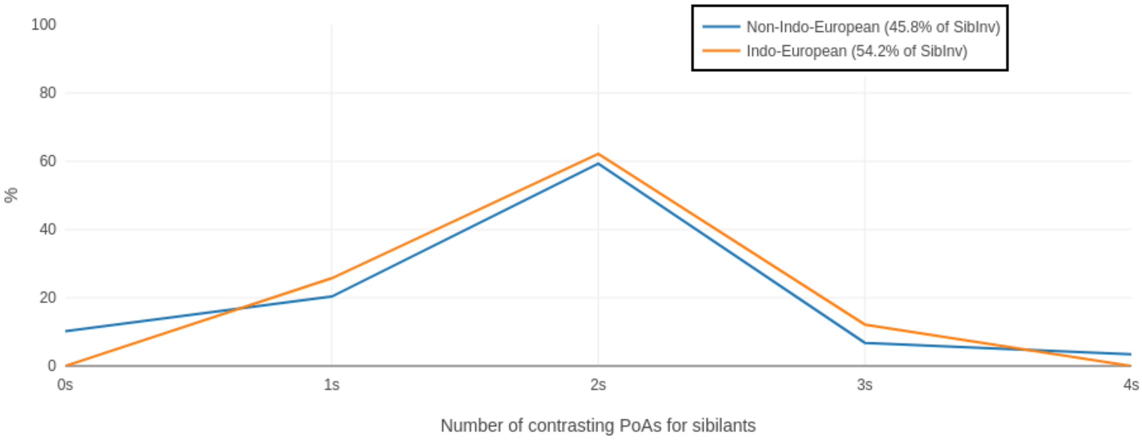


Figure XXXVI: *Comparison between the frequencies of sibilant inventories according to number of PoAs in Indo-European vs. non-Indo-European languages in the SibInv database.*

slightly more PoAs in IE than non-IE. The main priority of the SibInv database, next to favouring well-documented languages (see 3.1.4), is to include the widest possible variety of attested inventories of different kinds, in order to grasp all peculiarities of sibilant inventories. This purpose is summarised by Berns (2013: 52f.) as follows:

> A language database has to be primarily *descriptively* representative, i.e. it should include at least one instance of the different systems that occur in the world's languages. Therefore, the sample should be sufficiently large to assure that the likelihood that all possibilities are included is quite high. When this descriptive requirement is met, the sample can always be extended in such a way that it becomes more or less *statistically* representative, which makes it possible to examine relative frequencies of occurrence.

When talking about the different kinds of attested inventories (i.e. with respect to their different *qualities*), analysis will be based on the SibInv database; when talking about frequency of occurrence (i.e. wrt. *quantity*), data from larger databases like the UPSID (451 languages) and PHOIBLE (2186 languages) will be preferred. The successor of the

UPSID, called LAPSyD (Maddieson et al. 2014-2021), has not been designed to be a genetically or areally balanced database, unlike the UPSID; it will thus not be consulted for statistical queries.

3.1.3 Structure of the database

The database was created in the SQL format, with the intention to make it publicly accessible on the web at a later stage. The database is connected in the meanwhile to an own local web interface in PHP, Javascript and HTML allowing to visualise, update and customise it in a user-friendly way, as is visible in fig. XXXVII. Eight places of artic-

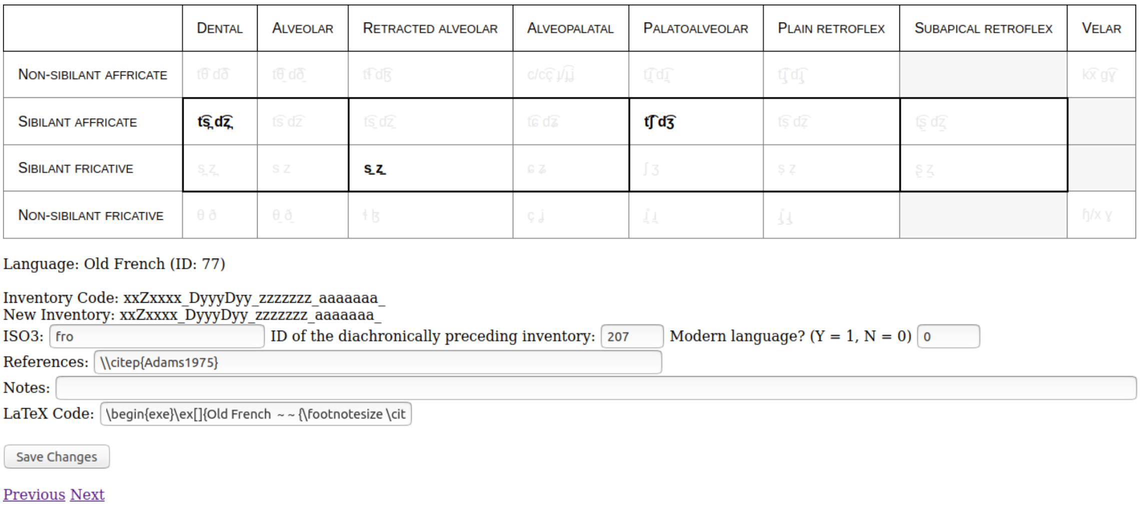


Figure XXXVII: *Screenshot showing the PHP interface of the SibInv database.*

ulation are distinguished in the database for phonetic precision, although not all PoAs can contrast in a language (Clements 2004: 4). The PoAs are ordered from left to right in function of their acoustics (as in table 1 in 2.1.1).

The thicker black lines in fig. XXXVII delimitating 4 groups of sibilant PoAs (dental and alveolar, retracted alveolar and alveopalatal, palatoalveolar and plain retroflex, subapical retroflex) correspond to the 4 possible place of articulation contrasts found in natural languages (see 4.2.1) and therefore also to the 4 possible PoAs in the OT analysis (5.2). Several reasons to consider the alveopalatal sibilants as pertaining to the second and not the third category (i.e. together with retracted alveolar and not palatoalveolar) are given in 2.2.1. Dental and alveolar sibilants never contrast in natural languages, but are distinguished in some language descriptions; they are distinguished here for the sake of precision. The distinctions made within the second and the third category are distinctions of distributedness (laminal alveopalatals and apical plain retroflexes).

Relevant non-sibilant fricatives and affricates were also taken into account, since they have been shown to influence or be influenced by acoustically close sibilant segments in terms of auditory dispersion (see e.g. Pétursson 1971: 213 for the contrasts [θ] - [s] vs. [θ̠] - [s̠]). This is however an influence at the level of phonetic detail (e.g. English /s/ being tendentially apical (Dart 1991) to contrast better with /θ/, but with permissible variation), and does not imply relevant effects on the structure of the sibinv; the non-sibilants are thus reported for the interest of phonetic detail (i.e. allowing to predict auditory dispersion and distributedness distinctions). Since auditory dispersion is active at the phonetic surface level rather than at the grammatical level, non-sibilant fricatives and affricates are included with their allophones, unlike sibilants (see 3.1.4).

The languages are stored with a major language family specification (e.g. Indo-European) and a minor one (e.g. Romance). The repartitions of major and minor language families in the database are represented in fig. XXXVIII and XXXIX, respectively.

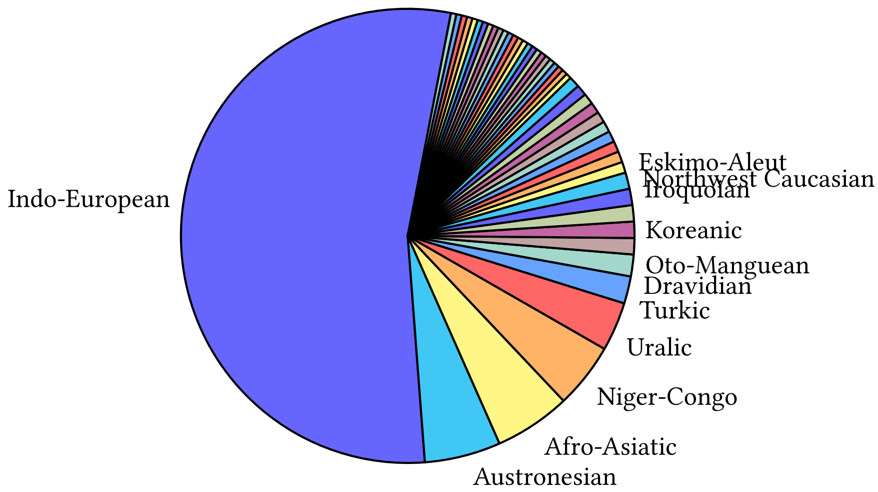


Figure XXXVIII: *Repartition of the languages in the SibInv database according to major language family. Families with frequencies under 1% are not labelled.*

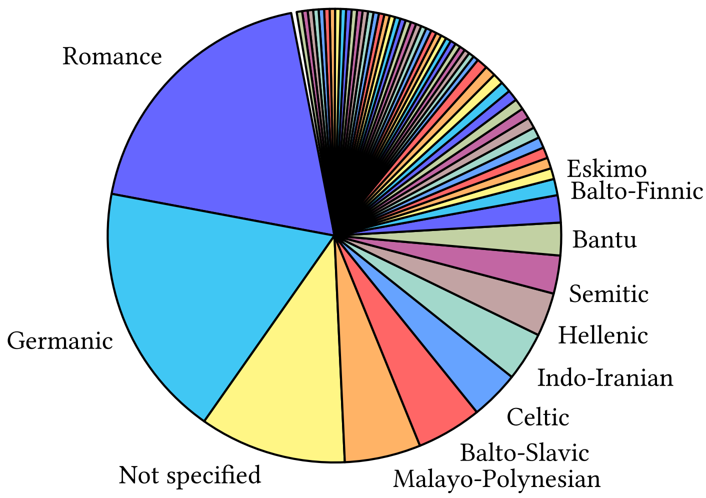


Figure XXXIX: *Repartition of the languages in the SibInv database according to minor language family. Families with frequencies under 1% are not labelled. 'Not specified' includes language isolates and proto-languages.*

Every language has an identification number (ID), and if a previous stage of the language is present in the database, it is referred to that language by means of its ID ('ID of the diachronically preceding inventory'). The parameter 'Modern language?' has the value *0* (= 'no') if the language has at least one descendant, *1* (= 'yes') if the inventory is the most recent exemplar of the language (note however that it could be extinct), and - if the database does not contain younger or older stages of the language. This makes it possible to visualise the diachrony of a sibilant inventory by navigating from the most recent inventory (i.e. starting from a modern inventory with *1*) to its oldest attested or reconstructed stage (i.e. with - for the diachronically preceding inventory).

ISO3 refers to the code of the language according to the ISO 639-3 standard. Some references are also provided for most inventories in the database, more or less proportionally to how underdocumented the language is; this means that languages without references in the database are likely to be e.g. French (my native language), German or English rather than Hawaiian or Abipón. The percentage of inventories without references is 29.45%. The optional notes for each inventory can be found in the list of all inventories in the appendix of this dissertation (8.8.1), as comments to the right of the inventories.

In total, there are 258 sibilant inventories in the database, among which there are 84 unique types (listed in 8.8.2). With 'unique types', it is meant 'all inventories when sibilant inventories that are perfect duplicates of another have been filtered out'.

3.1.4 Criteria for the inclusion of languages in the database

The following guidelines have been considered when adding languages to the database (ranked according to priority).

Selection:

1. Add a language to the database if its phonology has been described, preferably if its description is uncontroversial.
2. Add a language to the database if its sibilant inventory is peculiar and potentially unique among the other inventories of the database.
3. Add a language to the database if it is genealogically and/or geographically distant from the most represented language families/areas in the database.

Description:

1. If the language described has only one PoA for sibilants and the description does not make the distinction between alveolar and retracted alveolar, find audio samples of the language to judge acoustically.[1]
2. Only add phonemes, not allophones, to the sibinv in the database. For example, Japanese /ɕ/ is considered a phoneme because it occurs before non-high vowels

[1] This method might be prone to subjectivity or native language bias, but is more desirable than the other two options, namely 1) trusting blindly all language descriptions, even when contradictory and/or ostensibly inaccurate, or 2) trusting no language description and making a statistically significantly comprehensive acoustic study of all sibilants in all languages of the database, which would require years.

like /a/ since the disappearance of the trigger /j/; on the other hand, /ɕ/ in e.g. U.S.A. English *miss you* is considered an allophone and not included (Zsiga 1995).

3. Only add native phonemes, not marginal/loan phonemes (e.g. no [t͡s] in English, despite loanwords like *pizza*).

4. An exception to the previous rule is if there is evidence for a marginal phoneme that constitutes an independent PoA of its own. For example, Afrikaans /ʃ/ is considered a phoneme because of its contrast with alveolar /s/, despite its low statistical frequency. Omitting /ʃ/ would yield an otherwise unattested pattern and would be contradicted by the ability of native speakers to produce /ʃ/ (see 3.2.2.3).

5. Otherwise, be faithful to the original language description.

Several languages are currently experiencing or have recently experienced developments in their sibilant inventory, to the extent that it becomes neither legitimate or descriptively accurate to only add the older stage to the database, nor to only add the younger stage. In these cases, the two stages are represented as two distinct inventories in the database, named 'Older [language name]' and 'Younger [language name]'. For example, modern Övdalian (also called Elfdalian) has depalatalised the result of /k, g/-palatalisation to alveolar retracted sibilants (Sapir 2005: 21), while older stages of the language possessed those sounds as alveopalatal sibilant affricates. They are thus listed as follows in the database:

(11) Older Elfdalian (Sapir 2005; Garbacz 2008)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.				t͡ɕ d͡ʑ			
Sib. fri.			s̠				

(12) Younger Elfdalian (Sapir 2005; Garbacz 2008)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.			t͡s̠ d͡z̠				
Sib. fri.			s̠				

This distinction is especially needed when the two stages coexist in modern times, with an older generation of speakers fitting in the older stage pattern and the younger generation fitting into the more recent pattern.

Some sibinvs are much more stable than others: this difference is visible both in the statistical frequencies of stable and unstable inventories and in their diachronic stability itself. For example, 2s with a well-dispersed (denti-)alveolar - palatoalveolar contrast like English are more stable than 1s with a retracted alveolar - alveopalatal contrast like Hollandic Dutch (Boersma and Hamann 2008; Ooijevaar 2011). The former inventory is much more frequent than the latter (around 20 times more frequent in the Sibinv), and has remained a stable 2s for more or less 1000 years, while evidence from loanwords shows that the Hollandic Dutch posterior sibilants are about 400 years old and have not constantly been distinguished from retracted alveolars (e.g. Yiddish [ʃ] > Dutch /s̠/ in 1657 but Yiddish [ʃ] > Dutch /ʃ/ in 1765, then also sporadically Yiddish [ʃ] > Dutch /s̠/ in 1875 and 1901; cfr. van der Sijs 2001: 270f.).

(13) Standard Hollandic Dutch (Collins and Mees 1999: 160; Kwakkel 2008; Ooijevaar 2011; Seinhorst and Ooijevaar 2011)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.			t͡s̠				
Sib. fri.			s̠ z̠	ɕ ʑ			

The less-dispersed inventory represented here by Hollandic Dutch has furthermore been itself a short-lived intermediary stage from Proto-Anglo-Frisian to Old English. Despite this empirical distinction between stable and unstable inventories, I decided to consider the unstable ones to be possible inventories on an equal footing with stable ones, because any intermediary stage (as short-lived as it may be) should be considered a possible inventory since it existed or exists for at least some period of time.

3.1.5 Recurrent misanalyses of sibilant inventories

3.1.5.1 The forgotten/misunderstood retracted alveolar sibilants

Many of the (most widely spoken) languages in the world have a contrast between alveolar [s] and palatoalveolar [ʃ] (e.g. 6 out of the top 10). Accordingly, the number of linguists with English, French, German, American Spanish or Portuguese as a native language, and thus with [s - ʃ] as a native contrast, is high with respect to languages that have different sibilant PoAs. This has introduced a strong bias in the literature, where the many researchers having [s - ʃ] as a native contrast are likely to misidentify retracted alveolar sibilants as /s/ or /ʃ/. This problem is formulated by Gafos (1999: 154) and Bright (1978: 43) as follows:

> William Bright reports on his fieldwork on Karok in the mid 1950s: "I was having trouble with sibilants; in some words I consistently wrote an [s]; in others, I consistently wrote [š]" (Bright 1978: 43), the latter meant to be the same "sh"-type of fricative as the English [ʃ]. The problem was certainly not Bright's alone. Other linguists doing fieldwork in several aboriginal languages of California had gone through similar experiences, and Bright comments that the earlier literature on Californian languages, mainly that published in the U. C. Publications in American Archeology and Ethnology between 1900 and 1940, is particularly oblivious to this problem of categorization "seriously lacking in phonetic accuracy".
>
> At first, many words were written inconsistently with [s] and [š]; later, it

> was recognized that such words actually contained a sibilant which was somehow, articulatorily and acoustically, intermediate between [s] and [š].

It turns out that Karuk indeed has a retracted alveolar sibilant, which caused the many misinterpretations of the U.S. linguists:

(14) Karuk (Bright 1978)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.			t͡s̠				
Sib. fri.	s̪		s̠				

Even in more recent times, misinterpretations of retracted alveolars occur not so infrequently (e.g. in Quaresima 1957; Beale-Rivaya 2010; Árnason 2011: 122). Retracted alveolars are not only reinterpreted as either alveolar or palatoalveolar, as described above, but also sometimes identified as ‘retroflex segments’ (e.g. the Cimbrian [s̠] as [s̢] in Tyroller 2003 and Hall 2012, *idem* for Mòcheno in Meneguzzo 2017: 101f.). As mentioned earlier, databases such as the UPSID, LAPSyD and P-Base do not incorporate retracted alveolar as an independent place of articulation. The identification of retracted alveolars as such is crucial for any sibilant typology and any understanding of sibilant patterns: for example, it allows to explain why Mozarabic [s̪] (<c, z>) consistently shows up as /s/ in the contemporary Andalusian Arabic writing but [s̠] (<s>) variably surfaces as /s/ or /ʃ/ (Adams 1975: 285).

3.1.5.2 Alveopalatals as palatoalveolars

Padgett and Żygis (2007: 165) note that “sounds like [ɕ] and [s̪] are often transcribed as though palatoalveolars in grammars, for reasons of convention, typography, or a lack of interest in postalveolar details”. For example, Nogita (2016: 47) uses <ʃ> throughout to represent Japanese /ɕ/, and recognises at the same time that:

> [T]he Japanese /ʃ/ does not have lip rounding and dorsum elevation, unlike the [ʃ] English counterpart (Pan et al. 2004). Therefore, the Japanese /ʃ/ is transcribed as /ɕ/ by some linguists (Pan et al. 2004). As for /s/, as indicated by Li et al. (2009), the English /s/ is clearly alveolar while the Japanese /s/ is more laminal and possibly somewhat dentalized.

It can indeed be more convenient to use the more common IPA characters <ʃ> and <s> indistinctively for [ɕ] and [s̪] as much as for [ʃ] and [s], but this occurs at the cost of phonetic precision. When researchers then encounter [ɕ] and [s̪] as well as [ʃ] and [s] in the literature on e.g. Japanese, it is likely to lead them to confusion, if not error.

Another kind of misinterpretation is potentially caused by the placement of palatal segments *behind* posterior sibilants in the IPA chart (International Phonetic Association 2015). Alveopalatals, which are not included in the main IPA table itself but pattern phonologically with palatal segments, are then likely to be seen as more posterior than any other posterior sibilant. This is perhaps the reason for which e.g. Trask (2004: 19) defines 'alveopalatal' as "[a]rticulated in a position which is further back than palato-alveolar but further forward than palatal". While alveopalatal and palatoalveolar sibilants have a significant articulatory overlap in terms of PoA, where the most anterior and most posterior zones of contact of [ɕ] can be at the same time more anterior and more posterior than those of [ʃ], alveopalatals behave as more anterior in acoustic and phonological terms (see e.g. Jannedy and Weirich 2016, 2017).

3.1.5.3 Retroflexes as palatoalveolars

Like alveopalatals, retroflex segments are sometimes misrepresented as [ʃ]. For example, Żygis (2003: 177) says:

> It should be stressed that the retroflex status of the sibilant in Polish [ʂ] is by no means commonly accepted, especially by Slavic researche[r]s who transcribe this sound as postalveolar with the IPA symbol ʃ (...).

She then shows that the Slavic plain retroflexes are phonetically different from palatoalveolars. The other way around, some researchers have represented plain palatoalveolars as retroflexes, for example Flemming (2018) who lists Mazatec, Papago (O'Odham), Czech dialects and Slovene dialects as having a /s - ʂ/ contrast (and mixes *en passant* plain retroflexes with subapical retroflexes). Silverman et al. (2015: 83), Dart (1993), Šimáčková et al. (2012) and Šuštaršič et al. (1995) however speak in favour of the opposite interpretation /s - ʃ/. About Czech in particular, Hamann (2003: 40) mentions that she found "no phonological evidence but also no counterevidence for concluding that the Czech post-alveolar is retroflex", a consideration confirmed in Hamann (2004).

3.1.5.4 Impossible gaps

Some inaccuracies in transcriptions can lead to incorrect conclusions about phonemic gaps in sibilant inventories. For example, the Pama-Nyungan language Bandjalang is listed in the UPSID (Maddieson 1984) as having a voiced sibilant affricate /d͡ʒ/, but no voiceless counterpart. The reason for this categorisation is the fact that, as the UPSID mentions in the comments, "/dZ/ is described by Cunningham (1969) as 'identical with English "j"'". This contradicts the observation that the presence of a voiced sibilant tends to imply that of its voiceless counterpart (see 3.2.3.3). Yet, as much as all Australian languages (except those of the Torres strait), Bandjalang does not have /d͡ʒ/ but a non-sibilant palatal stop /ɟ/ instead (Dixon 2002). The researcher is thus likely to have perceived the voiced palatal stop as her native phoneme /d͡ʒ/, something which in turn influenced the data in the UPSID in a considerable way, something which in turn influenced the results of research based on the UPSID such as e.g. Berns (2013).

3.2 A synchronic typology of sibilant inventories

This section provides an overview of the attested sibilant inventories, giving first an estimate of their frequencies in natural languages (according to place of articulation, sibilant phoneme, voicing, and the fricative vs. affricate distinction), followed by a description of the different types of inventory attested. A typology of gaps in sibinvs is then presented, which shows where one can likely expect missing members of a sibilant pair.

3.2.1 Frequencies of sibilant inventories

3.2.1.1 Number of contrasting Places of Articulation

We will now compare the number of contrasting places of articulation for sibilants in three databases: the UPSID, the PHOIBLE and, for the reference (although one shall recall that this database is not designed primarily for statistical queries), the SibInv database. As figure XL illustrates, it results that languages with 2s are the most frequent in all three databases, with frequencies ranging from 47% to 61%; 1s comes second in all three, with values ranging from 23% to 38%. Then comes 0s, closely followed by 3s, with ranges between 5% and 17% and between 5% and 10%, respectively. 4s are very rare in all databases, and even absent in the UPSID.[2] In the PHOIBLE and SibInv, 4s are represented by the language Toda (Ladefoged and Maddieson 1996: 156-160) and Northwest Caucasian languages like Ubykh (Hall 1997*b*: 94). Regardless of the number of languages

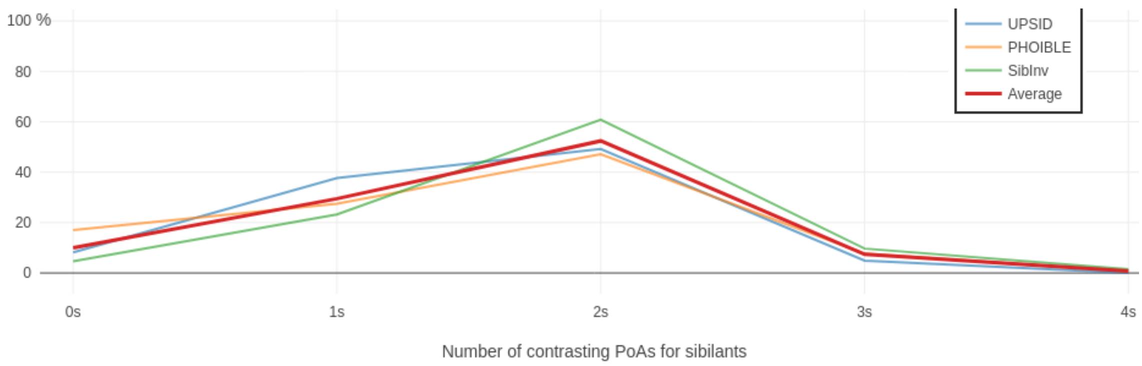


Figure XL: *Typological frequency of sibilant inventories according to the number of contrasting places of articulation for sibilants in the UPSID (Maddieson 1984), PHOIBLE (Moran and McCloy 2019) and SibInv databases.*

that are present in the respective databases (UPSID: 451; PHOIBLE: 2816[3]; SibInv: 258), figure XL clearly shows the general trend in favour of 2s and 1s among natural languages. The mean number of sibilant PoAs is furthermore very homogeneously located between 1.508 (UPSID), 1.472 (PHOIBLE) and 1.802 (SibInv). The latter slightly higher

[2] Acoma is listed there as a 4s, but is actually a 3s with an additional non-sibilant palatal stop.

[3] One shall remark that a significant number of languages is listed several times in the PHOIBLE, in the cases in which there are “distinct sources that disagree about the number and/or identity of that language’s phonemes” (Moran and McCloy 2019).

number (1.802) in the SibInv might be a consequence of the Eurasian/Indo-European bias described in 3.1.2, which is a bias in favour of sibinvs with many PoAs, because 0s are virtually absent in Europe and language contact throughout Eurasia favour the transmission of sibilants, making it so to say 'difficult' for a language to 'get rid' of all its sibilants.

The larger deviation of the numbers from the SibInv wrt. the average in figure XL can be attributed to the same bias ('fewer 0s, so necessarily more of the others'), and/or to the smaller number of languages in the sample, which tendentially allows for larger deviations from a certain pattern.

3.2.1.2 Frequencies of sibilant phonemes

Large discrepancies exist in the respective frequencies of sibilants in natural languages: although as many as 16 different sibilant phonemes can coexist in one and the same inventory (4 contrastable PoAs x 2 possibilities for voicing x 2 possibilities for fricative vs. affricate, as demonstrated by Toda in (15)), only a few of these sibilants are very frequent, while many are very infrequent. For example, retracted alveolar sibilants are found in Toda *and* in many other sibinvs, while subapical retroflex sibilants are found only in Toda and another handful of languages.

(15) Toda (Emeneau 1984; Ladefoged and Maddieson 1996; Boersma and Hamann 2008)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.	s̪		s̠		ʃ ʒ		ʂ̠ ʐ̠	

A look at the PHOIBLE reveals that one segment is extremely frequent in the world's languages, namely /s/ (which is used to represent both (denti-)alveolar [s̪/s] and retracted alveolar [s̠] in most of its language descriptions). This is detailed in fig. XLI, in which 'all variants of' indicates that secondary articulations (/sʲ, sˤ/ etc.) have been grouped together. All variants of /s/ taken together (i.e. after removing secondary articulations like /sʰ/ or /sː/ that are kept distinct in the database) reach 86% frequency in natural languages, which is practically equal to the proportion of languages that have sibilants *at all*. It can thus be safely assumed as a universal that:

(16) If a language has sibilants, it has at least one voiceless anterior (dental, alveolar or retracted alveolar) sibilant fricative.

The segment /s/ is followed in the frequency ranking by a group of five segments with frequencies between 30% and 60%, from more to less frequent: the posterior /t͡ʃ/, /ʃ/, the anterior /z/, /t͡s/, and the voiced posterior /d͡ʒ/. The fact that posterior /t͡ʃ/, /ʃ/ are more frequent than voiced /z/ seems to indicate that languages that have more than one sibilant segment prefer place contrasts rather than voicing contrasts. This prediction is borne out by the data in the UPSID: 63 out of 451 languages have voicing contrasts without PoA contrasts, while 99 languages have PoA contrasts without voicing contrasts; a *t*-test confirms this significance (t = -6.212, $p < 0.001$). If a language has no posterior

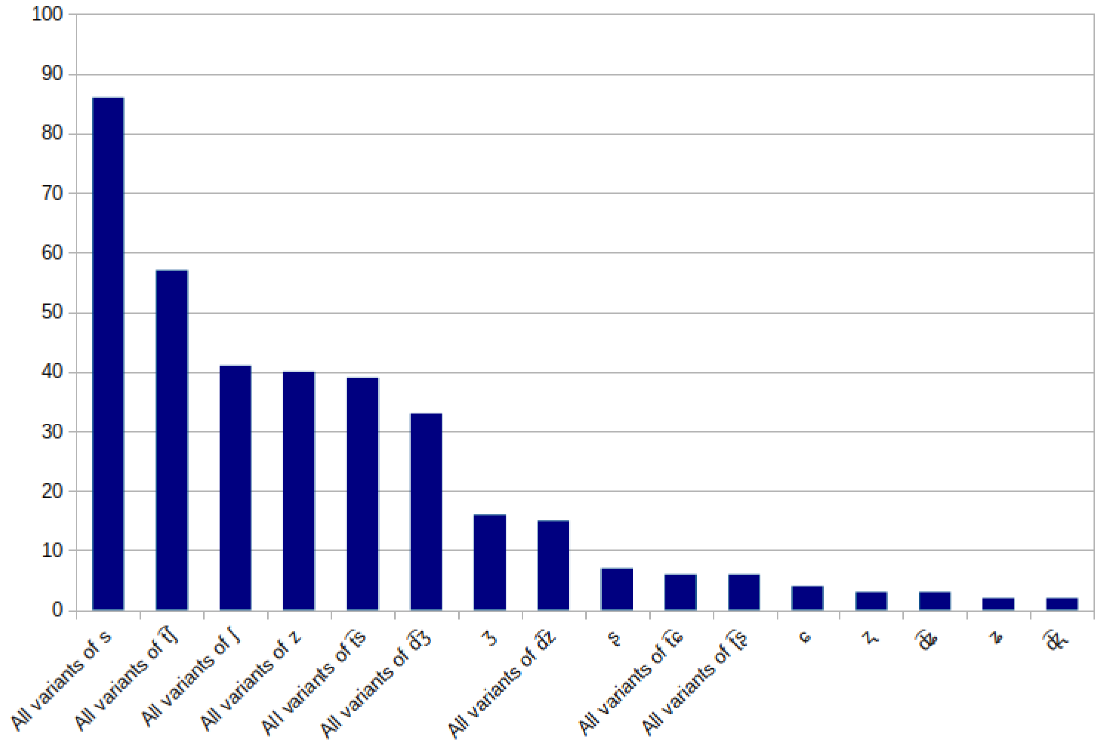


Figure XLI: *Typological frequency of single sibilant phonemes in the PHOIBLE database (Moran and McCloy 2019). Sibilants with a frequency below 1% have been omitted. For a detailed version of this figure, see the appendix (CXXII).*

sibilants, it has a 34.4% chance of having voicing distinctions, but if it has no voicing, it has a 45.2% chance of having posterior sibilants.

3.2.1.3 Voicing

As can be seen in fig. XLI, voiced sibilants are always less frequent in natural languages than their voiceless counterpart, without exception. This infrequency of voiced sibilants is observed within the class of obstruents in general, and widely acknowledged in the literature. Żygis, Fuchs and Koenig (2012: 301), basing their analysis on frequency counts in the P-Base database (Mielke 2007), summarise the infrequency of voiced obstruents as follows:

> For all three consonant types (stops, fricatives, affricates), voiced obstruents are less frequent than their voiceless counterparts. That is, these data are in accord with the traditional claim that voiced obstruents are marked relative to voiceless ones (e.g., Trubetzkoy 1939; Greenberg 1966; Chomsky and Halle 1968; Lombardi 1991, 1995, but see also Vaux and Samuels 2005 for a view that aspirated stops are the least marked segments). The voicing asymmetry is most extreme in the fricatives (245 vs. 491), and Maddieson 1984: 47 states that "generally, the existence of a given voiced fricative in the inventory implies the presence of a voiceless counterpart in the inventory."

This observation is also formulated as a weaker universal stating that the presence of [z] implies that of [s], as universal #1802 in the Konstanz Universals Archive (Plank and Filimonova 2000). The presence of a voiced sibilant in fact implies that of a voiceless sibilant counterpart in any sibilant inventory; this prediction is also borne out in the SibInv database, but not entirely in the other databases (see 3.2.3.3).

This does not, however, imply that inventories without voicing distinctions are more frequent than inventories with it. In the SibInv, only 32.11% of the languages with sibilants have no voicing contrast in sibilants. In the UPSID, this proportion amounts to 44.20%; in the PHOIBLE, it is 37.62%. With respect to the frequency of single segments, voiced sibilants as single segments are thus much less frequent than their voiceless counterparts, but at the level of the entire sibilant inventory, one is more likely to find a voicing contrast in at least one of the sibilants. A look at the proportion between fully voiceless inventories and inventories with voicing distinctions per number of PoAs reveals the interesting trend in fig. XLII. In the three databases, the difference in voicing

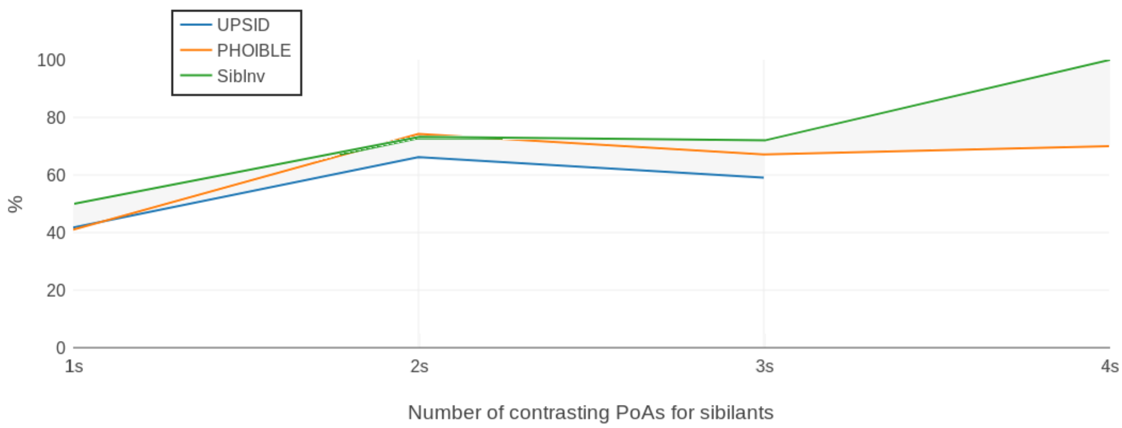


Figure XLII: *Proportion of sibilant inventories with at least one voicing contrast per number of places of articulation in the UPSID (Maddieson 1984), PHOIBLE (Moran and McCloy 2019) and SibInv databases.*

proportion as a function of number of PoAs is statistically significant (UPSID: t = -7.564, $p < 0.001$; PHOIBLE: t = -13.151, $p < 0.001$; SibInv: t = -5.418, $p < 0.001$). This correlation echoes the observation made above (3.2.1.2) that a sibilant inventory with two sibilant phonemes is more likely to have a PoA contrast than a voicing contrast. One can formulate the correlation between voicing and PoA more generally as follows:

(17) The more PoAs for sibilants a language has, the more it is likely to have phonemic voicing distinctions.

A possible motivation for this phenomenon could be that voicing contrasts are perceptually more salient than the smaller place distinctions in larger inventories: if a language has 2 or more PoAs for sibilants, keeping a voicing contrast proves to be more efficient in terms of successful communication than keeping a PoA contrast. This is because larger contrasts in terms of auditory dispersion are possible in inventories with less categories, and the more categories are added to the continuum, the more the acoustic distance between them decreases (Boersma and Hamann 2008: 222). The acoustic efficiency of voicing contrasts, on the other hand, is independent from the number of PoAs for sibilants (i.e. a /s-z/ contrast, for example, is as distinct in a 1s as in a 3s). In sibinvs with less PoAs, the benefit of having PoA distinctions is thus greater than that of having voicing distinctions, but as the number of PoAs increase, the tendency reverses.

Another possible explanation is that since voicing distinctions often appear in 'voiceless' sibinvs by means of the palatalisation of /t, d/ or /k, g/ (like e.g. in Latin > Romance and PIE > Balto-Slavic), the appearance of voicing contrasts correlates with the appearance of (voiced) posterior affricates, which themselves are likely to add more PoAs to the sibinv on the long term (for example, Latin 1s > Middle French 3s evolved exclusively through stop palatalisation). Telfer (2006: 74) provides an argument in favour of this possible explanation by noting the following:

> Strident coronal fricatives seem to have a voiced variant roughly one third of the time, as do dental and alveolar affricates. Voiced palatoalveolar [affricates] are far more common however, occurring in more than half of the languages that include *ʧ*. Maddieson (1984: 39) goes so far as to suggest that this discrepancy "may be related to the frequent historical descent of palatoalveolar affricates from velar or palatal stops".

3.2.1.4 Fricative vs. affricate

Only 3.99% (18/451) of the languages in the UPSID have at least one sibilant affricate but no sibilant fricative,[4] while 27.72% (125/451) have at least one sibilant fricative but no sibilant affricate (Maddieson 1984). As fig. XLIII shows, the number of languages without sibilant affricates is almost thrice the number of languages without sibilant fricatives. In the SibInv, there are 12 languages out of 258 without sibilant fricatives (corresponding entirely to the languages that have no sibilants at all), but 65 languages without sibilant affricates. In absolute numbers also, sibilant fricatives are more frequent than sibilant affricates: only a few languages (3.88% in the SibInv) have more sibilant affricates than sibilant fricatives, while almost half of the languages in the SibInv (43.41%) have more sibilant fricatives than sibilant affricates.

Żygis, Fuchs and Koenig (2012: 301) confirm this statistical predominance of sibilant fricatives based on the data in the P-Base database (Mielke 2007), in which coronal fricatives (N = 736) are 1.46 times as frequent as coronal affricates (N = 504). Affricates are more marked cross-linguistically than their fricative counterparts, to such an extent that it is claimed that affricates need the presence of fricative counterparts in order to exist in any given inventory, as explained by Żygis, Fuchs and Koenig (2012: 302):

> According to Greenberg (1963: 3), "There is an implicational universal that

[4] Even among these languages, many constitute doubtworthy cases such as e.g. Abipón and (South) Andamanese, where the affricate /tʃ/ is alternatively considered to be a (non-sibilant) palatal stop (Najlis 1966; Blevins 2007). Lev et al. (2012) say for example about Abipón: "It is uncertain whether the phoneme designated here as a palatal stop /c/ is better represented by a palatal stop /c/ or a post-alveolar affricate /tʃ/". The real number might thus be even lower than 3.99%.

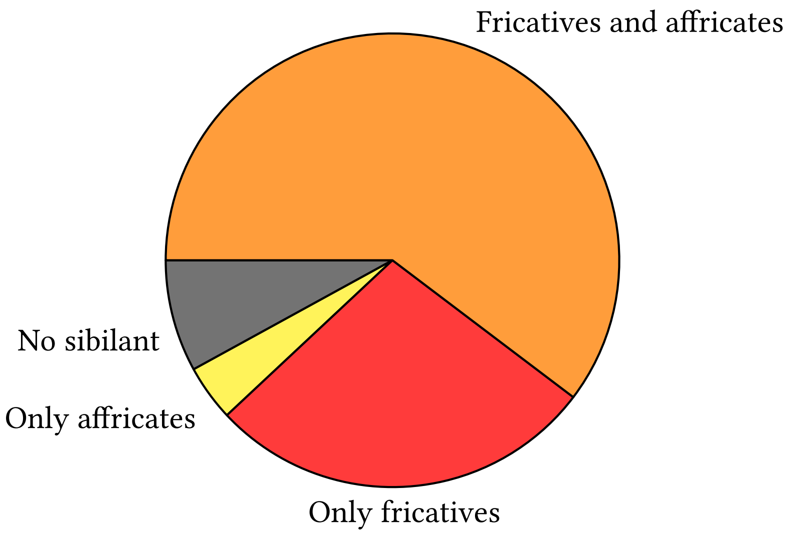


Figure XLIII: *Typological frequency of sibilant inventories according to manner of articulation (fricative or affricate), based on the data in the UPSID (Maddieson 1984).*

> an alveolar affricate such as /ts/ always implies the presence of /s/ in a language but not vice versa. (...) There are no languages with /ts/ that lack /s/." Jakobson (1972: 74) similarly states that the number of affricates in a given language never exceeds the number of fricatives in the same language. He also adds that the phoneme pair /t-ts/ implies the presence of the phoneme /s/ in a given language.

This implicational universal coincides with the observation that deaffrication, considered a kind of lenition (Marotta 2008: 255), is much more frequent in sibilants than its opposite (Kümmel 2007: 69f.), i.e. affrication (which either happens in very specific phonological contexts, e.g. after /n, l, r/, see B.7 in this chapter, or in loanword adaptation, e.g. Middle French > [z̺] > Middle Tuscan [ʒ] > modern standard Italian [d͡ʒ] in words like *prigione* 'prison' due to the historical correspondence Tuscan [ʒ] - standard Italian [d͡ʒ]; cfr. Adams 1975: 287).

The universal posited among others by Greenberg (1963) and Jakobson (1972) is however to be relativised: as detailed above, 3.88% of the languages in the SibInv have a number of sibilant affricates exceeding that of sibilant fricatives. Among them are e.g. Old French, which had the following sibinv besides the non-phonemic retracted alveolar [t͡s̺] occurring across morpheme boundaries (e.g. *des* [t͡ʃ]*a*[t͡s̺] 'cats'):

(18) Old French (Adams 1975)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.	t͡s̪ d͡z̪				t͡ʃ d͡ʒ		
Sib. fri.			s̺ z̺				

Modern examples include Veronese Italian, with dental affricates lacking corresponding fricatives, and standard Serbo-Croatian, with a gap at [ɕ] (except for Montenegrin, which fills that gap):

(19) Veronese regional standard Italian (Avesani et al. 2017)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠				
Sib. fri.			s̠ z̠				

(20) Serbo-Croatian (Morén 2006)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.	t͡s̪			t͡ɕ d͡ʑ		t͡ṣ d͡ẓ	
Sib. fri.	s̪ z̪					ṣ ẓ	

It is thus a strong universal tendency rather than an absolute universal. Despite not being a perfect universal, this tendency makes the prediction that inventories such as Old French, Veronese and Serbo-Croatian are both rare (as already shown above, recall the 3.88% frequency in the SibInv) and diachronically unstable. It is indeed the case that 1) Old French affricates became fricatives due to context-free deaffrication in Middle French, solving the 'problem' of fricativeless affricates; 2) Veronese and, more generally, Veneto dialects variably deaffricate and optionally merge the dental sibilants with the retracted alveolar sibilant fricatives; 3) The Montenegrin variety of Serbo-Croatian filled the gap with [ɕ] from /sj/, while Croatian merges [t͡ɕ] to [t͡ʃ] to eliminate the alveopalatal PoA (Ćavar 2011). Sibilant affricates without corresponding fricatives are thus possible, although rare and diachronically short-lived.

It is thereby also predicted that a language without any sibilant can get sibilant fricatives by means of stop palatalisation, with an unstable but possible intermediary step in which only [t͡ɕ] exists in the inventory (as exemplified in 3.3.1.1).

3.2.1.5 Other significant correlations

The correlation between voicing distinctions and number of PoAs discovered above is not the only statistically relevant phenomenon that can be observed in sibilant inventories. In the SibInv database, the following correlations have also been noted:

- **PoA and voicing**: Besides the number of PoAs for sibilants in general, the number of PoAs for sibilant fricatives also correlates with voicing (t = -4.983, $p < 0.001$), as well as the number of PoAs for sibilant affricates (t = -7.155, $p < 0.001$). Sibilant inventories with more PoAs for affricates are thus more likely to have voicing contrasts than those with more PoAs for fricatives. This might be linked to the explanation given above (3.2.1.3) that voicing distinctions are likely to come about in 'voiceless' inventories by means of the palatalisation of voiced stops. More affricates (which most often originate in stop palatalisation) thus means larger probabilities that a voiced stop has been palatalised and therewith that voicing has been introduced into the sibinv.

- **PoA, fricative PoAs and affricate PoAs**: The number of PoAs for sibilant fricatives is a better predictor (F = 332.637, $r = 0.915$, $p < 0.001$) of the number of PoAs for sibilants *überhaupt* than the number of PoAs for sibilant affricates (F = 34.810, $r = 0.586$, $p < 0.001$). This corresponds to the typological observation that sibilant affricate tend to align the articulation of their sibilantic part to that of sibilant fricatives rather than the other way around (see 4.7 for further details).

- **Family and voicing**: The factors major language family (F = 2.148, $p < 0.001$) and, to a slightly larger extent, minor language family (F = 2.584, $p < 0.001$) allow to predict if a sibilant inventory has voicing distinctions. This is predictably a consequence of language conservatism: for example, Proto-Balto-Slavic already had voicing distinctions, so that all its Baltic (N = 3) as well as Slavic (N = 9) descendants in the database have voicing distinctions.

- **Family and PoA**: The same holds for the prediction of the number of PoAs for sibilants based on language family (major family: F = 3.000, $p < 0.001$; minor family: F = 2.098, $p < 0.001$), the number of PoAs for sibilant fricatives (major family: F = 3.056, $p < 0.001$; minor family: F = 2.155, $p < 0.001$) and the number of PoAs for sibilant affricates (major family: F = 2.593, $p < 0.001$; minor family: F = 2.892, $p < 0.001$).

- **IE and voicing**: Besides the almost significant difference between Indo-European and Non-IE with respect to the number of PoAs, mentioned in 3.1.2, there is a significant correlation between being Indo-European and having voicing contrasts in sibilants ($t = -3.853$, $p < 0.001$).

Additionally, the following correlation can be observed in the UPSID (a correlation that could not be tested in the SibInv because this database lacks information about the total number of segments of its languages):

- **Number of segments and presence of sibilants**: Languages without sibilants are very likely to have few phonemes. The 36 languages without sibilants in the UPSID have an average of 21.25 segments overall in their phoneme inventory, while the 415 languages with sibilants have 31.81 phonemes on average. This difference is highly significant ($t = -5.424$, $p < 0.001$).

3.2.2 Types of sibilant inventories

The diverse sibilant inventories of the world's languages are classified here according to a hierarchy of three parameters: 1) the total number of places of articulation for sibilants, 2) the precise places of articulation and 3) the presence of fricatives vs. affricates. A fourth distinction could be made based on the presence of voicing distinctions, but this increases considerably the number of observed sibilant inventories without adding much relevant information. In fact, voicing distinctions always occur in pairs with a voiceless and a voiced member (see 3.2.3.3), and whether or not a sibilant inventory allows for voicing distinctions does not directly influence the rest of the inventory (e.g. how many or where PoA contrasts will be present). This is not the case for affricates: an inventory like Old French detailed above is a 3s precisely because of its affricates and would be a 1s without its affricates, but its voiced segments do not change anything to the inventory. Contrarily to affricates, voiced sibilants can be present or absent in a predictable way (all voiced sibilants have a voiceless counterpart, but some affricates can have no fricative counterpart). For example, a 2s with only sibilant fricatives can look like one of these:

(21)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.							
Sib. fri.		s			ʃ		

(22)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.							
Sib. fri.		s z			ʃ		

(23)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.							
Sib. fri.		s			ʃ ʒ		

(24)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.							
Sib. fri.		s z			ʃ ʒ		

These inventories correspond to the four logically possible inventories given these parameters (assuming that a voiced sibilant implies the presence of its voiceless counterpart, as shown in 3.2.1.3): one without voicing distinctions, one with voicing at the alveolar place, one with voicing at the palatoalveolar place and one with voicing at both places. All four are attested (e.g. Irish, Daju, Chalcotongo Mixtec[5] and French[6]), and more generally, whatever the number of PoAs is, the different voicing possibilities are attested without that there be any universal implication of the kind e.g. 'If a language has [ʒ], it has [z]'. Languages with [ʒ] and no [z] are in fact less frequent than the opposite pattern, but are not unattested; their rarity, in this case, is a consequence of the markedness of postalveolar sibilants and of that of voiced sibilants ([ʒ] having thus twice the markedness of [z]). In order to keep the number of attested inventories reasonable within this section, and since adding voicing distinctions does not add much to our un-

[5] If one omits the palatoalveolar affricate present in the language.
[6] In the analysis according to which French has no phonemic affricates (see 6.3).

derstanding of the typology, voicing distinctions will no longer be considered in this section.

In the next subsubsections, all attested sibinvs will be classified according to the hierarchy of three parameters detailed above. For each number of PoAs for sibilants (hereafter called *class*), the possibilities wrt. each precise PoA are illustrated (hereafter called *type* of sibinv), together with the possible presence of affricates (classifying each *type* into *subtypes*). For example, Irish exemplified in (21) belongs to the class of 2s, the type of 2s with a (denti-)alveolar - palatoalveolar contrast, and the subtype of 2s with a (denti-)alveolar - palatoalveolar contrast and no phonemic affricates.

3.2.2.1 0-sibilant inventories

For obvious reasons, 0s can only belong to one class, one type and one subtype (i.e. inventories with 0 PoAs, no precise nature for this inexistent PoA, and no affricates). They are per definition the inventories that have no sibilant phonemes, but as will become clear below, they can either have no sibilant allophone at all (like Hawaiian) or allow certain segment combinations to produce sibilants allophonically (like Gilbertese).

Languages without sibilants are very rare. In the UPSID (Maddieson 1984), 7.98% of the languages (36/451) have no sibilant at all, neither fricative nor affricate. The number of these sibilantless languages is very close to the number of languages in the same sample without any fricative (31). In fact, only 10 languages (i.e. 2.22% of the languages in the UPSID) have a fricative but no sibilant (which can then be [h] or [f∼ɸ], but no other fricative). One can thus infer from this that 1) languages are very likely (92.02%) to have at least one sibilant, 2) languages that have fricatives are even more likely (97.62%) to have sibilants (as confirmed by the WALS, Maddieson 2013*b*). This is formulated as the following statistical universal (Nartey 1979: 4, cited in Hyman 2008: 115), #767 in the Konstanz Universals Archive (Plank and Filimonova 2000):

> If a language has only one primary fricative its primary allophone is most likely to be /s/.

This universal is not unviolated (as shown by the 10 UPSID languages with fricatives but no sibilant) but a strong statistical tendency instead.

Importantly, within the small number of languages without sibilant phonemes, a certain number still possesses sibilants as allophones (e.g. Rotokas, Robinson 2006: 207 and Gilbertese, Blevins and Harrison 1999: 206), something which leads to the idea that *really sibilantless* languages might be even rarer. Both in the case of Rotokas and in that of Gilbertese, one sibilant fricative exists as a surface realisation of the underlying sequence /ti/, thus being the result of stop palatalisation. In my database, only 3 languages (Australian languages, Hawaiian and Tuamotuan) are known with certainty to be languages without sibilant allophones (even if stop palatalisation is reported sporadically in Hawaiian, cfr. Schütz 1994: 70); 4 certainly have sibilant allophones (Maori, Rotokas, Gilbertese and Proto-Bantu); and for 4 other 1s, it is unknown.

In 0s, sibilants in loanwords are reanalysed as non-sibilant segments such as stops.

The most famous example for this phenomenon is the Hawaiian expression *Mele Kalikimaka* as a nativisation of the English expression 'Merry Christmas', in which the foreign English [s] is reanalysed as /k/ (Golston and Yang 2001; Gussenhoven and Jacobs 2017). However, in 0s that have sibilants as allophones, sibilants in loanwords are instead adapted and represented as the phonemic combination that yields the allophonic sibilant in the language: for example, the English word 'Gilberts' has been adapted in Gilbertese as *Kiribati* (with the pronunciation -[bæs̪] and the grapheme <ti>). In Rotokas however, the allophonic sibilant is usually written <s(i)> (as in the name *Rotokas* itself, underlyingly /rotokati/), although it could very much be written <ti> like in Gilbertese.

Languages without sibilants are in fact especially interesting, because by being most often fricativeless in general, they demonstrate that sibilants are among the least marked fricative sounds in natural languages (perhaps the least marked, if it were not for /h/; cfr. e.g. Tagalog, Qualla Cherokee or Guaraní with only /h/ and sibilants). This has been explained as a consequence of their perceptual saliency (see Flemming 2008*a*: 126 and Clements 2009: 45) or as an articulatorily motivated phenomenon (Everett 2018: 133). In any case, this makes the interesting prediction that sibilants could only appear in a 0s from a very restricted set of consonants: [t] in /tj/ (A.1), [j] (A.4) or [h] in /hj/ (A.3), but not from other fricatives such as [θ], since their presence implies that of sibilant fricatives. It also predicts that the disappearance of sibilants in a language must yield a language without any fricative (except for [h]), and so that it can either be by means of debuccalisation (C.4) or stopping (i.e. transformation of sibilants into stops, C.5). These predictions are borne out by the observation of sound shifts, as detailed in 3.3.1.1.

Another important observation about sibilantless languages is that they tendentially occur in isolated regions (e.g. the Amazone forest, or islands in the Pacific Ocean) or in *clusters* of languages with 0s, but almost never in the proximity of/in contact with languages that have sibilants. Maddieson (2013*a*) stresses this isolated/clustering behaviour wrt. the lack of fricatives in general, even if the same can be said for the correlated lack of sibilants:

> The great majority of [fricativeless languages] are in Australia, with other notable clusters in New Guinea and in the interior of South America. Outside of these areas there are only a few sporadic examples of languages without fricatives, such as Gilbertese and Hawaiian (both Austronesian), the Nilo-Saharan languages Dinka and Lango, spoken in the Sudan and Uganda respectively, and the one surviving Great Andamanese language (also known as Pucikwar), as well as Aleut, as mentioned above.

The areality and borrowability of segments such as /f/ and /s/ (Stolz and Levkovych 2019; Grossman et al. 2019) can explain the scarcity of 0s in the less isolated parts of the world.

3.2.2.2 1-sibilant inventories

Attested inventories with one PoA for sibilants belong to two types (either only retracted alveolar, or retracted alveolar and alveopalatal) and to four subtypes ((25)-(28)). The possibilities are (as said above, irrespectively from voicing distinctions):

(25)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.							
Sib. fri.			s̠				

(26)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.			t͡s̠				
Sib. fri.			s̠				

(27)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.				t͡ɕ			
Sib. fri.			s̠				

(28)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.			t͡s̠				
Sib. fri.			s̠	ɕ			

Examples of the respective subtypes are Classical Latin, Finnish, Older Elfdalian and Hollandic Dutch. The subtypes shown in (27) - (28) are still to be considered a 1s despite the two neighbouring PoAs because, as shown in 2.2.1, alveopalatal and retracted alveolar sibilants behave phonologically as one PoA, being distinguished by an additional distributedness contrast that does not pertain to place contrasts. For example, Hollandic Dutch /ɕ/ is accompanied with formant transitions (Ooijevaar 2011) that help to distinguish it from the otherwise acoustically close /s̠/. Since the typology proposed in chapter 5 does not integrate formant transitions etc. into the model, the subtypes in (27) - (28) are listed here as 1s inventories.

Out of the many possibilities (theoretically, there could be 27 subtypes of 1s with these 7 PoAs, irrespectively of voicing), only the four in (25) - (28) are attested. The reason for which sibinvs with only sibilant affricates cannot be diachronically stable (i.e. that they are 'doomed to deaffrication') is detailed in 3.2.1.4 above. The fact that 1s with sibilants at other places than the alveolar retracted/alveopalatal ones are not attested is due to another phenomenon: preference for the centre of the continuum (called 'centralisation' in this dissertation). This preference for the central retracted alveolar sibilants in 1s has been observed by Boersma and Hamann (2008: 220), who argue that "[i]f a language has only one category on the continuum, it tends to be in the centre". Among others, Vijūnas (2010) and Adams (1975), building further on the original observation of Martinet (1955), confirm typologically that this holds for all the languages with 1s of which they are aware, including e.g. Proto-Indo-European, Finnish, Old Korean, Samoan, Greenlandic, Canadian Inuit and Kurux (Vijūnas 2010: 49). The articulatory motivations for this universal of 1s are given in chapter 4 (4.3).

A characteristic trait distinguishing 1s from all other inventories (except those 0s which have allophonic sibilants, see 3.2.2.1) is that they adapt sibilants in loanwords as retracted alveolar sibilants, regardless of their original PoA. For example, Latin was a 1s like in (25) and adapted the postalveolar /ʃ/ in loanwords like Old Hebrew *Yeshua* ('Jesus', first name) as /s/ (*Ie*[s̠]*us*), the same way as it adapted the (denti-)alveolar /s/ in Etruscan *Rasna* as /s/ (*Ra*[s̠]*enna*) (Kokkelmans 2019*a*). A modern example of this phenomenon from the Iberian Peninsula is the observation that Castilian "speakers are transferring their Spanish sibilant inventory to Basque and they are reducing the historical three-way contrast in Basque to a single sibilant" (Muxika-Loitzate 2017: 6). Among other examples, Wolof also maps all sibilants to its sole /s̠/, so that French *croiser* ([kʁ̥waze], 'to cross') becomes *koro*[s̠]*e* and Arabic *shari'a* ([ʃariʕa] '(Sharia) law') becomes [s̠]*arija* (Ngom 2006: 105).

As seen above in the loanword adaptation pattern of 0s with allophonic sibilants, 1s with allophonic sibilants (e.g. Icelandic [ɕ] before [j]; cfr. Pétursson 1971: 210 who stresses that the sibilant in /sj/ is *palatalisé* and not *palatal*) make use of the acoustically closest sibilant allophone when adapting a loanword with a foreign sibilant. For example, Russian большевик 'Bolshevik' with [ṣ] is adapted in Icelandic as *bolsévíki*[7] with the allophone [ɕ], closer to the original postalveolar sibilant of the donor language in terms of CoG than the usual realisation of /s/ as [s̠]. The same thing holds for *klisja* 'cliché', *ansjósa* 'anchovy', *sjarmi* 'charme' etc.

The type of 1s and its different subtypes are summarised in fig. XLIV.

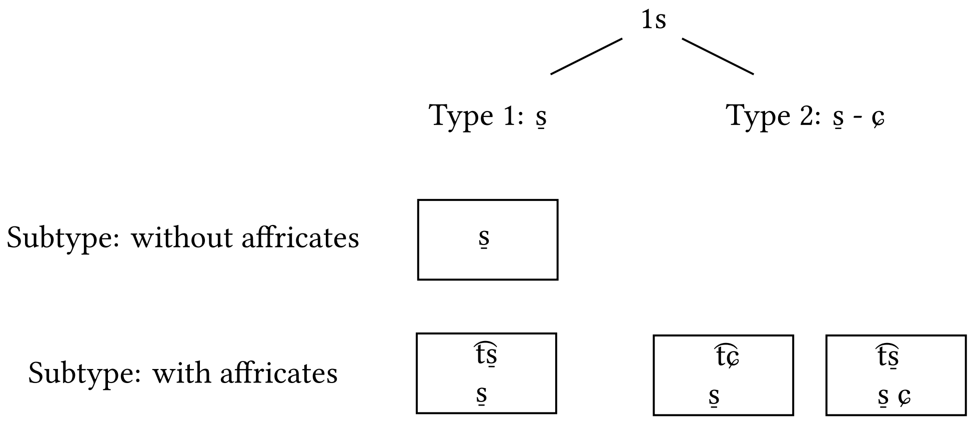


Figure XLIV: *Attested types and subtypes of 1s inventories.*

3.2.2.3 2-sibilant inventories

Transiting from the description of 1s to that of 2s, one must observe that there is one type of sibinv that might look like a 1s, but is in fact what one should consider a 'hidden 2s': sibilant inventories that have a phoneme /s/ pronounced as (denti-)alveolar [s]

[7] Note that <é> is pronounced [je] in Icelandic, so that *bolsévíki* = *bolsjevíki*.

and no /ʃ/ as a phoneme in native vocabulary. For example, Vijūnas (2010: 290) and Adams (1975) consider e.g. Mexican Spanish an exception to the centralisation universal by having a single phoneme /s/ articulated [s]. In fact, /ʃ/ is a phoneme in Mexican Spanish, although arguably a marginal phoneme because of its presence in loanwords (mainly from native Amerindian languages like Nahuatl). Regardless of the question of its status as a native or marginal/loan phoneme, /ʃ/ surfaces as [ʃ] in contexts that are not predictable as allophony (Lope 2004: 29). Another example is Younger Afrikaans, with a (denti-)alveolar /s/ (at least in the speech of young and/or white individuals, see Wissing et al. 2015) and a rare /ʃ/ in loanwords like *masjien* 'machine':

(29)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.					t͡ʃ		
Sib. fri.		s			ʃ		

Contrarily to 1s, which adapt any sibilant in loanwords as a retracted alveolar or alveopalatal one, these languages let loanwords surface as two different sibilants (alveolar /s/ or postalveolar /ʃ/); such inventories must thus be considered to be 2s, because they have 2 PoAs (regardless of whether one of these PoAs is frequently used or not).

There are four different types of 2s: one with an alveolar - palatoalveolar contrast, by far the most frequent type; one with a dental - retracted alveolar contrast; one with a dental - alveopalatal contrast; and one with a retracted alveolar - palatoalveolar contrast, as summarised in fig. XLV. Each of these four types has several attested subtypes, all described here below (again, without voicing distinctions). The first type of 2s is one

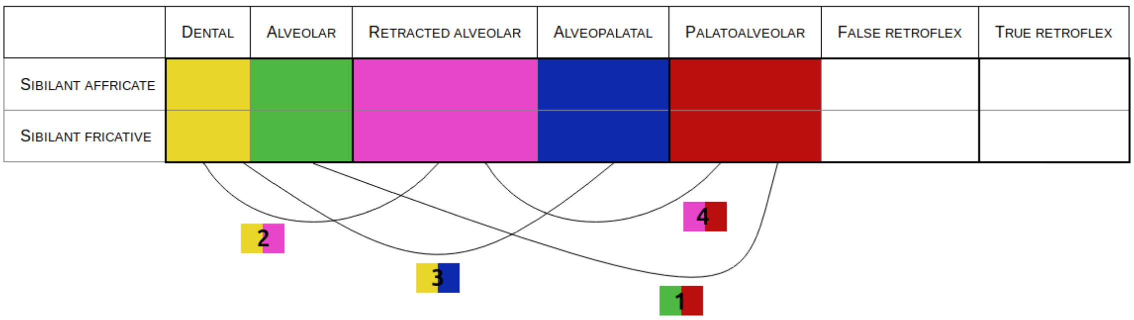


Figure XLV: *The three possible types of 2s inventories.*

with a contrast between (denti-)alveolar and palatoalveolar. The 2s illustrated in (29), Younger Afrikaans, is its subtype with only posterior affricates. The five other attested subtypes are exemplified in (30) - (34).

(30)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.							
Sib. fri.		s			ʃ		

(31)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.		t͡s			t͡ʃ		
Sib. fri.		s			ʃ		

Examples of these two sibinvs are Welsh ((30)) and Swiss German ((31)), respectively. These two inventories can also exist with a gap at /ʃ/. The first variant without phonemic /t͡s/ is illustrated by Aymara, in which allophonic alternations are observed between [t͡ʃ] and [ʃ] (Briggs 2007: 181f.), the former being deaffricated to the latter in consonant clusters.

(32)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.					t͡ʃ		
Sib. fri.		s					

The second variant with /t͡s/ is illustrated by Sandawe, in which it is not known (and at least not said in descriptive grammars) whether [ʃ] can surface allophonically (Hunziker et al. 2008; Steeman 2012).

(33)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.		t͡s			t͡ʃ		
Sib. fri.		s					

A sixth subtype is one with alveolar affricates, exemplified by Québec French.[8] It is rarer than the other subtypes because of the high frequency of palatalisation in natural languages, which results in postalveolar affricates being very frequent, as visible in fig. XLI; the frequency of palatoalveolar affricates makes it logically difficult to find languages with (denti-)alveolar affricates but without them.

(34)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.		t͡s					
Sib. fri.		s			ʃ		

The six attested subtypes just described are only a subset of the nine 2s inventories that are logically possible, given these two precise PoAs and an absence of voicing contrasts. The unattested subtypes are the three that have a gap at the place of /s/, showing

[8] Picard (1987: 46) shows that the alveolar affricates are phonemes, since [ti] and [di] occur in recent loanwords as well as across morpheme/word boundaries, and vowel deletion processes delete the phonetic trigger for affrication. On the opposite, postalveolar affricates are limited to non-native words.

that the only universal that can be observed in 2s is that there must be a (voiceless) alveolar sibilant fricative (as stated in (16)).

Some languages (e.g. Eastern Norwegian, White Hmong) are sometimes described as having a 2s with a contrast between (denti-)alveolar and plain retroflexes. ‘Sometimes’ is meant in the sense that the Eastern Norwegian postalveolar sibilant is e.g. labelled as ‘retroflex’ but written in the IPA with <ʃ> (Solhaug 2010: 17f.), labelled as ‘retroflex’ and written in the IPA with <ʂ> (Kristoffersen 2000: 22), or labelled as a simple palatoalveolar <ʃ> (Vanvik 1972: 146). This type only differs from the type with plain palatoalveolars to the extent that the main realisation of the palatoalveolars is described as apical (labelled as ‘retroflex’). Nevertheless, such types should not be considered as an independent type from the classical /s/ - /ʃ/ type, because the apicality of the postalveolar is a matter of allophony. In Eastern Norwegian, there is an alternation between plain palatoalveolars [ʃ], the historical result of palatalisation in clusters (Kristoffersen 2000: 23), and apical palatoalveolar i.e. plain retroflex [ʃ̺/ṣ], the historical result of apical assimilation in /rs/-clusters (Kokkelmans 2020*a*). Whether one should consider the contrast to be one between /s/ and /ʃ/ with allophonic [ṣ] or one between /s/ and /ṣ/ with allophonic [ʃ] is not a relevant question, since the postalveolar segment is one phoneme in any case and its different realisations are not contrasted (Solhaug 2010: 18). Thus, 2s with plain retroflexes do not need to be distinguished from 2s with plain palatoalveolar sibilants, the distinction being a matter of allophony.

The second type of 2s is one with dental and retracted alveolar sibilants. This type of 2s is less-dispersed, in the sense that its sibilant categories are acoustically near to each other when compared to the first type, and as a consequence of this, the contrast is often enhanced by distributedness distinctions (e.g. laminal dental vs. apical retracted alveolar; Alber and Kokkelmans in prep.). This type is exemplified by Romance dialects from the Swiss canton of Valais like those of Savièse and Évolène (Favre and Balet 1960; Elmiger et al. 2013):

(35)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.	t͡s̪		t͡s̠				
Sib. fri.	s̪		s̠				

The retracted alveolar sibilant fricatives are written <ch, j> as the result of the influence of French spelling conventions (Elmiger et al. 2013), but come from the historical retracted alveolar sibilant <s> in Latin and are still pronounced as such by most speakers (excluding the interference of the standard French inventory). This type of 2s is also attested with gaps at the dental fricatives or at the dental affricates in dialects from Veneto and Trentino in Italy (Marcato and Ursini 1998; Alber and Kokkelmans in prep.). As it is a rare type, subtypes with a gap at the retracted alveolar affricate, at both affricates or at the retracted alveolar fricative are not attested, something which does not however imply that they could not exist.

The third type is one with dental and alveopalatal sibilants. This type is also less-dispersed, and is thus likely to be accompanied by a distributedness distinction, as is the case in Japanese (Li et al. 2009; Flemming 2018).

(36)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.	t͡s̪			t͡ɕ			
Sib. fri.	s̪			ɕ			

This type is attested with the two possible gaps at the affricates (Forest Nenets without [t͡s̪] and Younger South Korean without [t͡ɕ]; cfr. Kim 1999, 2004), but not with a gap at both affricates or with any gap at a fricative.

The fourth type is one with retracted alveolar and palatoalveolar sibilants. It is exemplified by Catalan (Vijūnas 2010):

(37)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.			t͡s̠		t͡ʃ		
Sib. fri.			s̠		ʃ		

Here also, the less-dispersed contrast is likely to imply distributedness distinctions (e.g. Catalan /s/ as tendentially apical; Recasens and Rodríguez 2017). This type is also attested with a gap at the retracted alveolar affricate (Faroese and Asturian).

This type of sibilant inventory with /s̠ - ʃ/ is difficult to tell apart from inventories with /s̠ - ɕ/ such as the previously described Hollandic Dutch inventory. The presence of phonological rules like regressive palatal assimilation (e.g. Faroese *telgja* [tɛʎt͡ʃa] 'to carve', Árnason 2011: 126) can constitute evidence for an alveopalatal articulation, but also for a laminal palatoalveolar [ʃ̻] similar to the Basque phoneme /x/, which is also involved in palatalisation processes. From a historical perspective, both /s̠ - ɕ/ and /s̠ - ʃ/ are intermediate diachronic steps from a 1s with /s̠/ to a well-dispersed 2s with /s - ʃ/ (e.g. Irish and English), where /s̠/ gets palatalised in certain contexts and its output /ɕ/ is then retracted and depalatalised to /ʃ/. The different types of 2s and their different subtypes are summarised in fig. XLVI.

3.2.2.4 3-sibilant inventories

There are two types of 3s: one in which the PoA of the central sibilant has an alveopalatal articulation and the most posterior sibilant is a plain retroflex, and one in which the middle PoA has a retracted alveolar apical articulation and the most posterior sibilant is a laminal palatoalveolar. The first type is exemplified by standard Mandarin Chinese, the second by Basque. These two types will be referred to with the denominations 'retroflex 3s' and 'non-retroflex 3s', respectively.

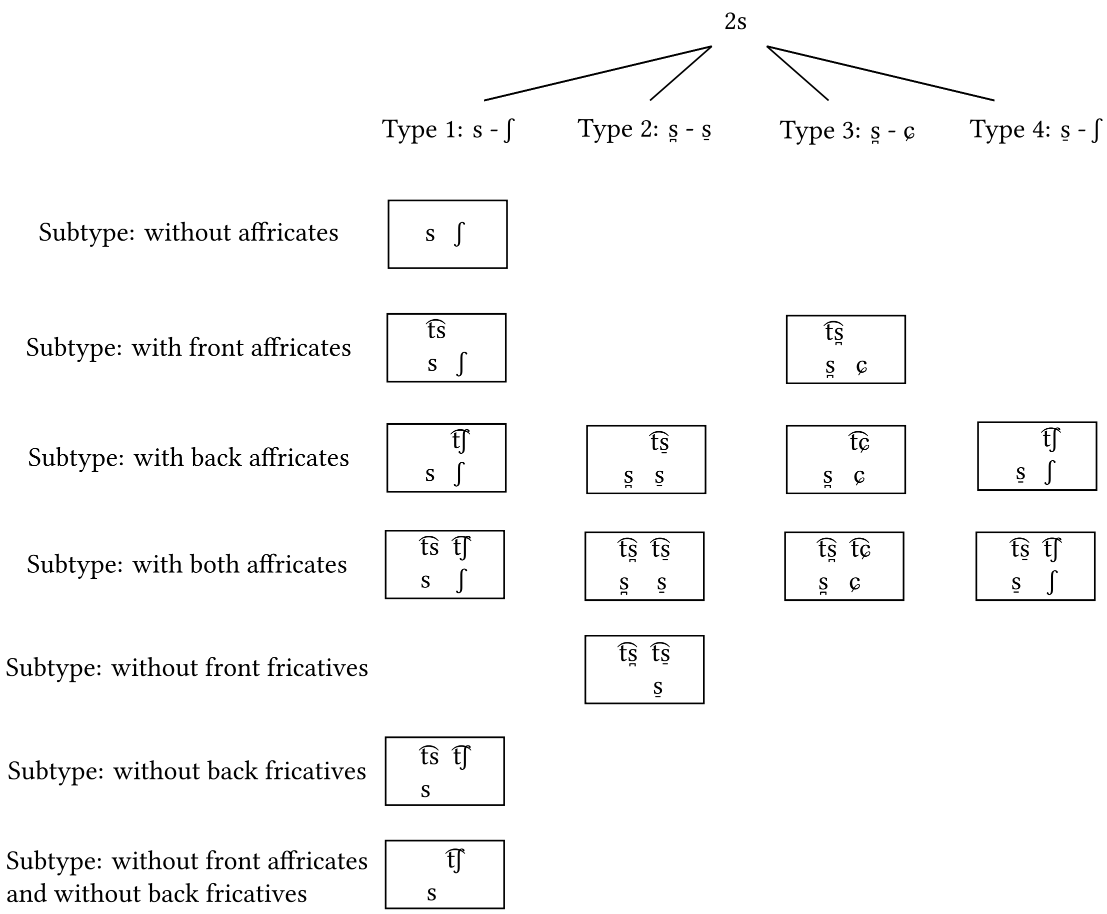


Figure XLVI: *Attested types and subtypes of 2s inventories.*

(38)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.	t͡s̪			t͡ɕ		t͡ṣ	
Sib. fri.	s̪			ɕ		ṣ	

(39)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.	t͡s̪		t͡s̠̺		t͡ʃ̻		
Sib. fri.	s̪		s̠̺		ʃ̻		

Contrarily to 2s, their front sibilant is consistently dental rather than variably (denti-)alveolar (Puppel et al. 1977 as cited in Ladefoged and Maddieson 1996: 154; Muxika-Loitzate 2019):

> The Polish [JK: a retroflex 3s] sounds s̪ and z̪ belong to dentalized sounds, i.e. those that are articulated in the alveolar region but with the blade of the tongue being very close to the inner side of the upper front teeth. Thus, the

> hissing effect is very strong. However, the English counterparts are articulated more in the purely alveolar region. Thus, in English, the tongue is more retracted for the articulation of these sounds.
>
> Basque has an apico-alveolar /s̺/, a *dental* /s̪/ and a prepalatal sibilant /ʃ/ (...). (my emphasis)

Also contrarily to the 2s inventories, the contrasts in 3s always imply distributedness distinctions (laminal vs. apical). Either the middle sibilant is laminal and the back sibilant is apical (retroflex type), or the other way around (non-retroflex type).

Turning to their respective subtypes, the retroflex type is attested with gaps at 6 places/manners: no gap at all (Polish), a gap at the alveopalatal fricative (standard Serbo-Croatian), a gap at the alveopalatal affricate (Beijing Mandarin Chinese), a gap at the apical palatoalveolar affricate (Russian), gaps at dental and apical palatoalveolar affricates (Sanskrit) and gaps at all affricates (Swedish).

The non-retroflex type is also attested with gaps at 6 places/manners: no gap at all (standard Basque), a gap at the retracted alveolar affricate (Middle High German), gaps at the dental and palatoalveolar fricatives as well as at the retracted alveolar affricate (Old French), gaps at the dental fricative and the retracted alveolar affricate (Old Spanish), gaps at the dental and retracted alveolar affricates (Mirandese), and gaps at all affricates (Middle French). Some Southern Bantu languages have been described as languages having whistled sibilants as some special kind of labialised segments, but these whistled sibilants have in fact been shown in the meanwhile to be apical retracted alveolar sibilants instead, with a periodic character obtained by whistling as a means of perceptual enhancement (Lee-Kim et al. 2014; Shosted 2011). They have thus been classified accordingly as 3s of the non-retroflex type.

The different types of 3s and their different subtypes are summarised in fig. XLVII. To summarise, 3s inventories thus consist of an invariably dental sibilant, together with retracted alveolar and palatoalveolar sibilants (non-retroflex type) or alveopalatal and plain retroflex sibilants (retroflex type). Distributedness distinguishes one posterior segment from the other, and the difference between the retroflex and non-retroflex types resides in which segment is apical and which is laminal.

3.2.2.5 4-sibilant inventories

Inventories with 4 PoAs for sibilants are exemplified by Toda:

(40) Toda (Emeneau 1984; Ladefoged and Maddieson 1996; Boersma and Hamann 2008)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.					t͡ʃ d͡ʒ		
Sib. fri.	s̪		s̠		ʃ ʒ		s̢̠ z̢̠

This Dravidian language is described by Ladefoged and Maddieson (1996: 156-159) as a language with a laminal dental sibilant, an apical retracted alveolar sibilant, a lami-

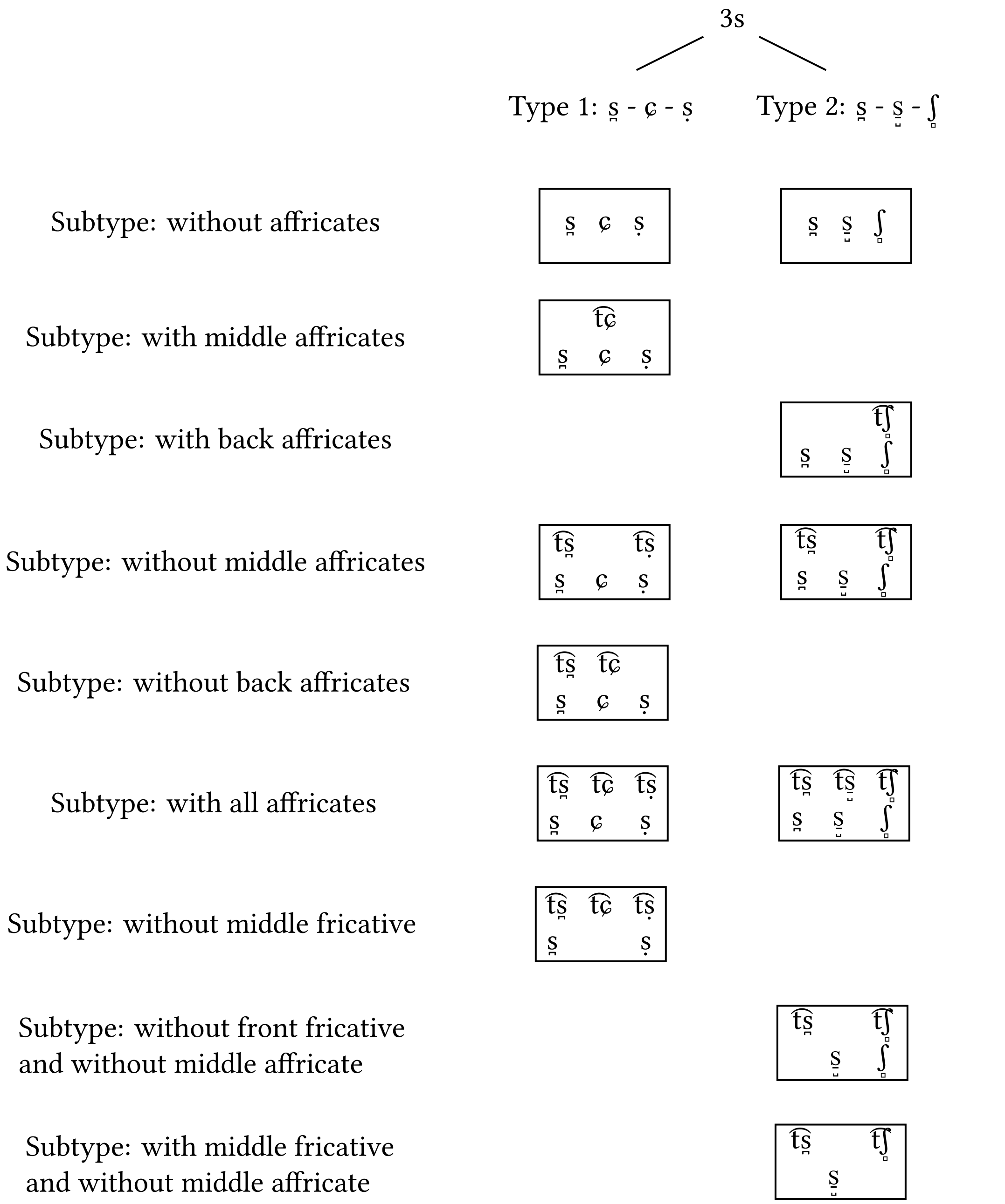


Figure XLVII: *Attested types and subtypes of 3s inventories.*

nal palatoalveolar and a subapical retroflex. Like in 3s, distributedness distinctions are obligatorily exploited to maximise the perceptual distinctiveness of the contrasts. Two other subtypes are attested: the Northwest Caucasian language Ubykh had the same PoAs as Toda with a gapless voicing and affricate series and literary Abzhywa Abkhaz is reported to have a gap at the retracted alveolar fricatives. In fact, a very large number of subtypes should be possible considering the many possible gaps, but 4s are so rare in general that only these three subtypes are attested. The attested types and subtypes of 4s are summarised in fig. XLVIII.

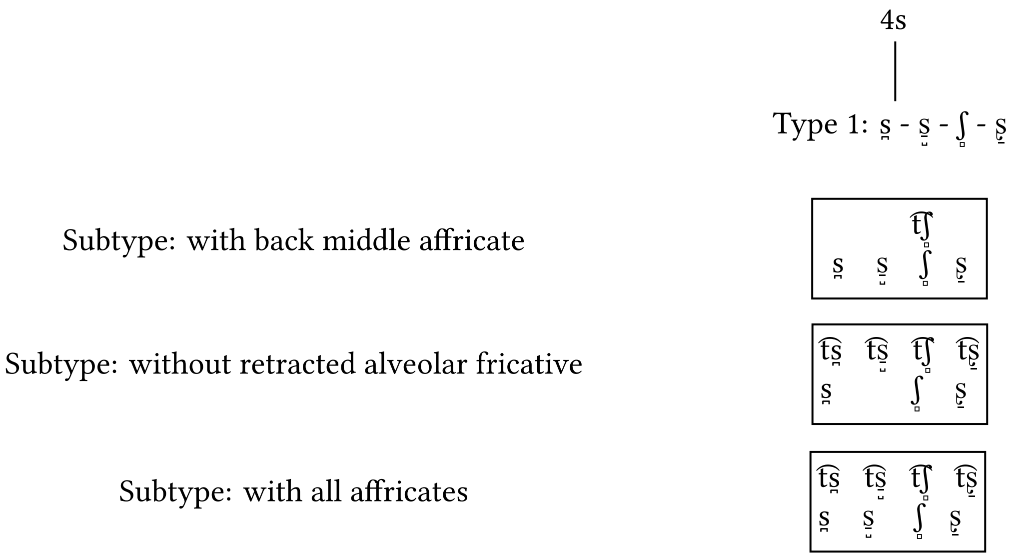


Figure XLVIII: *Attested types and subtypes of 4s inventories.*

3.2.3 Gaps

Gaps in phonological inventories can be accidental (i.e. due to historical 'accidents', synchronically unmotivated) or systemic (i.e. synchronically and phonologically motivated; Kirby and Yu 2007). Crystal (2008: 205) provides the following definition of *gap*:

> **gap** (*n.*) (1) A term used in LINGUISTICS to refer to the absence of a linguistic UNIT at a place in a pattern of relationships where one might have been expected; also called an **accidental gap** or **hole in the pattern** (...). An example of a **phonological gap** would be seen in a language where the PHONEMIC CONTRASTS /p/, /b/ and /t/, /d/ were not matched by a corresponding velar pair, only /k/ being found.

The example he provides is in fact one of a systemic gap, because [g] is universally marked (Boersma 1998: 226f.). An example of the opposite, an accidental gap, is the absence of [ɕe]-sequences in conservative (19^{th} century) Tokyo Japanese, which is due to the merely historical reason that palatalisation did not occur in the combination /se/ while it did in /si/ (Itō and Mester 2003; Pintér 2015: 124ff.).

With respect to sibilant inventory structure, one can posit e.g. that the lack of /d͡z/ in Northern standard German is an accidental gap: its voiceless counterpart /t͡s/ appeared in the sibinv with the reanalysis of aspiration ([tʰ]) as stridency ([tˢʰ]) in Old High German, but there was no aspirated voiced [dʰ] to assibilate at all (Kokkelmans 2019*b*). It is not clear, however, if sibilant inventories can have systemic gaps: Żygis (2008) and Żygis, Fuchs and Koenig (2012) seem to consider that it is possible with voiced affricates,[9] while

[9] To be precise, they do not use the terms 'accidental' or 'systemic' but do compare voiced affricate gaps to the /g/-gap. In a sense, their articulatory account of the difficulty in the production of voiced affricates corresponds to the presence of a constraint against voicing and one against affricates in the OT sibilant typology I propose in chapter 5.

I claim that those gaps are accidental and in fact likely to occur, considering certain universals of phoneme inventories (see 5.4.2 for an argumentation in favour of accidental gaps).

The following subsubsections examine generalisations about what seems and seems not to be a possible gap in a sibilant inventory.

3.2.3.1 Fricative gaps

When considering the precise place of articulation of the sibilant affricates that have no corresponding fricative, the data presented in the previous subsection (3.2.2) reveals that four sibilant fricatives were able to constitute a gap in the inventory: [s̪] (Veronese regional standard Italian), [s̠] (Literary Abzhywa Abkhaz), [ɕ] (standard Serbo-Croatian) and [ʃ] (Aymara). The three sibilant fricatives that never constitute a gap in my database are alveolar [s], plain retroflex [ṣ] and subapical retroflex [ṣ̠]. Considering the rarity of both subapical and plain retroflex sibilants in the world's languages (only 7% in fig. XLI), it is very much possible that this observed absence of gaps is simply the consequence of the low number of sibilant inventories that have these PoAs *überhaupt*, and would thus be able to have a gap at these PoAs. In other words, finding an inventory with subapical or plain retroflex sibilants is already so unlikely that finding one with a gap at those PoAs is even more unlikely.

On the other hand, the unattested gap at [s] (which has 86% frequency in fig. XLI, although not distinguishing dental, alveolar and retracted alveolar) is highly likely to be motivated by articulatory and/or perceptual constraints in grammar rather than to be a mere accident. There is evidence for this in the diachronic evolution of sibinvs, which allows for accidental gaps at e.g. [d͡z] like in the German example above but does not allow for a gap at [s] or more generally, the absence of a voiceless anterior sibilant fricative whenever the conditions are united to create such an impossible gap. For example, Turkmen fronted /s, z/ to [θ, ð], leaving /ʃ/, /ʒ/, /t͡ʃ/ as sole sibilants, but Russian loanwords with [s, z] fill the gap by being realised as [s, z] (Clark 1998: 37; Hoey 2013). Similarly, rural *ceceo* dialects of Southern Spain have merged the historical dental sibilants with the historical retracted alveolar sibilants (like in *seseo* dialects in general), and then fronted that sibilant to interdental fricatives. Since this would lead to an inventory with /ʃ/ as a sole sibilant, the interdental fricative cannot maintain itself and it is thus reported that *ceceo* speakers "alternate between /θ/ and /s/ in free variation" (Moore 2005: 4), without using one or the other consistently (Dalbor 1980). This impossibility of gaps at /s/ (or more generally, of having gaps at all anterior sibilant fricatives simultaneously) confirms the universal stated in (16):

> If a language has sibilants, it has at least one voiceless anterior sibilant fricative.

When considering the relative position of the missing sibilant fricative (i.e. whether a gap occurs at the frontmost sibilant, backmost, middle etc.), gaps are possible at any place except one: at the front of the sibinv (Old Spanish), at the middle front (Literary Abzhywa Abkhaz), at the middle (Serbo-Croatian), at the back (Old French). The only unattested gap position is the middle back position in a 4s, which is most probably unattested accidentally rather than for some reason (as explained above).

3.2.3.2 Affricate gaps

Gaps in sibilant affricates are possible at all PoAs: [t͡s̠] (standard Korean), [t͡s] (Middle French), [t͡s̠] (Faroese), [t͡ɕ] (Swedish), [t͡ʃ] (Middle French), [t͡ʂ] (Russian) and [t͡s̠̺] (Toda). With respect to the relative position within the sibinv, all positions are attested. It thus becomes apparent that no principle governs the presence or absence of a gap at affricates. Some gaps might be more frequent than others, but this is a predictable consequence of the markedness and rarity of specific segments resulting from constraints applying to all sibilants of a certain kind (e.g. affricate or voiced), not constraints that apply to "one segment in particular but not another" (see 8.2.1.3).

3.2.3.3 Voicing gaps

Voicing gaps in the form of a missing voiced sibilant in a sibinv that has voicing distinctions can occur at any PoA. For example, many languages have a gap at /ʒ/ but do possess a phoneme /z/. The reverse is also possible, as witnessed by e.g. Chalcotongo Mixtec in the SibInv database (/s/ vs. /ʃ/ and /ʒ/, cfr. Macaulay 1996: 19; Botma and van 't Veer 2013: 47). Among the 3s, Russian exhibits a voicing gap in the middle (/ʑ/; Yanushevskaya and Bunčić 2015: 224) and at all voiced affricates; Burushaski (Jammu and Kashmir) exhibits a voicing gap at the palatal fricative and at the dental and retroflex affricates; Sanskrit exhibited voicing gaps at all fricatives; standard Serbo-Croatian only at /d͡z/; Old Spanish at /d͡z̠/ and /d͡ʒ̠/, etc. It becomes apparent that no implicational hierarchy is to be found in voicing gaps. The only striking regularity is that gaps at /ʒ/ are more frequent (19 occurrences in the SibInv) than at the *anterior* /z/ (4 in the SibInv).

When it comes to a voicing gap in the sense of a voiced sibilant lacking a voiceless counterpart at the same PoA and with the same manner, it is unclear whether the universal formulated in 3.2.1.3 (stating that the presence of a voiced sibilant implies that of its voiceless counterpart) holds in all of the world's languages. While there is little doubt that it is a valid generalisation if intended as a statistical tendency (e.g. 0% of the inventories in the SibInv database having such a gap), two cases (Arabic and Ancient Greek) seem to challenge this universal tendency (cfr. Berns 2013: 68f.).

Historical **g* in Arabic surfaces as [d͡ʒ] in several modern varieties of Arabic, among others in Egypt (Hary 1996; Woidich and Zack 2009), the Levant and Modern Standard Arabic (Al-Deaibes 2016: 42), but not in Sudan, in Libya and the Maghreb (Elhija and Davis 2016: 89f.) nor in Oman (As-Sammer 2011). The gap at /g/ created by its historical context-independent affrication constitutes a systemic gap (Boersma 1998).

In many of the varieties with /d͡ʒ/, /t͡ʃ/ does also exist (mainly as a result of the affrication of **k*; Ferguson 1969: 116) but a few varieties with /d͡ʒ/ and without /t͡ʃ/ such as San'ani (Elhija and Davis 2016) or Beduin Arabic (Freeman 2014) do exist. The voicing gap universal thus holds only as a statistical tendency in the languages of the world and within Arabic varieties themselves, since languages with /d͡ʒ/ and without /t͡ʃ/ are rare both in the world and within Arabic dialects. The gap at /t͡ʃ/ indeed tends to be eliminated by affricating /k/, by devoicing /d͡ʒ/ or deaffricating /d͡ʒ/ (Freeman 2014). Additionally to the tendency of varieties with /d͡ʒ/ and no /t͡ʃ/ to repair this gap, Elhija and Davis (2016) show that in San'ani Arabic, derived [t͡ʃ] sequences behave as a single segment, based on evidence from phonotactic restrictions, stress patterns and loanwords. They demonstrate that in a variety with no historical affricate /t͡ʃ/, "there is systemic pressure for sequences of [tʃ] to pattern as a single segment; given that the voiced affricate [d͡ʒ] occurs, [t͡ʃ] would constitute its voiceless counterpart" (Elhija and Davis 2016: 94). In dialects that do not have the monosegmental /d͡ʒ/, they show inversely that [tʃ] behaves as an underlying cluster /t+ʃ/ (e.g. by allowing epenthesis in the middle of the cluster). They conclude that Arabic dialects either tend towards an inventory without any affricate, or towards a sibinv with /d͡ʒ/ *and* /t͡ʃ/. In the light of these observations, it is thus unclear whether a typology of sibilant inventories should predict all possible voicing gaps at any voiceless sibilant only because of these varieties with an apparent gap at /t͡ʃ/ (see 6.3).

Ancient Greek is known to have had the letter *dzeta*, representing (at least originally) a voiced affricate [d͡z̠] originating e.g. in the palatalisation of /dj/. The absence of a letter for its voiceless counterpart might suggest that there was no /t͡s̠/ in Ancient Greek; however, /t͡s̠/ also appeared by means of palatalisation in /tj/ (Teodorsson 1979; Vijūnas 2010: 41). The lack of a graphemic representation for the voiceless affricate is ascribed by Teodorsson (1979: 328) to Boeotian influence (a variety of Greek that had a geminated /t:/ for historical /tj/-sequences). After the deaffrication of both affricates, the earlier /t͡s̠/ was represented with the pre-existing *sigma* for /s̠/ in the Ionian orthography, while keeping a distinct letter *dzeta* for /z̠/. This apparent exception to the voicing universal is thus in fact no exception, but looks like it because it is masked by the orthography.

3.2.4 Summary

To summarise the main findings of the overview of attested sibilant inventories presented in this section, it has become apparent that:

- Languages have between 0 and 4 PoAs for sibilants.
- There are two kinds of 0s: one in which sibilants do not exist, neither as phoneme nor as allophone; and one in which there is no sibilant phoneme but sibilants are allowed as allophones. In the former case, foreign sibilants in loanwords surface as non-sibilants; in the latter, sibilants in loanwords are mapped to the underlying form that yields a sibilant allophone on the surface. In languages with 0s and sibilant allophones, the allophones behave like in a 1s (i.e. they are retracted alveolar or alveopalatal fricatives/affricates).
- Languages with 1s have retracted alveolar sibilants (and, optionally, alveopalatal sibilants).
- Languages with 2s are well-dispersed ([s] - [ʃ]) or less-dispersed ([s̪] - [s̠/ɕ] or [s̠] - [ʃ]).

- Languages with 3s are either of the retroflex ([s̪] - [ɕ] - [ṣ]) or the non-retroflex type ([s̪] - [s̠̺] - [ʃ̠̻]).
- Languages with 4s are of the non-retroflex type ([s̪] - [s̠̺] - [ʃ̠̻] - [ṣ̠]).
- Gaps for sibilant affricates or fricatives are always accidental, because they can seemingly occur at any place of articulation. It remains unclear whether gaps at /s/ can occur in well-dispersed 2s as well as 3s and 4s, since this pattern is unattested.
- All languages that possess sibilants have at least a dental [s̪], an alveolar [s] or a retracted alveolar [s̠].
- Gaps for voicing contrasts are possible at any PoA, and the presence of a voiced sibilant implies that of a voiceless sibilant at the same PoA and with the same manner (with the exception of some Arabic varieties, as explained above).

A summary of the attested inventories, taking into account PoA and the fricative - affricate distinction, is provided in table 2.

# PoAs	Example inventory	Fully-detailed 7 PoA version							Reduced to 4 PoAs
0	Hawaiian	–	–	–	–	–	–	–	_ _ _ _
1	Athens (standard) Greek	–	–	T/S	–	–	–	–	_ B _ _
	Hollandic Dutch, Older Elfdalian	–	–	T/S	T/S	–	–	–	
2	Xhosa	–	T/S	–	–	T/S	–	–	A _ C _
	Catalan	–	–	T/S	–	T/S	–	–	_ B C _
	Évolène dialect	T/S	–	T/S	–	–	–	–	A B _ _
	Japanese	T/S	–	–	T/S	–	–	–	
3	Standard Basque (Euskara Batua)	T/S	–	T/S	–	T/S	–	–	A B C _
	Polish	T/S	–	–	T/S	–	T/S	–	
4	Toda	T/S	–	T/S	–	T/S	–	T/S	A B C D

Table 2: *Summary of the attested sibilant inventories, not taking into account possible gaps and voicing distinctions.* T *stands for an affricate,* S *for a fricative; the 7 PoAs (4 PoAs) are: (dental, alveolar), (retracted alveolar, alveopalatal), (palatoalveolar, plain retroflex), (subapical retroflex).*

3.3 A diachronic typology of sibilant inventories

This section is structured into two subsections: 3.3.1 gives a non-exhaustive overview of the ways in which sibilant inventories change as a whole, and 3.3.2 lists sound shifts that provoke these changes as exhaustively as possible. The purpose of this section is mainly descriptive, which includes centralising existing literature on sibilant-related sound shifts and the ways in which these affect the structures of sibilant inventories.

3.3.1 Attested transitions between inventories

This section provides an non-exhaustive overview of the ways in which sibilant inventories have changed in the languages of the world, illustrating each change with one example. Considering the very large number of possible changes (especially in larger sibilant inventories), this part focuses on the most prototypical changes. It starts from the smallest sibilant inventories and advances towards the larger ones.

Importantly, changes take place *minimally*: rather than seeing sibilant inventories move abruptly from e.g. 3s to 1s (cfr. fig. LXXII in 5.2.2.2), they move in minimal steps according to the predictions made in chapter 5 (cfr. Alber 2001 et seq.).

3.3.1.1 From 0s to 1s and back: Birth and death of sibilants

The simplest sibilant inventory is obviously one without any sibilant, like Hawaiian (Gussenhoven and Jacobs 2017: 46):

(41) Hawaiian (Gussenhoven and Jacobs 2017; Schütz 1994)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.							
Sib. fri.							

In such inventories, sibilants could appear from several other sounds (e.g. /t/, /j/). The only appearance of a sibilant from a 0s of which I am aware is through stop palatalisation, described in A.1 in this section for Rotokas. As described in 3.2.2.1, it is predicted that only sibilantless languages with /t/ and /j/, with /h/ and /j/ or with /j/ can phonemicise sibilants out of these sounds. The Polynesian language Maori, related to Hawaiian, is an example of a 0s with allophonic sibilant affricates: underlying sequences of /tj/ and /ti/ surface as [t͡ɕ], and this could become a phoneme as soon as the sequence /tj/ or /ti/ reappears as such in the language.

(42) Hypothetical Maori in which the palatal allophone got phonemicised

	Dent.	Alve.	Retr.	Alve.	Pala.	Fals.	True.
Sibi.				t͡ɕ			
Sibi.							

The birth of sibilants happens through a diachronic stage in which a sibilant affricate exists without a corresponding sibilant fricative (e.g. [ɕ]). Yet, even before phonemicisation occurs, the initial palatal stop is observed to also have depalatalised allophones realised as plain retracted alveolar sibilants, as shown e.g. by the Papuan language Rotokas (Robinson 2006: 207) and the Micronesian language Gilbertese (Blevins and Harrison 1999), in which the allophone of /t/ before /i/ is a fricative /s/.[10] Since 0s in general and the birth of sibilants in particular are rare, it is uncertain whether all 0s that become 1s transit or not through a stage with a sibilant affricate, and if there is always allophony with a sibilant fricative. Considering the typologically observed markedness of affricates (cfr. 4.6), the change [c] > [t͡ɕ] > [ɕ] > [s̠] or [c] > [t͡ɕ] > [t͡s̠] > [s̠] (or the allophonic distribution of these segments) that must have occurred in some form in Rotokas is highly likely.

In any case, the transition from a 0s to a 1s leads to a sibilant inventory in which the only sibilant phoneme is a retracted alveolar [s̠] (which might or might not have affricated allophones):

(43) Proto-Polynesian (Elbert 1964; Rutter 2001)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.							
Sib. fri.			s̠				

Such an inventory can then undergo sibilant debuccalisation to [h], as occurred precisely from Proto-Polynesian to Hawaiian. This leads back to a 0s. The possible sound shifts leading to a 0s that are predicted (see 3.2.2.1) are debuccalisation to [h] (C.4) and stopping (C.5).

3.3.1.2 Transitions within a 1s

Rather than losing its only sibilant, a 1s with a single sibilant fricative (as e.g. Late West Germanic, which had the same sibilant as Proto-Polynesian) can phonemicise a new affricate, either by means of stop palatalisation like in Anglo-Frisian (from /ki, gi/, as in *church* vs. German *Kirche*) or by means of the reanalysis of a (denti-)alveolar stop as a strident stop, as in early Old High German:

(44) Late West Germanic (van der Hoek 2010)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.							
Sib. fri.			s̠				

[10] The same can be said of the endangered Amazonian language Wari', in which the affricate freely alternates with the fricative. They are transcribed as <tʃ, ʃ> by Everett and Kern (2002), although little is said about their concrete phonetic realisation. The affricate is represented interchangeably as both <ts> and <tʃ> throughout the book and the researchers are native speakers of English and thus predictably likely to perceive retracted alveolar sibilants as postalveolar (Adams 1975).

(45) Proto Anglo-Frisian (van der Hoek 2010: 47f.)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.				t͡ɕ d͡ʑ			
Sib. fri.			s̠				

(46) Old High German (Adams 1975)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.			t͡s̠				
Sib. fri.			s̠				

The modern outcomes of these affricates in English and German are a backed [t͡ʃ] (e.g. *cheese*) and a fronted [t͡s̪] (e.g. *Salz* 'salt'), respectively.

Depalatalisation can occur in the subtype with [t͡ɕ], yielding a retracted alveolar [t͡s̠].

Deaffrication to sibilant fricatives is possible with both [t͡s̠] and [t͡ɕ]; in the first case, one goes back to a 1s with one sibilant fricative ([t͡s̠] > [s̠]), and in the second case, one can either go back to a 1s with one sibilant fricative through a merger of [ɕ] to [s̠] or end up with a 2s with [ʃ] (see further below for 2s). Staying within the domain of 1s, one observes that changes from Late Proto-Finnic to modern Finnish exemplify the possibility of depalatalisation followed by deaffrication for singletons (singleton [t͡ɕ] > [t͡s̠] > [s̠], geminate [t͡ɕ] > [t͡s̠]; Pajusalu 2012; Koivulehto 1986: 293):

(47) Late Proto-Finnic (Pajusalu 2012; Koivulehto 1986: 293)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.				t͡ɕ			
Sib. fri.			s̠				

(48) Finnish (Adams 1975; Suomi et al. 2008: 27; Vijūnas 2010: 49)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.			t͡s̠				
Sib. fri.			s̠				

In Ancient Greek, this occurred regardless of gemination, leading back to a 1s with a single [s̠] (Teodorsson 1979; Vijūnas 2010: 41).

In a 1s with retracted alveolar fricatives and affricates, stop palatalisation can occur once more, yielding an inventory like Late Vulgar Latin (e.g. constan[t͡s̠]a 'constancy' and [t͡ɕ]aesar 'Caesar', except in Norman and Picard):

(49) Late Vulgar Latin (Pope 1934; Bateman 2007)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.	t͡s̪ d͡z̪			t͡ɕ d͡ʑ			
Sib. fri.			s̠				

In this case, the retracted alveolar affricates are fronted to avoid merging with the new alveopalatal sibilants, leading to a 2s (or, alternatively, they could simply merge to [t͡s̪]). This is also possible with the palatalisation of both /tj/ and /sj/, which happened from Old Japanese to modern Japanese and yields a 2s with symmetrical dental and alveopalatal sibilants:

(50) Japanese (Toda 2009)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.	t͡s̪ d͡z̪			t͡ɕ d͡ʑ			
Sib. fri.	s̪ z̪			ɕ ʑ			

The possibility in which only /sj/ would palatalise and transform the inventory into a 2s with /s - ʃ/ is not known to me. However, from Old Korean to modern standard Korean, the palatalisation of both /tj/ and /sj/ without having had an affricate at first yielded a 2s inventory like this:

(51) Older (South) Korean (Ko 2013; Cho 2016)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.				t͡ɕ			
Sib. fri.	s̪			ɕ			

Turning back to a 1s with only sibilant fricatives, it can develop voiced allophones e.g. intervocalically and phonemicise them, as happened from Late West Germanic illustrated in 3.3.1.2 to Middle Dutch, yielding:

(52) Middle Dutch

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.							
Sib. fri.			s̠ z̠				

This voicing can be undone again, as occurred from Middle Dutch to Afrikaans:

(53) Older Afrikaans (Wissing et al. 2015)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.							
Sib. fri.			s̠				

Voicing distinctions can likewise appear and disappear in any of the inventories described above and below (compare for example standard Mandarin Chinese and Polish, which have the same inventory with [(t͡)s̪ - (t͡)ɕ - (t͡)ṣ] except for voicing). It is therefore not relevant, if not not particularly interesting, to illustrate every inventory without voicing distinctions together with all possible inventories with one or several corresponding voiced sibilants in it; like in section 3.2.2, transitions from 'voiceless inventories' to 'voiced inventories' and vice-versa will not be detailed here.[11] Affrication and deaffrication, on the contrary, can change inventories radically and are therefore regarded in this subsection.

The transitions between the inventories described here are summarised in fig. XLIX. To summarise the generalisations that we can deduce from the observed transitions,

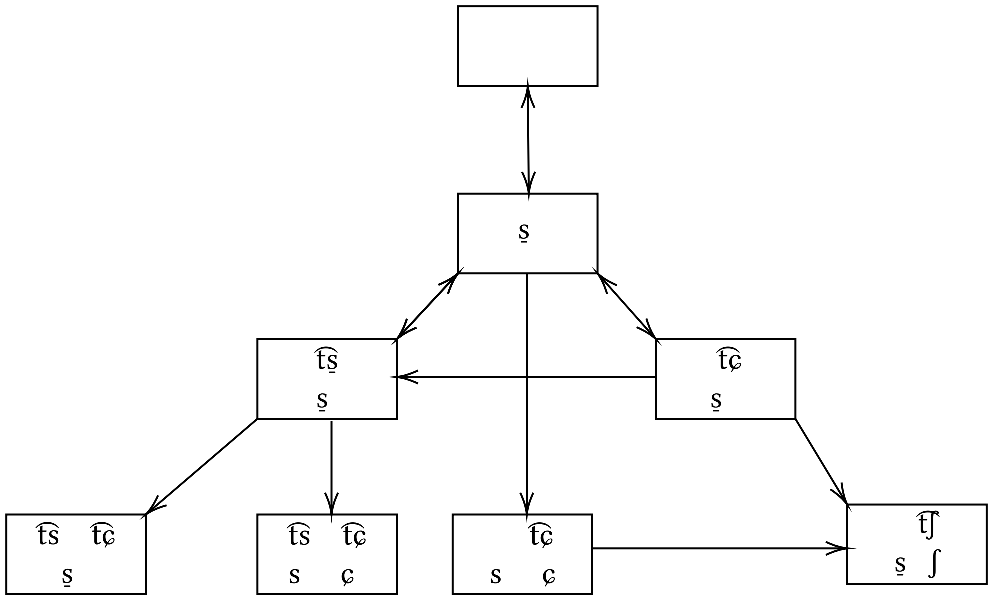


Figure XLIX: *Attested transitions between and within 0s and 1s inventories (with the lower row being 2s).*

sibilant inventories are targeted by sound shifts that can create as much as change or delete a sibilant from the inventory, and inventories 'react' to similar changes (recall e.g. the Late Latin affricate in CONSTANTIA being fronted to prevent auditory confusion with the new affricate in CAESAR in many Romance varieties). New sibilants can appear from non-sibilants as well as from sibilants (e.g. [t͡ɕ] from /ti/ in Maori and [z̠] from [s̠] in Middle Dutch), and sibilants can disappear by becoming non-sibilants (e.g. /h/ in Hawaiian) as well as by merging with other sibilants.

3.3.1.3 From 1s to 2s and back again

Under the influence of neighbouring 2s languages with a dispersed [s] - [ʃ] contrast, a 1s like Afrikaans can undergo what I call 'sibilant redistribution', i.e. the fronting of [s̠] to [s] and the optional backing of [s̠] in certain contexts to [ʃ]. In the speech of younger and higher educated Afrikaners, the inventory illustrated in 3.3.1.2 has become the following:

[11]The relevant cases in which a sound shift only applies to one of the two members of a voiceless - voiced sibilant pair are described below in 3.3.2 in any case, making voicing distinctions unnecessary in this subsection.

(54) Younger Afrikaans (van Wyk et al. 1989; Wissing et al. 2015)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.					t͡ʃ		
Sib. fri.		s			ʃ		

Crucially, [ʃ] in loanwords is realised as [ʃ], showing that there is an underlying postalveolar phoneme although it occurs marginally in native vocabulary (cfr. (29)).

Inversely, a 2s language can become a 1s under the influence of speakers of a 1s. This happened with standard French being learnt on Maurice Island by populations speaking (among others) Malagasy, leading to the merger of French palatoalveolars (e.g. *cheveu* 'hair') and alveolars to alveolar retracted (Mauritian Creole *seve* 'hair'; Gussenhoven and Jacobs 2017: 58).

(55) Older Mauritian Creole (Baker and Kriegel 2013: 253; Dufour et al. 2014)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.
Sib. aff.				t͡ɕ d͡ʑ			
Sib. fri.			s̠ z̠				

A 2s can also become a 1s by debuccalising the posterior sibilant to [h], something which happened from Early Proto-Finnic (or Balto-Finnic) to Late Proto-Finnic (Hyllested 2014: 11), yielding the inventory with only retracted alveolar sibilants presented in 3.3.1.2.

(56) Early Proto-Finnic (Pajusalu 2012; Koivulehto 1986: 293)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.							
Sib. fri.		s			ʃ		

Further transitions within 2s and towards 3s as well as 4s are possible, but are not detailed here because of the innumerable possibilities and attested transitions (especially considering that the number of possibilities increases with the number of PoAs).

3.3.2 Attested sibilant sound shifts

In this subsection, attested diachronic changes that create (= A), modify (= B) or delete (= C) sibilants are described, accompanied by evidence from languages that exemplify the relevant change (similarly to what Hall and Żygis 2010: 11-15 undertake). The subsection is divided in three subsubsections (A, B, C) based on whether sibilants appear, change or disappear, respectively. An additional distinction is made within the subsubsections between articulatorily motivated and perceptually motivated sound shifts.

Here and in the first subsubsection on assibilation (A.), I claim that the inputs of assibilating sound shifts and their *observed* outputs on the surface can vary widely, but that each assibilating sound shift yields one single *direct* output at first (which can then be reanalysed or changed by other sound shifts). This helps clarifying the disorientating if not misleading abundance of different outputs observed for one and the same sound shift. For example, a historical input /tj/ or /tʲ/ may surface as /t͡ɕ/ in some language (e.g. Polish), but also as /t͡s/ in another (e.g. Québec French), or as /t͡ʃ/ in yet another (e.g. U.S.A. English *tune*); nevertheless, I claim (together with e.g. Recasens 2011) that *every* palatalisation of /tj/ goes through the intermediary step [c] (yielding its sibilant counterpart [t͡ɕ] in Polish, having been depalatalised and fronted in Québec French, and depalatalised and backed in U.S.A. English). The *direct* output (here e.g. [c]) is claimed to be the only natural phonetic output of the sound shift. This claim contradicts several analyses that see (what I consider to be) 'unnatural' outputs as being able to be the direct result of a sound shift (e.g. Yoshida 2001; Telfer 2006; Youngberg 2013, among others, positing [t͡s] to be a possible direct output of stop palatalisation). Positing for example the direct change [kj]/[tj] > [t͡s], I argue, is as little meaningful and as little phonetically explainable as positing the direct change [ki] > [θj] (e.g. from Latin to Castilian Spanish).

My claim of a unitary assibilation (in this example, for [c]) can be represented schematically as illustrated in fig. L, showing that many different inputs and outputs are possible, but all must transit via the intermediary step [c]. As what sibilant phoneme [c] will be reanalysed perceptually is then dependent on the structure of the sibilant inventory. Assibilating stop palatalisation is only one example among several: for example, I

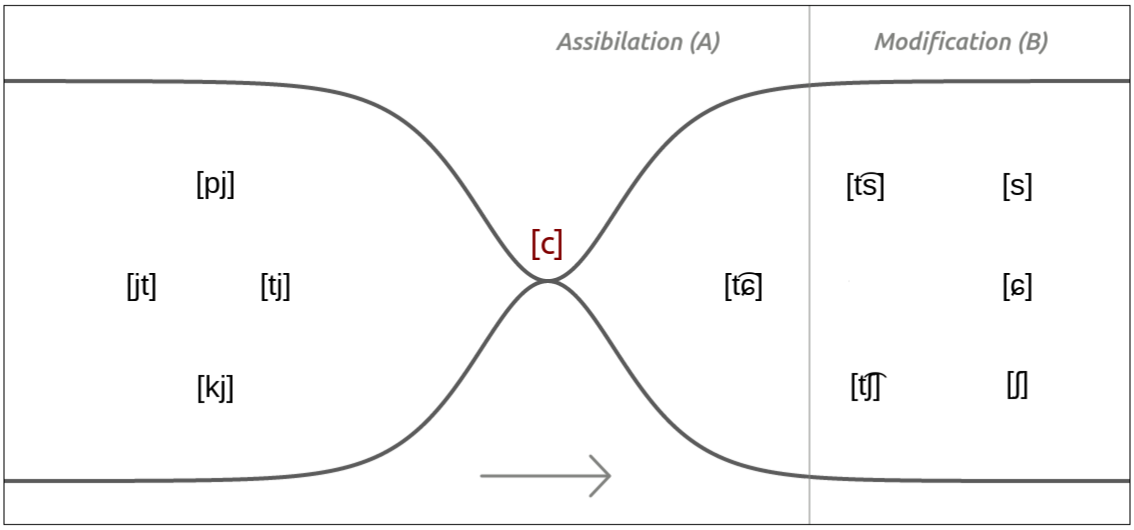


Figure L: *Schematic example of the claim of unitary assibilation for stop palatalisation.*

also claim that s-retraction in /rs/-clusters always yields an apical retracted alveolar [s̺] when triggered by a retracted alveolar [r] (cfr. Kokkelmans 2020*a*).

A. A unitary view on assibilation

This subsubsection details the attested sound shifts that create sibilants. Each of the sound shifts is first introduced by a nonce word example illustrating of what precisely it consists. The following sound shifts are types of assibilations:

Assibilation is a sound shift that transforms a non-sibilant phonetic input into a sibilant phonetic output. Some studies use the term 'assibilation' as a more specific term for e.g. stop palatalisation (Hall et al. 2006); however, in this dissertation, assibilation refers to any sound shift that creates sibilants from exclusively non-sibilant segments.

A large-scale typological account of assibilation in languages of Russia can be found in Sieber (2020)'s dissertation, and an overview of sound shifts that yield retroflex sibilant affricates is given in Nikolaev and Grossman (2018: 599).

In what follows, I will show that assibilation can be understood in this way:

(57)

1) A phonetic input, because of a primary (e.g. [ç]) or secondary (e.g. [tʰ]) aspect of its pronunciation, is articulated with turbulent noise.
2) This turbulence is reanalysed as stridency (Ohala and Solé 2010).
3) The direct phonetic output of this reanalysis (the sibilant output) is determined by the precise articulation of the original phonetic input.

This means for example that if a language has an aspirated (denti-)alveolar [tʰ] and it undergoes assibilation (i.e. if the glottal friction is reanalysed as stridency), it yields a (denti-)alveolar sibilant affricate [t͡s] as *direct* output. If it has a palatal stop [c], it yields [t͡ɕ] at first; if it has an apical postalveolar [tɻ] or [tɹ̠], it will yield [t͡ʂ]; if it has a voiced palatal [ɟ], it will yield a voiced [ʑ], and so on. This is summarised in fig. LI.[12]

A.1 Regressive stop palatalisation (stop + j/front vowel) to [c] > [t͡ɕ]

Example: /kjatjakiti/ > [t͡ɕat͡ɕat͡ɕit͡ɕi]

The general concept of palatalisation is characterised by many different observed inputs, outputs, contexts and linguistic interpretations, as acknowledged by Krämer and Urek (2016*a*: 1):

[12] I do not deem it necessary to include in the figure the option of dental stops/rhotics as contrasting with alveolar stops/rhotics, since they do not contrast in natural languages (Hall 1997*b*: 127). Apical postalveolar rhotics are not considered to be able to provoke sufficient friction to be reanalysed as sibilants.

		Dental	Alveolar	Apical retr. alveolar	Alveopalatal	Apical palatoalveolar	Subapical retroflex
Affricating	Aspiration	[t̪ʰ]	[tʰ]	[t̺ʰ]	[cʰ]	[tʰ]	[t̠ʰ]
	Palatalisation				[c]		
	Rhotic clusters			[tɹ]		[tɻ]	[tɻ̠]
	Lateral affricates			[t͡ɬ]			
	Direct output	[t͡s̪]	[t͡s]	[t͡s̺]	[t͡ɕ]	[t͡ṣ]	[t͡s̠]
Fricativising	Rhotics		[r̝]	[r/r̝]			
	Laterals		[ɬ̞]	[ɬ]	[ʎ̝/ɬ̠]	[l̞]	[l̠˔]
	Other fricatives	[θ]	[θ̠]		[ç/ʝ]		
	Direct output	[s̪]	[s]	[s̺]	[ɕ]	[ṣ]	[s̠]

Figure LI: *Overview of the possible assibilation paths, from stops and non-sibilant fricatives to sibilants.*

> [I]t is by no means clear which processes count as palatalization phonologically or if the term refers to one or several phenomena. Palatalization processes exhibit great variation within and across languages, regarding the triggers of the process, the targets, and the output of the process. Most studies focus on a certain palatalization process in a specific language.

Large-scale studies on regressive stop palatalisation are rather numerous (for example, non-exhaustively: Bhat 1978; Hall et al. 2006; Hall and Hamann 2006; Telfer 2006; Bateman 2007, 2011; Kochetov 2011; Krämer and Urek 2016*b*; see Kokkelmans in prep. and references therein). Different terms are used in the literature (e.g. 'coronalisation', 'affrication', 'assibilation'...) for the phenomenon of stop palatalisation; I will exclusively use the term 'stop palatalisation'. A distinction is made between *full palatalisation*, defined by Bateman (2007, 2011: 589f.) as a sound shift that alters entirely the PoA and/or the manner of articulation of a segment, as opposed to *secondary palatalisation* which pertains to secondary articulation (e.g. [t^j, n^j]). I only address the former phenomenon here.

I define stop palatalisation as a phonetic articulatory change from a combination of segments to the palatal stop [c]. This palatal stop is disfavoured articulatorily (cfr. Lee 2000 cited in Telfer 2006: 57), is typologically marked (Jacobs and Berns 2013) and prone to considerable spontaneously emerging friction (Jacobs and Berns 2013; Hall et al. 2006; Recasens 2013: 17), which both lead very frequently to phonological perceptual reanalysis as a sibilant. Importantly, the definition of stop palatalisation I give consists of a phonetic-articulatory change followed by a phonological-perceptual reanalysis (similarly to Jacobs and Berns 2013): stop palatalisation always yields the phonetic output [c], but the following reanalysis as a sibilant category depends exclusively on the sibilant inventory. The facts that the first motivation for stop palatalisation is articulatory and that it yields the direct output [c] are both confirmed by evidence from phonetic measurements in Recasens and Espinosa (2009) and Recasens (2011). They refer to the stop realisations generated through (velar) stop palatalisation "as (alveolo)palatal, and transcribed as [c] if voiceless and as [ɟ] if voiced, in line with the fact that they show most typically two articulatory patterns: alveolopalatal involving a simultaneous alveolar and palatal closure; and palatal proper with closure occurring exclusively at the palatal zone" (Recasens and Espinosa 2009: 190). This PoA coincides with the sibilant affricate [t͡ɕ].

I agree with Bateman (2011: 591f.) in seeing labial palatalisation as consisting of hardening followed by labial deletion: perseverative coarticulation leads e.g. [bj] to become [bʝ] with a fricative glide, then to be reanalysed as [bd͡ʑ] before the deletion of the labial component (e.g. Old French *rouge* 'red' from Latin RUBEUM). Labial palatalisation is thus different from stop palatalisation, in the sense that it rather works as the process of assibilation of palatal fricatives (detailed in A.3).

The articulatory motivation for stop palatalisation is understood as gestural blending between the trigger and the target, resulting acoustically in temporal overlap (Bateman 2007: 201ff., Bateman 2008: 27, Recasens 2018: 130ff.): applied to stops, the tongue shape of a high or front vowel/glide is anticipated in the articulation of a preceding alveolar or velar stop, so that the stop itself becomes a palatal stop [c]. If the stop is [t], the front of the tongue is backed and the body of the tongue is raised; if it is [k], the front of the tongue is fronted and the body of the tongue remains raised, as shown in fig. LII.

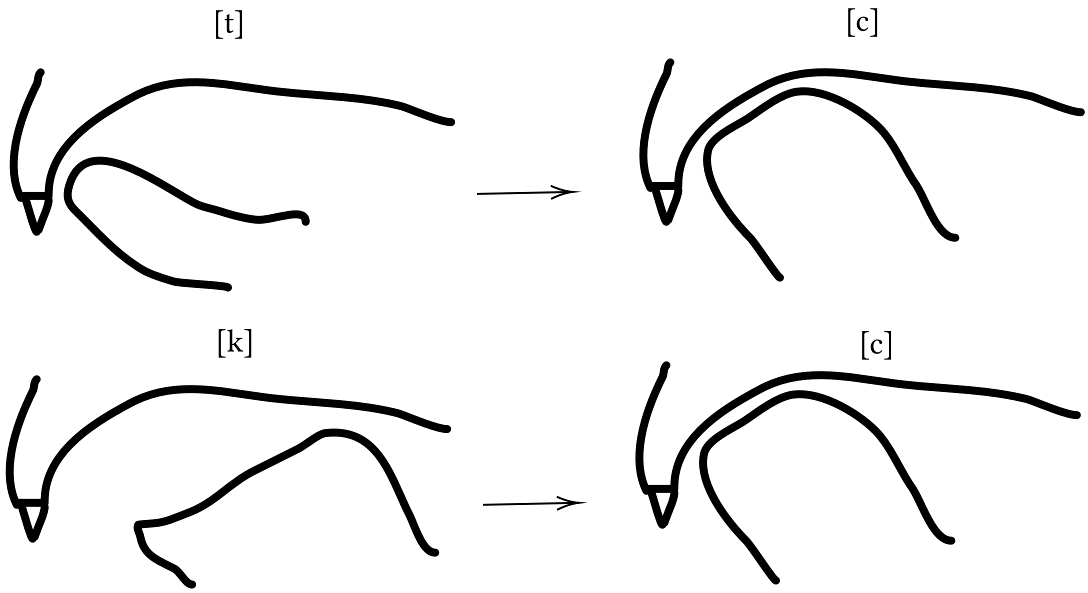


Figure LII: *(Own) schematic representation of coronal and velar stop palatalisation.*

This explains why "coronals are more readily palatalized by high vowels and velars are more likely to be affected by front vowels" (Krämer and Urek 2016*b*: 7): being [+high] suffices for a trigger to make the target [t] raise the tongue body to become [c], while being [-back] suffices for a trigger to make the target [k] front the tongue to become [c].

A common misconception about stop palatalisation is that, since it looks phonologically like a process that converts "a (coronal) stop to [JK: any] sibilant affricate or fricative before high vocoids, e.g. /t/ is realized as [ts], [s] or [tʃ] before /i/" (Hall et al. 2006: 59), it should also be understood phonetically as such. Cases like *t > s /_i* are in fact cases of *rule telescoping* (Anderson 1981: 521): in the absence of evidence for the intermediary stage, some phonologists deduce that the historical input can yield the contemporarily observed output directly. Against that, I argue that the sole phonetic motivation of stop palatalisation is the gestural overlap described above, and thus that its sole direct phonetic output is [c], in accordance with phonetic evidence (Recasens and Espinosa 2009) and some other analyses of the phenomenon (Jacobs and Berns 2013; Ramírez Martínez 2017: 24f.). This contradicts e.g. Telfer (2006) and several others who assume that phonetic stop palatalisation can circumvent the intermediary stage [c]. To illustrate that [c] is the direct output of stop palatalisation, I provide some arguments here in favour of the interpretation that any stop palatalisation can be explained as having yielded a [c] which, entering the sibilant inventory through assibilation, is reanalysed as the sibilant affricate phoneme acoustically nearest to [t͡ɕ].

In a well-dispersed 2s, [c] is either reanalysed as [t͡s] or as [t͡ʃ], most often as [t͡ʃ]. This is because the alveopalatal [t͡ɕ] is situated in between [t͡s] and [t͡ʃ] acoustically, and can thus be reanalysed as the former or the latter; however, since it is acoustically closer to [t͡ʃ], it tends towards [t͡ʃ] more often. For example, Cypriot Greek with its [s] - [ʃ] contrast has undergone a stop palatalisation through which the palatal [c] has been reanalysed as [t͡ʃ] (Manolessou and Pantelidis 2012: 275). In Northern American French, stop palatalisation has yielded both [t͡s] (mainly in Québec French, but also in Cajun, see Emmitte 2013) and [t͡ʃ] (mainly Acadian/Cajun, see Poirier 1994 and Morin 2002: 65). The palatal stop [c] is considered by Poirier (1994: 85) and Juneau (1972: 119-138) to be the original phonetic output of the sound shift, and Morin (2002: 66) even mentions that [c] can still be heard in Québec. In contemporary English, post-lexical palatalisation (e.g. *what you* > [wat͡ɕu]) has become a rule transforming [tj, dj] into the pre-existing phonemic categories /t͡ʃ, d͡ʒ/, as much as it gave /t͡ʃ, d͡ʒ/ within words (e.g. *tune*, *soldier*; cfr. Zsiga 1995).

In some Basque dialects (e.g. Arbizu, Hualde 2004: 46), palatalisation affects /t, n, l/ after /i, j/ (as detailed in A.2 below), so that e.g. standard Basque *aita* 'father' becomes [ajca]. In this case, the direct output of stop palatalisation is preserved as [c]; in other dialects, it is reanalysed as the laminal /t͡ʃ̻/ sibilant phoneme (yielding *atxa* for the same word, de Rijk 2008: 14), which is its closest acoustic category (especially having laminality in common with [c/t͡ɕ]).

An inventory undergoing stop palatalisation can also obtain a new phoneme category from it. In this case, this new category is precisely /t͡ɕ/ (e.g. Japanese: Eckman and Iverson 2012; Chinese: Li 1999: 173; Polish and Russian: Padgett and Żygis 2007; Sanskrit: Hall 1997*a*; Swedish: Asu et al. 2015).

A final argument for why stop palatalisation yields [c/t͡ɕ] comes from the observation of successive palatalisations in Arpitan (French) dialects, compared to Romance languages in general. The Late Latin palatalisation of /tj/ created a palatal affricate that got depalatalised and fronted as soon as the palatalisation of /kj/ and /k + front vowel/ occurred. If such an inventory became a 2s, this yielded a sibinv like that of standard Italian, with the first affricate at the front PoA and the second affricate reassigned to the back PoA.

(58) Standard Italian (Vietti 2019)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ		
Sib. fri.		s z			ʃ		

Standard French however, as a Gallo-Romance variety, experienced one more stop palatalisation (namely that of /ka, ga/). The affricate from the second Latin palatalisation ([t͡s̠]) was fronted and merged with the front affricate from the first palatalisation (e.g. *constan*[t͡s̠]*e* 'constancy', [t͡s̠]*ésar* 'Caesar'). Contemporary standard French thus has [s] (< Old French [t͡s̠]) where standard Italian has [t͡ʃ]. Alpine dialects like that of Évolène in Switzerland, however, underwent an additional palatalisation while still being a 1s, therefore fronting once more the existing affricates (e.g. *lu tchyó* 'your', phonetically [t͡s̠∼t͡ɕ] and corresponding to French *le tien* with [tj]; compare *zornîgva* 'day', corresponding to French *journée* with [ʒ], cfr. Elmiger et al. 2013).

If the direct phonetic output of stop palatalisation were [t͡s̠], we could not explain why the output of the three first palatalisations in the dialect of Évolène has become [t͡s̠] but not that of the fourth ([t͡s̠]): it would either be predicted that they all merge to [t͡s̠], or that the results of the three first palatalisations should have been backed to [t͡s̠]. If the direct phonetic output of stop palatalisation were [t͡s], English could not have post-lexical palatalisation to /t͡ʃ/, since it would obligatorily have become /t͡s/ symmetrically with the existing /s/. If the direct phonetic output were [t͡ʃ], one could not explain why the Northern American French palatalisation did not yield exclusively [t͡ʃ] but also [t͡s].

It is thus possible to predict to a certain extent how stop palatalisation appearing in a sibilant inventory will behave: if the sibinv is a 0s, an assibilated palatal stop is predicted to deaffricate and depalatalise to [s̠] (as in Rotokas), due to the markedness of alveopalatals and affricates in general; in a 1s, it is predicted to remain an alveopalatal affricate at first and optionally to depalatalise to a retracted alveolar, if /sj/ is not also palatalised (as in Proto-Greek); if however a 1s also palatalises /sj/ to [ɕ], it becomes a 1s with alveopalatals (like Old Japanese) and then (with the help of auditory dispersion) a less-dispersed 2s (like modern Japanese). In a well-dispersed 2s that does not palatalise /sj/ nor other consonants, [c] has strong chances to be reanalysed as /t͡ʃ/ (as in English), and smaller chances to be reanalysed as /t͡s/ (as in Québec French). The same holds for a less-dispersed 2s with palatoalveolar sibilants (e.g. palatalised stops became /t͡ʃ/ in Faroese). In a less-dispersed 2s with dental sibilants, [c] becomes obligatorily the more posterior affricate. In a well-dispersed 2s that palatalises /sj/ and other coronal consonants, the palatals become a third sibilant category, yielding a retroflex 3s (like Polish). In a well-dispersed 2s that does not palatalise /sj/ but does palatalise other consonants, it also becomes a 3s (like Serbo-Croatian).

Stop palatalisation has also been explained in perception as the acoustic enhancement of a palatal stop [c]. Flemming (1995: 10) explains the following:

> [C]oronalization of velars by front vowels (e.g. ki ⟶ tʃi) is a further enhancement of vowel or glide F2 distinctions. Palatalizing a velar stop results

> in a palatal stop, which is liable to affrication, probably due to the length of the contact between tongue and palate. This affrication provides a further cue to the frontness of the following vowel or glide. Changing the palatal to a palato-alveolar affricate enhances this difference in affrication because the sibilant palato-alveolar frication is louder than non-sibilant palatal frication, but a palato-alveolar is otherwise acoustically similar to a palatal (…).

The palatalisation of stops in high/front contexts is thus an articulatory-based phenomenon, but the change of a palatal stop [c] to a sibilant affricate is the result of a perceptual preference for sibilant posterior stops rather than non-sibilant posterior stops (Clements 2009: 50).

A.2 Progressive stop palatalisation (j/front vowel + stop) to [c] > [t͡ɕ]

Example: /ajkajtikit/ > [at͡ɕat͡ɕit͡ɕit͡ɕ]

Progressive stop palatalisation is a phenomenon similar to regressive palatalisation explained in the previous section, except for the order of the segments (e.g. /jt/ instead of /tj/). It remains unclear whether it can occur directly as such ([jt] > [c]) or if it requires an intermediary step in which the trigger is transferred from before to after the target ([jt] > [jt^j] > [jc] > [c]); both have been posited (e.g. Recasens 2014 vs. Baker 2004). Regardless of this question, the articulatory explanation as gestural blending holds in any case.

Some varieties of Basque have progressive palatal assimilation, as anticipated in A.1 (e.g. *baina* 'but', older spelling: *baña*; de Rijk 2008: 13). Among these, a few also apply this process to [jt]-clusters, yielding what would be represented orthographically as <tx> (i.e. a laminal [t͡ʃ̻]): "Palatalized *t* may then become indistinguishable from the affricate *tx*, so that *aita* 'father' sounds as [*atxa*]" (de Rijk 2008: 14). The reason for which the primary output [t͡ɕ] is reanalysed as [t͡ʃ̻] rather than [t͡s̺] is that Basque sibilants contrast in distributedness, and thus laminal [c∼t͡ɕ] is reanalysed as laminal [t͡ʃ̻] because both laminal [t͡ʃ̻] and apical [t͡s̺] are acoustically situated one step away from [t͡ɕ], but the laminal one has the additional advantage of having laminality in common.

Additionally to this, most varieties of Basque have the semantically triggered process of expressive palatalisation, which expresses 'smallness' by palatalising the coronal consonants and /k, g/ in a word (indicated in writing with a doubling of the letter, e.g. *ttotta* [coca] 'brandy', lit. 'small drop', from *gota*). Interestingly, it surfaces as laminal [t͡ʃ] in expressive words that have been lexicalised and have replaced the original non-expressive counterpart (e.g. *txiki* 'small' having replaced *tiki*), but it surfaces as palatal [c, ɟ] in those alternations with /t, d/ within the paradigm that are still synchronically productive, e.g. in *tontor* 'peak' vs. *ttonttor* 'hump' (de Rijk 2008: 14f.).

The same phenomenon is observed in the varieties of Spanish that transformed Latin /kt/ into [t͡ʃ] across the intermediary step [jt] (e.g. *noche* 'night' from Latin NOCTEM), if it is indeed true that they went through that step. Evidence in favour of this intermediary step is the presence of the corresponding sound [jt] in other Iberian varieties such as Portuguese (*noite* 'night') (Szigetvári 2008: 122). Baker (2004: 110), for example, posits the steps:

(59) *nok.te* > **nox.te* > **noj.te* > **no.tje* > *no͡tʃe*

As a remark, one should also add to this path the necessary intermediary step **noç.te*, accounting for the fronting from **nox.te* to **noj.te*.

Recasens (2014: 162), on the other hand, explains this sound shift as follows:

> Blending in sequences of dental or alveolar and velar consonants involve the superposition of the tongue front and back contact areas for C1 and C2. The superposition between [k] and [t] in the sequence [kt] may give rise to the (alveo)palatal stop [c] as in N.W. Tuscan [ˈteco] TECTU and Romansh [fac] FACTU (Rohlfs 1966: 366-367; Haiman and Benincà 1992: 70). At a later stage, the outcoming (alveo)palatal stop may be categorized as an affricate, as exemplified by the lexical variants [notʃ] NOCTE in Northern Italian dialects, [fatʃ/ts] FACTU in Romansh and [fretʃ] FRIGIDU in Occitan areas (Ronjat 1932: 181; Rohlfs 1966: 367; Haiman and Benincà 1992: 70).

The same phenomenon is found e.g. in Croatian dialects (Galović 2013), in which older speakers keep the cluster /jd/, while the others have an alveopalatal affricate /d͡ʑ/ (Vulić 2008: 332). Here also, the intermediary step [dj] is posited (Galović 2013: 133). Velars in Proto-Slavic have furthermore experienced progressive palatalisation after front high vowels (Vermeer 2012).

A.3 Assibilation of [ç] to [ɕ]

Example: /açaʝa/ > [aɕaʑa]

This sound shift is precisely like the secondary step of stop palatalisation, in the sense that it consists of the perceptual reanalysis of a palatal non-sibilant as a sibilant. Instead of the palatal stop/affricate [c͡(ç)] being reanalysed as an alveopalatal sibilant affricate [t͡ɕ], the palatal fricative [ç] is reanalysed as an alveopalatal sibilant fricative [ɕ]. The articulatory motivation for this change to a sibilant is to be found in the markedness of palatal segments (Chomsky and Halle 1968; Saib 1974; Schmid 2011; Hall 2013).[13]

This sound shift occurs e.g. in Berlin, Middle German dialects and Luxemburgish (Jannedy and Weirich 2016; Hall 2013) as well as in the speech of young people in several Norwegian cities (Skjevik Jacobsen 2015; Kristoffersen 2000: 23). Manolessou and Pantelidis (2012: 277) also report that "[f]ronting of [x] > [ʃ] is attested for very few lexical items in the Greek dialect of Puglia".

When palatalisation of /hj/-clusters yields [ç] phonetically (as in English, cfr. Ohala and Solé 2010: 306), this [ç] can also become [ɕ]; this is what is likely to have happened in the English word *she* and what certainly happened with /hj/-clusters in Late Old English (Smith 1925; Samuels 1975), transforming e.g. Old Norse *Hjaltland* into *Shetland*. The direct phonetic output [ɕ] was then reanalysed as the phoneme category /ʃ/ in English.

[13] Hall considers [ɕ] as less marked than [ʃ], because he analyses the latter as bearing the feature [labial] in Middle German; nevertheless, [ɕ] is more marked than a palatoalveolar sibilant that does not bear any other feature than [-son, +stri, cor, +cont, -ant], because it bears [+high] in addition to these features.

A.4 Assibilation of [j̥] to [(d͡)ʑ]

Example: /aʝa/ > [ad͡ʑa]

This sound shift looks identical to the previous one, except of course that it concerns the voiced counterparts of the respective segments *and*, crucially, that it can give rise to a sibilant affricate. This sound shift occurs without yielding an affricate e.g. in several Greek dialects (Manolessou and Pantelidis 2012: 273) and in the Taiwanese language Atayal (Huang 2019), with the direct output [ʑ] being observed contemporaneously in both. Northern Mazovian and Kurp, two Polish dialects, exemplify the same change: the former has [pç] and [bʝ] for historical and/or underlying /pj/ and /bj/ respectively, while the latter has [pɕ] and [bʑ] in precisely the same contexts (Ségéral and Scheer 2008: 151). Here again, the direct output is [ʑ].

In the Greek dialect of Astypalaia however, [ʝ] surfaces as a sibilant affricate (Manolessou and Pantelidis 2012: 280). This is also what happened in Late Vulgar Latin (Posner 1996; Bateman 2007: 117): [j] in a syllable onset, and also after labials in Gallo-Romance (Ségéral and Scheer 2008: 153), became fricativised, assibilated and reanalysed as the already existing phoneme /d͡ʒ/, yielding e.g. Italian *gennaio* 'January' from Latin IANUARIUS.

In fact, the change [j] > [ʝ] > [ɟ] can also occur without assibilation, as witnessed by e.g. "Sanskrit *dváyōs* ∼ Gothic *twaddjē*, Old Norse *tveggja* 'of two'" (Szigetvári 2008: 123); the interpretation [ɟ] for Gothic *twaddjē* and Old Norse *tveggja* is especially likely to be accurate considering that it is reflected by both /d/ and /g/ respectively, showing the liminality between [t∼k] that is characteristic of the palatal stop [c] (also, the reflex of Old Norse /ggj/ is now [t͡ʃ] in Faroese, cfr. Árnason 2011). This pattern in Germanic languages thus allows us to understand what assibilation of [ʝ] to [(d͡)ʑ] is: when the direct output is [ʑ], it is simply the voiced variant of the sound shift described in A.3, i.e. [ʝ] > [ʑ]; when the direct output is [d͡ʑ], it is in fact the fortition process [ʝ] > [ɟ] > [d͡ʑ], and thus no different with respect to assibilation from stop palatalisation, described in A.1. Despite their similar inputs and outputs, these two processes are manifestations of these two different sound shifts, rather than a process of its own. Late Vulgar Latin [j] in onsets and Astypalaia Greek must thus have experienced the change [j] > [ʝ] > [ɟ] > [d͡ʑ] (Brandão de Carvalho 2008: 209).

The same way fortition of [j] to [d͡ʑ] can happen, it is accordingly predicted that its voiceless counterpart [j̊], the voiceless palatal approximant, can be assibilated to [c] > [t͡ɕ]. It has indeed happened in historical /hj/-clusters in Faroese, where [hj̊] has yielded /t͡ʃʰ/ (Vrieland 2014; Árnason 2011: 116, 122). This is not surprising in the sense that it parallels the historical development of voiced /j/ (recall *tveggja* 'of two' pronounced [tfɛt͡ʃa]).

A.5 Perceptual reanalysis of aspirated stops as sibilant affricates

Example: /atʰa/ > [at͡sa]

This sound shift is reminiscent of stop palatalisation, but doesn't share its articulatory motivation, having only the perceptual reanalysis part in common. For example, the affricated [t͡s] in Danish reflecting a historical /t/ (Grønnum 2005: 148) and the affricated [t͡s] in Québec French reflecting a historical /t/ might look similar at first, but are the results of different sound shifts. A closer look reveals indeed that the Danish affricate occurs in all contexts in which /t/ is aspirated, regardless of the following vowel, while the Québec French affricate only occurs before /i, y, j, ɥ/, i.e. high front vowels and glides. The Québec French pattern is thus obviously the result of stop palatalisation, while the Danish one is the result of the perceptual reanalysis of aspiration as stridency; to distinguish both, it suffices to see if the affrication takes place e.g. before low back vowels (compare Danish *tom* [t͡sʌmˀ] 'empty' vs. Québec French *tonne* [tɔn] 'ton(ne)'), because stop palatalisation never takes place before low back vowels.

The affrication of aspirated coronal stops has in common with stop palatalisation that it consists of the perceptual reanalysis of a coronal segment articulated with a turbulent airflow as a sibilant affricate. Similarly to how the turbulence in the articulation of [c] provokes the reanalysis as a sibilant affricate [t͡ɕ], the turbulent airflow due to the aspiration in [tʰ] provokes its reanalysis as an assibilated aspirated [tˢʰ]. As predicted by the hypothesis of unitary assibilation (see fig. LI), the precise articulation of the non-sibilant input determines the precise articulation of the direct sibilant output, before it can be reanalysed according to the sibilant inventory. This means that a dental aspirated [t̪ʰ] yields dental [t͡s̪] as a direct phonetic output after assibilation, while an apical retracted alveolar [t̠̺ʰ] yields apical retracted alveolar [t͡s̠̺].

In Danish, the stop is an alveolar one (Basbøll 2005: 61), but since Danish is a 1s, the sibilant part of the affricate is retracted to retracted alveolar [t͡s̠] after assibilation. Besides Danish, languages such as Old High German (Höfler 1957; Lange 2007; Kokkelmans 2019*a*) as well as contemporary English, although at an allophonic level (Buizza and Plug 2012), experience this sound shift. In English, assibilation yields precisely an alveolar [t͡s], as is exemplified on the spectrograms in fig. LIII, since both [t] and [s] are alveolar (Dart 1998). In all cases, it is observed that while [tʰ] is shifted to [tˢʰ], its voiced counterpart [d] becomes devoiced to [t] to fill the gap (and potentially passively voiced in sonorant environments). This happened in Old High German to such an extent that historical /d/ even replaced it to become aspirated [tʰ] in modern Northern standard German (e.g. Dutch [t]*and* vs. German [t͡s]*ahn* 'tooth' showing [tʰ] > [tˢʰ], and Dutch [d]*enneboom* vs. German [tʰ]*annenbaum* 'fir tree', showing [d] > [tʰ]). In Danish and English, historical /d/ shows up as voiceless unaspirated [t], passively voiced in sonorant contexts (Basbøll 2005: 61 and Honeybone 2005: 331).

The Austronesian language Woleaian is perceived in the literature as providing a seemingly unexplainable case of stop palatalisation: historical **t* surfaces as /s/ before any vowel except /a/ (Hall and Hamann 2006: 1198, Telfer 2006: 77). In fact, not only **t* has become a fricative (Tawerilmang and Sohn 1984: 184), but also **p* > [f] and **k* > [x]; it is thus a clear example of perceptual reanalysis of [tʰ] as [tˢʰ], and not of stop palatalisation. Ironically enough, the gap created by [tʰ] > [t͡s] > [s] has not only been filled by loanwords with [t] but also with the historical **s* itself, which shifted to /t/ (Tawerilmang and Sohn 1984: 187). One can thus reconstruct: 1) the spirantisation of the stops /p, t, k/ to /p͡f, t͡s, k͡x/, followed by 2) this gap being filled (e.g. /s/ > /t/), and 3) the deaffrication

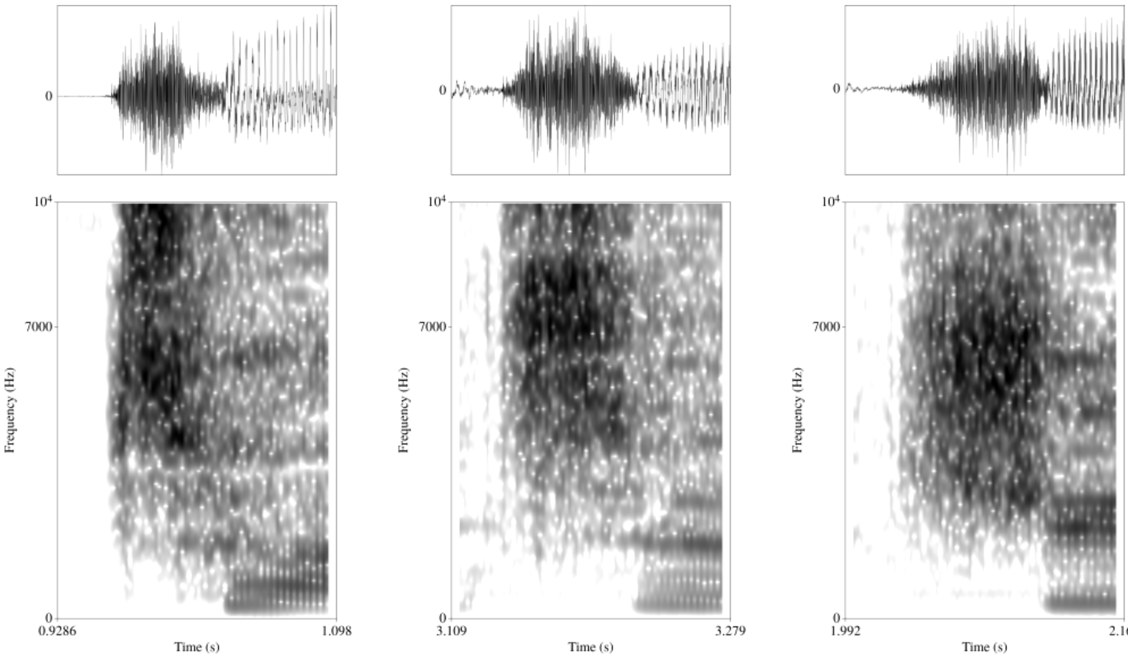


Figure LIII: *Spectrogram extracts in Praat (Boersma and Weenink 2019) showing English /tʰ/-assibilation in the words* 'tired'*,* 'to' *and* 'what you' *(the latter being post-lexical stop palatalisation across word boundaries, for comparison), pronounced by Chester Bennington in the voice-only version of the song* Numb*. The sibilant parts of [tˢʰ] in* 'tired' *and* 'to' *have CoGs of 7679 Hz and 7508 Hz respectively, comparable to the 7699 Hz of the [s] in* 'faithless'*, vs. 5538 Hz for [t͡ɕ] in* 'what you'*, comparable to the 4931 Hz of [ʃ] in the word* 'pressure'*. This is in accordance with the hypothesis that the articulation of the sibilant part in [tˢʰ] gets aligned with that of the nearest PoA for sibilants (here: /s/).*

of the affricates to /f, s, x/.

Assibilation of aspirated coronal stops also happens in Western Andalusian Spanish varieties, in which debuccalised /s/ in /st/-clusters yields [ʰt] > [tʰ] > [tˢʰ] (Del Saz 2019). It is specifically reported that this variety has a dental [s̪], and that the resulting affricate is a dental [t͡s̪] (Del Saz 2019: 760).

A generalisation one can make is that this sound shift seems to be favoured by the presence of [θ] and/or [ð] in the inventory; otherwise, it is more likely to yield [(t͡)θ] rather than a sibilant output. For example, the perceptual reanalysis of aspiration as stridency has given /pʰ, tʰ, kʰ/ > /f, θ, x/ from PIE to Proto-Germanic, probably because [θ] did not exist beforehand in PIE. In Ancient Greek, /pʰ, tʰ, kʰ/ > /f, θ, x/ has given exactly the same result in modern Greek. On the other hand, Pre-Old High German and English both had/have [θ, ð], which leads to think that the assibilation of [tʰ] is motivated by the avoidance of a merger with a pre-existing phoneme /θ/. The picture is less clear in Danish, since it cannot be said not to have interdentals (because it has [ð̞]), but not [θ] (Basbøll 2005: 63). In Woleaian, no *θ is reconstructed either, which also relativises this generalisation. In Andalusia, there is much variation between inventories with and without [θ] (Ruch 2018), which makes it difficult to verify this generalisation; nevertheless, Del Saz (2019) reports that one of her participants uses [θ] and has affrication also.

The Japanese sound change /tɯ/ > [t͡sɯ] is also explainable as the perceptual reanalysis of [tʰ] as [tˢʰ], and not as stop palatalisation (it would then be unexplainable why /tɯ/ does not surface as [t͡ɕ] like /ti/). In fact, the vowel /ɯ/ is devoiced in many contexts, among others next to voiceless segments; /tu/ is then realised as aspirated ([+spread glottis], cfr. Tsuchida 2001); aspiration following the stop is then reanalysed as stridency, yielding [tʰɰ̥] > [t͡sɰ̥] (Ohala and Solé 2010: 306ff.). Since non-high vowels are much less devoiced (Labrune 2012: 34f.) and /t/ is only partially aspirated in Japanese (Riney et al. 2007), combinations like /ta/ or /to/ yield [ta] and [to] rather than [tˢʰa] and [tˢʰo]. Considering the larger occurrence of vowel devoicing and thus of turbulence in /tɯ/ and /dɯ/, one can explain this process as perceptual reanalysis of turbulence as stridency.

A.6 Assibilation of [θ, θ̠] to [s̪, s]

Example: /aθ̠a/ > [asa]

Interdental or alveolar fricatives [θ, θ̠] (cfr. Pétursson 1971) can become assibilated in natural languages (Saib 1974: 18f.). The hypothesis of unitary assibilation predicts that an interdental [θ] should be assibilated to the direct phonetic output [s̪] and a retracted [θ̠] to [s], since their PoAs correspond respectively.

The majority of cases in which [θ, θ̠] disappear from a language are cases in which they become stops. For example, in many varieties of Arabic such as e.g. Moroccan Arabic, the historical interdentals have been shifted to [t, d] (Guerrero 2019: 495); however, in some varieties such as Urban Hijazi Arabic, they surface both as stops and as denti-alveolar [s, z] (Omar and Frith 1975: XIVf.). The assibilation of interdentals seems to also have occurred in Anti-Atlas Tashlhiyt Berber (Ridouane 2013), in which singleton dental stops are realised as sibilant fricatives whenever they are not in a cluster with another coronal consonant (note that the language furthermore has voiceless geminate stops, surfacing as voiceless stops). Other Berber dialects have a general rule of spirantisation of singleton stops [t] > [θ] (Saib 1974: 18) in the same environment, leading to think that [s, z] in Anti-Atlas Tashlhiyt Berber is probably the historical result of singleton /t, d/ > [θ, ð] > [s, z].

Languages in which interdental or alveolar fricatives have become sibilant fricatives furthermore include: "Burmese, Arabic, Breton, Hebrew, Akkadian, and Laconian (Ancient Greek)" (Li 1996), as well as dialects from Veneto (Marcato and Ursini 1998).

A.7 Assibilation of lateral fricatives to sibilant fricatives

Example: /aɬa/ > [as̠a]

The Proto-Semitic voiceless lateral **(t͡)ɬ*, which has been reconstructed both as an affricate or as a fricative lateral (e.g. Faber 1990: 632 vs. Versteegh et al. 2007: 189), has yielded /ʃ/ in modern Arabic and /s/ in modern Hebrew (Ronén 2019). It had a voiced counterpart in Arabic, represented by the letter *ḍād* (Versteegh et al. 2005: 544f.). While the latter merged to an interdental voiced fricative, **(t͡)ɬ* yields /ʃ/ in all Arabic dialects (except South Arabian, cfr. Ronén 2019).

Since the original PoA and precise articulation of **(t͡)ɬ* is unknown and lateral fricatives can have as many different PoAs as sibilants, e.g. dental or retroflex (Ronén 2019), one cannot do much to verify the hypothesis of unitary assibilation in this case. Nevertheless, a retracted alveolar [ɬ̠] is the most probable reconstruction in Proto-Semitic for two reasons: 1) like for sibilants, the unmarked PoA for laterals is likely to be retracted alveolar (e.g. the presence of a posterior lateral fricative implies that of an anterior one, cfr. Ronén 2019), and Proto-Semitic is likely to have had one PoA for lateral fricatives; 2) this [ɬ̠] has been reanalysed as an alveolar sibilant fricative in Hebrew but a postalveolar sibilant fricative in Arabic, and this is more likely to happen if the original sound is situated right in between; a postalveolar lateral fricative would probably have yielded postalveolar sibilant fricatives in both languages, and an alveolar one would probably have yielded an alveolar sibilant in both languages.

A.8 Assibilation of the lateral alveopalatal [ʎ] to [ʑ]

Example: /aʎa/ > [aʑa]

Besides the lateral fricatives, the lateral alveopalatal [ʎ] can also be assibilated to a palatal sibilant. In the history of Spanish, this has happened e.g. with Late Latin *fi*[lj]*um* 'son' that has become *hi*[x]*o* in modern Spanish, with modern [x] reflecting [ʒ] in Late Old Spanish (Allen 2002: 92). Considering that [ʎ] was originally voiced and palatal, the direct phonetic output must have first been [ʑ] and then [ʒ] after depalatalisation, before being devoiced to [ʃ] when all sibilants were devoiced and backed to [x] when the backing of postalveolars occurred.

In certain varieties of Spanish (e.g. Argentinian), [ʎ] from the historical geminate *ll* changes to [ʒ] (once more), a sound shift for which the intermediary steps [ɟ͡ʝ] and [ʝ] are furthermore attested in various dialects (Martínez-Celdrán et al. 2003: 258).

A.9 Assibilation in /tr/-clusters to apical sibilant affricates

Example: /atɻa/ > [at͡ʂa]

Clusters of a coronal rhotic preceded by a coronal stop display affrication in several languages of the world. The phonetic motivation is that a coronal /r/ becomes acoustically similar to a sibilant affricate due to the friction that follows the release of the preceding coronal stop. It is then predicted to yield an apical direct phonetic output, since /r/ is apical. The hypothesis of unitary assibilation predicts that the PoA and the distributedness of the rhotic should determine that of the direct phonetic output of the sibilant, since /r/ in a /tr, dr/-cluster determines that of the preceding stop (Baković 2006). For example, a retroflex /r/ will lead to /tr/ being pronounced [t͡ʂ] as a direct phonetic output (which can then be reanalysed by the sibilant inventory).

Assibilation in /tr/-clusters happens in languages that have an approximant /r/, like English: in certain varieties of American English, /tr/ is assibilated to /t͡ʃ/, as visible in spelling errors like *chrain* for *train* (Lance and Howie 1997: 356, cited in Janda and Joseph

2003: 216; Stevens and Harrington 2016: 119). This also happens in voiced /dr/-clusters (e.g. *jug* for 'drug', cfr. Baković 2006). The direct output [t͡ʃ̺] is reanalysed as /t͡ʃ/ (as shown by the misspellings), but still realised as an apical sibilant affricate (e.g. apical [t͡ʃ̺]*ue* 'true' vs. unspecified/tendentially laminal [t͡ʃ]*ew* 'chew' are not homophones).

It happens in languages that have an alveolar trill/tap/flap, like Chilean Spanish (Canfield 1981: 8): /tr/ is assibilated to /t͡ʃ/, and in Chilean varieties that deaffricate it to [ʃ], /tr/ predictably surfaces as [ʃɹ] (Subiabre Ubilla 2015: 56). However, as the form [ʃɹ] shows, the /r/ is preserved as a coronal *approximant*, rather than as the apical trill/tap/flap [r/ɾ] that is otherwise used in Chilean Spanish. The assibilation might thus be triggered by the approximant allophone of /r/ in /tr/, as is posited by Whitley (2003: 83). Here again, assibilation is thus triggered by an approximant /r/.

In Old Chinese, /trj/-clusters became affricated to retroflex sibilant affricates (Baxter 1992). The change also affected /srj/-clusters (see B.5). An assibilated version of retroflex [ʈ] is furthermore reported in Sicilian, with the direct phonetic output being [t͡ʂ] (Celata 2006: 2).

A.10 Assibilation of coronal rhotics to apical sibilants

Example: /ar̠a/ > [az̺a]

The motivation for this sound shift is that, whereas alveolar rhotic trills/tap-flaps are more marked cross-linguistically than sibilants (43.68% frequency in the UPSID (Maddieson 1984) vs. 92.02%, cfr. Kokkelmans 2020*a*: 221 and Pultrová 2013: 22), voiced sibilants are more marked than their voiceless counterparts (while the opposite holds for rhotic trills and tap-flaps). If a language does not have voicing in sibilant fricatives (i.e. does not have [z] or [z̠] etc.), markedness does thus not favour the shift from e.g. [r] to [z̠], because voiced sibilants are more marked. If a language already has voiced sibilant fricatives, the perceptual reanalysis of a coronal rhotic as a sibilant is however possible, since the markedness of voiced sibilants is already 'violated'.

Coronal rhotics have been reanalysed as voiced apical sibilants e.g. in Andes Spanish, among others Peruvian (Alvord et al. 2005), and in Polish (Szamryk 2014: 51). In the former case, spectrograms show that the phonetic output is acoustically close to a retracted alveolar [z̠] (because of large amounts of energy around 2000∼4000 Hz, see Alvord et al. 2005: 35 and Canfield 1981: 7), something which is predicted by the fact that [r] is a retracted alveolar rhotic trill. In the latter case, it is a diachronic development that has led from historical /r^j/ to modern standard Polish apical [ʒ̺] (written <rz>), attested also in Jita and Oceanic languages (Hall and Żygis 2010: 4). In both cases, the motivation for the sound shift seems to be a perceptual reanalysis of friction as stridency.

Not only voiced coronal rhotics, but also voiceless ones can be reanalysed as sibilants. In Eastern Tyrolean dialects, a historical /r/ in /rt/ clusters was first devoiced to [r̥] under the influence of the following voiceless /t/, and then the friction arising from this voiceless rhotic was reanalysed perceptually as stridency, yielding [s̺]. Like other [s̺t]-clusters, preconsonantal s-retraction then transformed it to /ʃt/. Examples include the word *Kårtn* 'map, card', which is pronounced [kɑʃtn̩] (Alber 2020: 13; Seeber 2017:

150, cited in Rabanus 2019).

B. Changes within sibilant inventories

This subsubsection lists the changes that transform sibilants into other sibilants.

B.1 Regressive sibilant palatalisation (sibilant + j or front vowel) to [(t͡)ɕ]

Example: /asjaʃjasiʃit͡sit͡ʃi/ > [aɕaɕaɕiɕit͡ɕit͡ɕi]

This sound shift is the sibilant correspondent of regressive stop palatalisation, described in A.1. It has precisely the same articulatory motivations: as with /t/, gestural blending applies the tongue raising gesture characteristic of palatal /j/ to the preceding sibilant, so that it becomes [(t͡)ɕ] (Bybee 2007: 949). Crucially, irrespectively of the PoA and distributedness of the original sibilant, the direct output is always alveopalatal [(t͡)ɕ] (e.g. [s̪j] > [ɕ] in Korean and Japanese, cfr. Eckman and Iverson 2012; Bybee 2007: 948, [sj] >[ɕ] in Afrikaans, [s̺j] > [ɕ] in Danish, [ṣi] > [ɕi] in Gujarati, cfr. Hamann 2003: 96). This sound shift often targets sibilant fricatives like [s] (e.g. Romanian historical /s/ > /ʃ/

in /si, se/ as in the words *și* 'and' from Latin SIC 'thus' and *șapte* 'seven' from Latin SEPTEM), but can also occur with affricates like [t͡s] (e.g. Middle Chinese [t͡si] > [t͡ɕi], cfr. Baxter 1992: 52).

It is generally observed that if it occurs productively across word boundaries, it predictably yields [(t͡)ɕ]. Even in the case of the English post-lexical palatalisation of /sj/, despite the output being categorised acoustically and phonemically as a /ʃ/ (and being widely referred to as [ʃ] in the literature), it has been shown to rather match precisely the articulation of [ɕ]. Zsiga (1995: 73) says:

> For *s+you*, however, the /ʃ/ template is not a good fit: there is too much contact at the front and center of the palate. This poor fit illustrates that although the acoustics in the *s+you* sequence at this point in time may be /ʃ/-like, the articulation is not. Rather, (...) the pattern in the *s+you* sequence is what would be expected from an /s/ and /j/ being articulated at the same time.

Phonetic evidence of gestural blending thus reveals that post-lexical /sj/ is in fact realised as [ɕ], while lexical /ʃ/ from historical /sj/ (e.g. in the word *pression*) is realised as [ʃ].

Besides English, regressive sibilant palatalisation also occurs across word boundaries e.g. in Danish (Grønnum 2005: 146, Basbøll 2005: 65), in Afrikaans and Hollandic Dutch (Collins and Mees 1999: 215, e.g. *as jy kom* and *als jij komt* 'if you come' respectively) and Norwegian (Kristoffersen 2000; Bøe 2015, e.g. *hvis jeg* 'if I'). The phonetic output of this productive rule is [ɕ] in all cases.

In languages in which it has been a historical process, it has been reanalysed as the acoustically closest sibilant phoneme (e.g. lexicalised palatalisation in English: *mi*[ʃ]*on*, compare modern standard French *mi*[sj]*on*). Besides English, the product of palatalisation in /sj/-clusters is e.g. [ʃ] in Faroese (Árnason 2011: 176) and Tuscan Italian among other Romance varieties (Canalis 2017; Zampaulo 2019: 93).

It appears that languages can palatalise /tj/ without palatalising /sj/ (e.g. Serbo-Croatian), the other way around (e.g. Afrikaans *wa*[tj]*y* 'what you'), and palatalise both /tj/ and /sj/ (e.g. Korean). There is thus seemingly no universal implication here, while there exist several implications with respect to other aspects of palatalisation (see among others Bateman 2007; Hall and Hamann 2006).

B.2 Progressive sibilant palatalisation (j or front vowel + sibilant) to [(t͡)ɕ]

Example: /ajsajʃisiʃit͡sit͡ʃ/ > [aɕaɕiɕiɕit͡ɕit͡ɕ]

Progressive sibilant palatalisation is a rather rare phenomenon (like progressive palatalisation in general). It consists of a sibilant becoming alveopalatal after a /j/ or a front vowel. The *RUKI*-rule in *satem* descendants of Proto-Indo-European corresponds partially to the description of this sound shift. This rule consisted of the retraction of PIE **s̩* after any of the segments /r, u, k, i/ (Pedersen 1895: 74), something which is in fact a height assimilation in the case of /u, k, i/ (but not in the case of /r/, which represents another sound shift explained in B.4), and which however occurs irrespectively of frontness (i.e. also with posterior [u] and [k]). Contrarily to that in /r, u, k/, the shift of PIE **s̠* to postalveolar consonants after both /i/ and /j/ (Beguš 2017) in itself is a perfect example of progressive sibilant palatalisation: the sibilant fricative in /is/ and /js/ has been palatalised to [ɕ]. It is indeed posited to have yielded the direct phonetic output [ɕ], which has however had several thousands of years of time to merge with palatoalveolar /ʃ/, and has been shifted further to a retroflex /ṣ/ later on in the *satem* languages that became a 3s inventory. Considering the age of the *RUKI*-rule, it has however to remain a hypothesis whether [ɕ] was the direct phonetic output of the rule.

Besides the *RUKI*-rule, this sound shift occurs in a variety of Kannada: Telfer (2006: 59) says that "Bhat (1978: 62) notes that the causative suffix -su of the Havyaka dialect of Kannada is pronounced -ʃu when preceded by i, e and j". The nature of the triggers clearly indicate that it is a case of progressive sibilant palatalisation, triggered by [j] and front vowels (here [e] and [i]).

B.3 Height assimilation to [+high] sibilants [(t͡)ɕ] and [(t͡)ʃ̻]

Example: /aksasisus/ > [akʃ̻asiɕuʃ̻]

Height assimilation is a process in which a non-high sibilant becomes a sibilant articulated with a raised tongue gesture ([+high] in DFT terms), triggered by gestural overlap with an adjacent [+high] segment. Contrarily to progressive and regressive sibilant palatalisation (B.2 and B.1), it does not apply only in (high) front vowel/glide contexts such as [i]/[j], but in any [+high] context instead.

As explained in B.2, the *RUKI*-rule occurred in *satem* IE daughter languages such as e.g. Sanskrit or Balto-Slavic, transforming PIE **s̠* into postalveolar sibilants after /r, u, k, i/ (Pedersen 1895). The development with /rs/ is a completely different one in terms of phonetic motivation (apical retracted alveolar assimilation rather than height assimilation), which is very interestingly reflected in the fact that it did not take place at the same time as /u, k, i/ but happened in a different, ulterior stage of its own (Whitehead 2012: 163). It is analysed in B.4. The articulatory motivation for the original **s̠* surfacing as a postalveolar sibilant after /u, k, i/ can be explained as the gestural overlap of a high and *front* tongue position with the articulation of the sibilant in the case of [i], and as that of a high and *back* tongue position with the articulation of the sibilant in the case of [u] and [k]. The direct phonetic output is thus dependent on the anteriority/backness of the trigger of the sound shift: it is predicted to yield [ɕ] after [i], and [ʃ̻] after [u] and [k] (with [ʃ̻] being the postalveolar counterpart of [ɕ], i.e. standing for a laminal palatoalveolar sibilant fricative with a tongue raised against the palate, like the Basque laminal palatoalveolar <x>).

The example given with the /ks/-cluster was a case of progressive height assimilation; I will give here an example of regressive height assimilation in the cluster /sk/.

The historical Germanic cluster /sk/ surfaces in English and German as /ʃ/ regardless of the position in a word or the neighbouring vowel/consonant (Hogg 1996; Hall 2008; e.g. *shine* and *scheinen* 'to shine' respectively, both from PG **skīnaną*). It might look like the straightforward hypothetical change [s̠k] > [ʃk] > [ʃ], but evidence suggests that it occurred as [s̠k] > [s̠kx] > [s̠x] > [ʃ̻x] > [ʃ] (Kokkelmans 2020*b*).[14]

B.4 Assimilation in coronal rhotic + sibilant fricative clusters to apical sibilants

Example: /arsaɹ̠sa/ > [as̺̠as̺a]

Perseverative coarticulatory assimilation in clusters like /rs/ yields direct phonetic outputs that can get reanalysed as a different phoneme category. In fact, even in a 1s, coarticulation to [rs̠] (both being already alveolar retracted, /s̠/ being apical under the influence of the preceding apical rhotic) already favours gestural blending to the point that the trill is likely to disappear, leaving an apical retracted alveolar [s̺̠] behind. Sporadic examples are Late Latin RUSSUM for RURSUM 'backwards' (Lloyd 1987: 139), Latin URSUS 'bear' > Old Spanish *osso* and the unetymological Middle French *jalou*[s̺̠] 'jealous' > Middle Dutch *jaloers*, considered to be a hypercorrection (cfr. Philippa et al. 2005: "zwakke articulatie of zelfs assimilatie aan de -*s*, zodat er in de uitspraak geen verschil was tussen -*oes*, -*ous*, en -*oers*, -*ours*", i.e. assimilation to the extent that the variants with and without /r/ are undistinguished). An example with /rʃ/ is standard German *haarscharf* "by a hair's breadth" vs. Ahrntal Tyrolean *hou*[ʃ]*orf* (Seeber 2017: 150), showing a historical assimilation process in /rʃ/ to [rs̺̠] with r-loss, followed by the merger of /s̺̠/ to /ʃ/ after [r] also attested in most German varieties.

The PoA of the rhotic becomes the PoA of the sibilant fricative, and the apical tongue shape of the rhotic is transmitted to the sibilant fricative. In languages with 2s or more, a change of phoneme category can occur if there is a mismatch between the coarticulated PoA and/or distributedness of the sibilant fricative after /r/ and that of its original phonemic category. This is observed e.g. in Faroese, which contrasts retracted alveolar sibilants and palatoalveolar ones. Speakers who preserve the original retracted alveolar trill [r] in their pronunciation also preserve the retracted alveolar sibilant [s̠] as such, but speakers who use a postalveolar plain retroflex approximant [ɹ̣] retract a following /s̠/ to postalveolar plain retroflex [ṣ] (Árnason 2011: 116).

In other languages, an apical retracted alveolar [s̺̠] has been shown to follow a trill [r] while originally belonging to a different phoneme category: in e.g. Afrikaans (Ewald 2015; Wissing et al. 2015), Flemish Dutch (Kokkelmans 2020*a*), Western Basque (Trask 1996: 77) and some varieties of Italian (Baroni 2014*a*: 143f.). The same process occurred

[14]A first observation is that in Dutch, the change took place only partially so that it halted at the intermediary stage [s̠x]. Furthermore, the intermediary steps [s̠kx] and [s̠x] are reflected in orthography: besides the historical orthography <scinen> (for [s̠k]), the verb 'to shine' appears from Old to Middle English as <schinen> (reflecting [s̠kx]) and <shinen> (reflecting [s̠x]). Since <c> was used in Latin for [k], <s> for [s̠] and <h> for [x∼h], the spellings <sch> (identical in German) and <sh> are most likely to reflect the intermediary stages of the sound shift. In several Scandinavian varieties, /sk/-clusters have been transformed into /ʃ/ before /j/ and front vowels, and Bøe (2015) notes that the intermediary stages [s̠ç] and [ɕç] (which he writes <sç> and <ʃç>, respectively) are found in recordings of older Norwegian speakers in Western Norway. One can thus understand this change from /sk/ to /ʃ/ as an affrication of the initial cluster to [s̠kx] for English/German and [s̠kç] in Scandinavia (due to 1) aspiration in English and German /kʰ/ to [kx] and 2) affrication in the front vowel/glide context of Scandinavian /skj/ to [kç], respectively), followed by the deletion of /k/ and then the working of this sound shift, i.e. height assimilation to [ʃ̻x] triggered by the high [x] for English and German and height assimilation to [ɕç] triggered by the high [ç] in Scandinavian.

historically in Middle High German (Hall 2008), Limburgish (Kokkelmans 2020*b*) and some French dialects near Germany and Luxemburg (Remacle 1992). The coarticulation of [r] + sibilant is an underresearched subject, which is among others probably a consequence of the rather limited awareness among linguists that retracted alveolar sibilants like [s̠] exist and can be referred to with that name and that IPA character (see 3.1.5.1). Due to the influence of a retroflex flap, Norwegian and Swedish dialects have experienced assimilation and s-retraction in /rs/ to a plain retroflex sibilant (Stausland Johnsen 2012; Hamann 2003: 83), similarly to Faroese.

In Kokkelmans (2017, 2020*a*), I showed that s-retraction in /rs/ happened in Middle High German similarly to how it happens today in Flemish regional varieties of standard Dutch, yielding the same direct phonetic output [s̠̺] and only after an alveolar trill [r]. The first analysis (Kokkelmans 2017) had the wrong hypothesis that the sound shift /rs/ -> /rʃ/ could be due to the *Obligatory Contour Principle* (van Oostendorp 2005*a*) or a similar kind of dissimilation, but the second (Kokkelmans 2020*a*) rectified it in seeing it as an apical tongue shape and retracted alveolar place of articulation assimilation to [rs̠̺]. Neither analysis attempted to determine if other sibilant fricatives are subject to the same assimilation pattern. It seems to me however that the Hungarian postalveolar /ʃ/ is realised (at least by some speakers) as an apical retracted alveolar [s̠̺] after the alveolar trill [r]; this is only based on my impressionistic evaluation, but could be quantified in further research.

The reader is referred to Kokkelmans (2020*a*) for an overview and demonstration that [r] triggers the coarticulation of a following sibilant fricative as [s̠̺], as well as further details on this topic.

Despite yielding one and the same phonetic output [s̠̺], the phonemicisation of this coarticulatory bias has lead to different surface outputs in various languages, because of their different sibilant inventories at the time of phonemicisation:

- In Western Basque (Trask 2004: 77), a [s] (<z>) after [r] has been reanalysed as the phoneme category /s̠/ (<s>), which is itself *also* an apical retracted alveolar sibilant fricative in its realisation. This is especially good evidence for the fact that [s̠̺] is the only direct phonetic output of assimilation in [rs], because it has not become the more postalveolar /ʃ/.
- In Limburgish, standard German and the French dialects mentioned above (Kokkelmans 2020*b*), a /s/ after [r] has been reanalysed as the postalveolar sibilant category /ʃ/. This is a consequence of the fact that all these varieties lost the phonetic realisation [s̠] for the phoneme /s/ when their sibilant inventories became a well-dispersed 2s with [s] vs. [ʃ].
- In many Scandinavian varieties (Solhaug 2010; Stausland Johnsen 2012), a /s/ after [r] has merged with the apical plain retroflex [ṣ] that arose from clusters containing the retroflex rhotic flap [ɽ]. In Swedish, the retraction of [s̠̺] to [ṣ] is motivated by the phonemicisation of /ɕ/, which pushed the apical sibilant backwards to maximise their perceptual distance, and by the appearance of [ɽ] in some dialects.

B.5 Assimilation in sibilant fricative + coronal rhotic clusters to apical sibilants

Example: /sra/ > [s̺a]

A similar phenomenon as the one in /rs/ can happen in /sr/-clusters, as witnessed by some varieties of English (e.g. *di*[ʃ̺]*respect*, Janda and Joseph 2003: 212), by Pashto (Hall 1997*b*: 212) and Old Chinese (Baxter 1992: 271). Both languages have a retroflex rhotic (Hall 1997*b*: 215), which leads to think that this sound shift can have been triggered by the retroflex nature of the rhotic. It is thus probably a retroflex assimilation, occurring in the opposite direction as that in retroflex /rs/-clusters.

B.6 Depalatalisation of alveopalatal sibilants

Example: /aɕat͡ɕa/ > /aʃat͡ʃa/

As seen in the previous sections about palatalisation, a palatoalveolar [ʃ] often surfaces where a diachronic change is predicted to have yielded [ɕ] instead. This is due to the process of depalatalisation, which transforms sibilants with a raised tongue body ([+high] in DFT terms) into sibilants unspecified for height (i.e. in most contexts, they will not be high, except adjacent to [+high] segments). Depalatalisation is active in several languages as a synchronic rule (e.g. Slovak, Rubach 1993 and Spanish, Bessett and Colina 2017) when the palatal allophone of a segment alternates with its non-palatal counterpart (e.g. Castilian Spanish *do*[ɲ]*a* 'Lady' vs. *do*[n] 'Sir'), but is also an irreversible diachronic process (recall e.g. the observation in B.1 that historical sibilant palatalisation in English /sj/ is reflected in articulation by a [ʃ] for /ʃ/ indistinguishable from [ʃ] for /ʃ/ from historical /sk/, but that the same productive post-lexical process yields [ɕ]).

B.7 Post-coronal stop intrusion generating sibilant affricates

Example: /ansarsalʃa/ > [ant͡sart͡salt͡ʃa]

Stops emerge spontaneously in consonant clusters consisting of a coronal sonorant (/n/, /l/, /r/) followed by a sibilant (Recasens 2012; Marotta 2005: 12). In clusters such as /ns/, /ls/ and /rs/, the articulatory transition from one consonant is characterised by a brief occlusion period that results acoustically in a short /t/-like sound (a phenomenon known as *stop intrusion*, cfr. Bateman 2007: 199). This has been phonologised e.g. in Tuscan and Roman Italian, in which /s/ after coronal sonorants is realised as an affricate [t͡s] (Marotta 2008: 237; Calamai 2016: 221; Marotta 2005: 11), corresponding to the phoneme /t͡s/ (resulting e.g. in the near-neutralisation of *orso* 'bear' and *orzo* 'barley', pronounced *or*[t͡s]*o* and *or*[d͡z]*o*, respectively). It also occurs in certain varieties of Spanish in the same contexts (Widdison 1997: 256). In early French, it happened together with a series of other intrusive stops (e.g. /lr/ > [ldr]), yielding e.g. *filts* 'son' from FIL(IU)S (Berns 2013: 128f.). The affricate, now lost due to deaffrication in Middle French, is still reflected in loanwords such as English *Fitzgerald* (from (Norman) French *filts* + *Gérald*).

The direct phonetic output depends exclusively on the sibilant fricative: a sequence [ns] will give [nt͡s], a sequence [nʃ] will give [nt͡ʃ], a sequence [nʒ] will give [nd͡ʒ] etc. The direct phonetic output is not reassigned to another phoneme in terms of PoA, because according to the principle of sibilant fricative and affricate symmetry (see 4.7 and 5.4.2), if a language has a sibilant fricative at a PoA, it can have an affricate at the same PoA or a gap, and vice-versa. A sequence /ns/ will thus not yield [nt͡ʃ], since the existence of /s/ implies that /t͡s/ is a possible *and* perceptually closer phonemic category.

B.8 Stop + sibilant fricative coalescence to sibilant affricates

Some languages contrast sequences of a stop and a sibilant fricative (e.g. Polish *trzysta* [tʂ̣ɨ.sta] 'three hundred') with sibilant affricates (e.g. Polish *czysta* [t͡ʂ̣ɨ.sta] 'vodka (colloquial)', illustrated in fig. XXII). However, such a distinction is not necessarily made in all languages, where a stop + sibilant fricative sequence can also be reanalysed as a sibilant affricate. Berns (2013: 128) exemplifies this case with early Gallo-Romance:

> The deletion of unstressed vowels in Gallo-Romance created affricates if it took place between a coronal plosive and a sibilant fricative. As noted by Pope (1934: 147), the two consonants that became juxtaposed after the erasure of the vowel merged into one single consonant: "[...] the resultant group t + s was replaced by the affricated denti-palatal ts [...]".

Examples given are e.g. *nets* 'naked' from NATUS and *grants* 'tall' from GRANDIS. Another example is German *Deutsch* 'German', where the affricate appeared after the disappearance of the /i/ that originally separated the *t-* and *sch* parts. The [t͡ʃ] is acoustically undistinguishable from a [t͡ʃ] in English loanwords into German (e.g. *Match*).

As for post-coronal stop intrusion, the direct phonetic output will adopt the PoA of the sibilant fricative, so that e.g. [ts] yields [t͡s], [dʒ] yields [d͡ʒ], and so further.

B.9 Deaffrication to sibilant fricatives

Example: /t͡sat͡ʃa/ > [saʃa]

In contexts that are not 'after coronal sonorants' as in B.7, there exists an opposite tendency to deaffricate sibilant affricates, because they are more complex and effortful segments than fricatives, and their acquisition presupposes that of stops and fricatives (Fletcher 1989: 746). The direct phonetic output of deaffrication corresponds precisely to the sibilant part of the original affricate (e.g. [d͡ʒ] > [ʒ]).

For example, /t͡s/ and /t͡ʃ/ as well as their voiced counterparts got deaffricated starting in the 13th century in Middle French (Tifrit and Voeltzel 2016: 2), yielding a 3s inventory with /s/, /s̺/ and /ʃ̺/ (plus their voiced counterparts). The dental affricates and /d͡ʒ/ in Old Spanish were also deaffricated in Middle Spanish to their fricative counterparts (Allen 2002: 97ff.), leaving /t͡ʃ/ as the only affricate.

There seems thus to be no absolute predictable pattern in deaffrication, if not that voiced affricates are deaffricated if their voiceless counterparts are deaffricated and disappear from the inventory (this to avoid an impossible voicing gap). If not all contexts are concerned, it still is possible that only voiceless affricates are deaffricated: this is exemplified by Roman Italian (Marotta 2005: 12), in which only /t͡ʃ/ is deaffricated between vowels (e.g. [t͡ʃ]*ena* 'dinner' but *la* [ʃ]*ena* 'the dinner'; and [d͡ʒ]*ente* irrespectively of the context).[15] In Tuscan Italian, both /t͡ʃ/ and /d͡ʒ/ deaffricate in the same context.

Romanian illustrates the case in which only /d͡z/ (and not /t͡s, t͡ʃ, d͡ʒ/) got deaffricated (Żygis, Fuchs and Koenig 2012; Telfer 2006: 84), and Walloon illustrates the case in which only the anterior affricates /t͡s, d͡z/ (and not /t͡ʃ, d͡ʒ/) got deaffricated.

Statistically, gaps at voiced and anterior affricates are more frequent (Żygis, Fuchs and Koenig 2012), but this does not predict with certainty which affricate will deaffricate if another is deaffricated. In terms of statistical likelihood, anterior voiced affricates are most likely to deaffricate (as witnessed by the many examples given in Żygis, Fuchs and Koenig 2012: 307ff.), followed by posterior voiced affricates (*idem*) and anterior voiceless affricates (exemplified by Walloon). No other pattern with deaffrication of /t͡ʃ/ without that of /d͡ʒ/ than the Roman Italian one are known to me. The statistical trend in terms of deaffrication in a 2s thus seems to be:

(60) d͡z > d͡ʒ, t͡s > t͡ʃ

In a 1s, only the voiced affricate deaffrication tendency is relevant, and in a 3s, too little data is available to make any statistical claim.

B.10 Voicing of sibilants in intervocalic/sonorant contexts

Example: /ansaʃaʃ/ > [anzaʒaʃ]

This sound shift is motivated by the fact that in voiced contexts (e.g. intervocalically or adjacent to voiced segments), the articulatorily least effortful option is to allow the vocal folds to vibrate also during the production of the sibilant. In those particular contexts, the universal markedness of voiced sibilants is reversed in favour of voiced sibilants (and of voiced obstruents in general).

Languages in which sibilants are allophonically passively voiced between vowels or adjacent to voiced segments (e.g. /s̺b/ realised as [z̺β] in Spanish and Basque, cfr. García 2013 and Hualde 2004: 100) can phonemicise this passive voicing, something which happened e.g. in Gallo-Romance varieties (Martinet 1952): the intervocalic allophonically voiced /s/ became /z/ while the intervocalic geminate /sː/ became singleton /s/ (e.g. modern French *ca*[z]*e* 'box, square' vs. *ca*[s]*e* 'breaks', written *case* and *casse*, respectively).

Other examples include Old Latin (e.g. *genu*[s̠] (nom.), *gene*[z̠]*is* (gen.) 'birth, origin', later yielding rhotacism of [z̠]), Proto-Germanic (*idem*), Northern and Middle German, English, Dutch and Limburgish (Kokkelmans 2020*b*).

There is no implication with respect to the distinction between the intervocalic context and the context adjacent to voiced consonants: in southern Italian varieties as well as in Spanish and Basque, /s/ is voiced in /zv, zn, zm, zl, zb/ etc. but not intervocalically (compare e.g. *ca*[s]*a* 'house'), while historical German, Limburgish and Dutch illustrate

[15] Deaffrication only targets singletons, so that the geminates /t͡sː, d͡zː, t͡ʃː, d͡ʒː/ remain unaffected. There is thus still no voicing gap at /t͡ʃ/ in Roman Italian.

the opposite pattern (e.g. Dutch *hui*[z]*en* 'houses' but [s]*langen* 'snakes'). Middle French illustrates the case in which both happened (Ohala and Solé 2010: 314f.). An interesting pattern is furthermore provided by 17th-century Dutch, in which /s/ in /sw/ but in no other consonantal context (such as /sl, sm, sn/ etc.) was voiced to /zw/ (Donaldson 1983: 151). This adds to the observation that the context in which sibilant fricatives will become voiced is seemingly not predictable.

Besides anterior sibilant fricatives, intervocalic posterior /ʃ/ is also voiced to /ʒ/ in Eys Limburgish (phonemically in *ange*[ʒ]*öm* 'the other way around', allophonically in e.g. *wie*[ʃ∼ʒ]*e* 'if you were'). Voicing of intervocalic affricates is also attested, namely in Gallo-Romance (Allen 2002: 95).

In originally voiceless sibilant inventories that gain voiced sibilant affricates (e.g. trough palatalisation), voicing of /s/ is observed to occur at least allophonically. For example, Old English palatalised /ki, gi/ to /t͡ʃ, d͡ʒ/ and then voiced /s/ intervocalically (e.g. *hou*[z]*es*). Nevertheless, a language can also phonemicise sibilant fricative voicing without having phonemicised voiced sibilants earlier (e.g. Dutch).

B.11 Devoicing of sibilants

Example: /and͡zaʒ/ > /ant͡saʃ/

This sound shift is motivated by the general markedness of voiced sibilants (see 4.5).

Several languages illustrate the full merger of voiced sibilants with their voiced counterparts: Spanish lost all its sibilant voicing distinctions around the 17th century (Moore 2005; Widdison 1997); some Dutch dialects, especially in Northern Holland, are devoicing their voiced sibilants (De Wulf and Goossens 2005; Taeldeman 2006); Afrikaans also lost its voiced sibilant, and by means of rhotacism, Proto-Germanic as well as Old Latin lost their sole voiced sibilant.

When this happens, all sibilants are devoiced at once: to my knowledge, no examples are found in which e.g. [d͡z] merges with [t͡s] but [d͡ʒ] remains distinct from [t͡ʃ]. This indicates that the switch occurring in the speakers' grammars is a radical one between 'allow voiced sibilants' and 'do not allow voiced sibilants at all', rather than one like 'allow voiced sibilant X' vs. 'do not allow voiced sibilant Y'.

B.12 Long-distance assimilation: sibilant harmony

Example: /and͡zaʒat͡ʃ/ > [and͡zazat͡s]

Several languages exhibit a pattern like the following ones (Hansson 2001*a*: 189):

(61)

(1) /d͡ʒaʃ-(u)s-esːe/ > [d͡ʒaʃuʃeʃːe] (Koyra)

(2) /s-api-t͡ʃʰo-it/ > [ʃ-api-t͡ʃʰo-it] (Ineseño)

These examples are cases of sibilant harmony: the PoA of all sibilants in a word match that of the first sibilant (i.e. progressive, as in Koyra) or the PoA of all sibilants in a word match that of the last sibilant (regressive, as in Ineseño). Besides the progressive/regressive distinction, some sibilant harmony patterns are sensitive to the PoA of the sibilants rather than their linear order (i.e. there are several languages in which only postalveolar sibilants trigger harmony, and one in which only alveolar sibilants trigger harmony). Sibilant harmony is part of the larger-scope phenomenon of consonant harmony (Rose and Walker 2004). The interested reader is referred to e.g. Walker et al. (2006); Hansson (2010); Berkson (2013); Bennett (2015) for more literature on sibilant harmony. This process is typically observed in 2s, but also in 3s (e.g. Basque and Benchnon, cfr. Hansson 2010: 52f.).

An important observation is that sibilant harmony does not necessarily target all the sibilants in a language, and does not necessarily occur regularly in a language. For example, sibilant harmony is found in most of Moroccan dialects, except Eastern dialects (Guerrero 2019: 495), but only with the *voiced* postalveolar sibilant fricative as a trigger (regressively, e.g. *zuʒ* > *ʒuʒ* 'two, pair' but *sbaʕtaʃ*, not **ʃbaʕtaʃ* 'seventeen'; cfr. Zellou 2013; though compare also *ʃəmʃ* 'sun' for standard Arabic *ʃams*). It happened in (at least) one word in Middle French, but not regularly at all for the rest (Hock 2009: 75): *chercher* 'to search', historically [s̺]*erchier* (with the original pronunciation without sibilant harmony still preserved in the English loanword). The fact that it is only triggered by the voiced postalveolar sibilant in Moroccan Arabic reveals something especially interesting: sibilant harmony tends to assimilate the less marked segments to the more marked (Hansson 2001*b*: 316), as a kind of *emergence of the marked* or preference for the marked (McCarthy and Prince 1994; Albright 2004).

The recurring pattern in sibilant harmony is that either all sibilants can be triggers, or only postalveolar ones can be triggers (e.g. /*saʃa*/ > [ʃaʃa] but /*ʃasa*/ > [ʃasa]), with only one known exception in natural languages (Hansson 2001*a*: 192ff.; Kosa 2010). This speaks in favour of the "interpretation that phonology encodes this correlation as an initial bias for palatalizing sibilant harmony (or for preserving palatal sibilants over alveolar sibilants)" (Kosa 2010: 14): in section 4.4, it will be argued that the explanation for the postalveolar bias is to be found in the markedness of postalveolars, because marked segments are more likely to spread their features.

The motivation for sibilant harmony is understood teleologically[16] as a means of avoiding difficult tongue position transitions, or non-teleologically as the emergence of harmony patterns from frequent tongue slips. In the first case, a constraint or rule favours using one and the same tongue gesture for sibilants inside a word, while in the latter case, tongue slips (which are more often postalveolar-favouring) lead to their reinterpretation as the main pattern in a language (Hansson 2001*a*). In perception, it was shown that a following sibilant makes the identification of a sibilant as /s/ or /ʃ/ more difficult (Fleischer et al. 2013), something which might reinforce the bias towards sibilant harmony: the neutralisation of sequences of sibilants is not only favoured articulatorily, but is also more easily tolerated because it is less noticeable perceptually.

[16]In this dissertation, I intend to use the term 'teleological' similarly to its use in Boersma and Hamann (2008). 'Teleological' refers to a principle encoded in the synchronic grammar of the speaker (in OT, as a constraint), as opposed to an 'emergent' principle, which arises grammar-externally.

B.13 Direct sibilant assimilation

Example: /aʒsast͡ʃ/ > [aʃːaʃt͡ʃ]

Besides the long-distance assimilation phenomenon of sibilant harmony, sibilants with different voicing and/or PoAs are also dispreferred as adjacent segments in the languages of the world. This tendency for sibilant sequences to undergo voicing assimilation and PoA assimilation is formulated as universal #813 in the Konstanz Universals Archive (Plank and Filimonova 2000), following (Greenberg 1965):

> In regard to sibilants there is a strong tendency for different types of sibilants not to combine. If it occurs, then it is with morpheme boundary.

The motivation for this phenomenon is to be found in the reduction of tongue gesture transitions and thus of articulatory complexity (Pouplier et al. 2011: 25). Like for long-distance assimilation, it can be regressive vs. progressive or postalveolar-favouring.

In modern standard French, the assimilation pattern is optional and favours postalveolar consonants asymmetrically. Voicing assimilation is always regressive in standard French (Niebuhr et al. 2008; Niebuhr and Meunier 2011, e.g. *je sais* 'I know' with /ʒs/ reduced to [ʃːe] and *trei*[ʃː]*ambres* 'thirteen rooms' with /zʃ/ > [ʃː] because of postalveolar place assimilation and regressive voicing assimilation). Direct sibilant assimilation behaves identically in Catalan (de Lacy 2002: 426). In Upper Sorbian, sibilant assimilation is regressive irrespectively of the anterior - posterior distinction, yielding e.g. *jeničce* 'only' pronounced as [jeːnʲit͡st͡sɛ] instead of [jeːnʲit͡ʃt͡sɛ] (Jocz 2015: 177). The phonetic research on English /ʃs/- and /sʃ/-sequences led by Pouplier et al. (2011: 24) shows that both /ʃs/ and /sʃ/ assimilate regressively (i.e. towards the PoA of the last sibilant), but assimilation favours more often /ʃ/ (to the point of being almost indistinguishable from /ʃʃ/-sequences) as compared to /s/ (Clayards et al. 2014). English thus belongs to a mixed pattern between regressive and postalveolar-favouring.

Note: From here onwards, perceptually motivated sound shifts are detailed (as opposed to articulatory motivated sound shifts, detailed in the previous paragraphs).

B.14 Auditory dispersion

Example: /ʃanas̺a/ > /ʃanasa/

It has often been observed that contrasting sibilant phonemes tend to be articulated towards diverging extremities of the perceptual continuum (mainly in terms of CoG), this to avoid perceptual confusion (Martinet 1955; Adams 1975: 290; Flemming 1996; Boersma and Hamann 2006; Padgett and Żygis 2007; Boersma and Hamann 2008; Flemming 2018). This is reflected, for instance, in the predominance of the well-dispersed type of 2s with respect to the less-dispersed types of 2s (in the SibInv database, 129 out of 149 sibinvs i.e. 82.17% are well-dispersed).

Examples showing the effects of auditory dispersion in the transition from a 1s to a 2s are Irish, standard Italian and English, in which /ʃ/ was phonemicised (from /s̺ʲ, s̺k and s̺k + front vowel/, respectively), resulting in the fronting of the historical /s̺/ to the (denti-)alveolar [s] it is today. Examples from a well-dispersed 2s to a 3s are provided by Serbo-Croatian, Polish and Sanskrit, in which the phonemicisation of /ɕ/ led to the backing of the historical /ʃ/ to the plain retroflex /ṣ/ (Hall 1997*a*; Seinhorst 2012: 7).

Auditory dispersion does not obligatorily create well-dispersed 2s: Catalan and Faroese are examples of sibilant inventories that maintain a contrast between /s̠/ and /ʃ/. Those inventories allow for a larger perceptual confusion (recall e.g. from 3.1.4 the Yiddish loanwords with [ʃ] variously adapted as /s̠/ and /ɕ/ in Hollandic Dutch, cfr. van der Sijs 2001: 270f.), but can remain stable across the centuries (e.g. approximatively 400 years of alternation between /s̠/ and /ɕ/ in Hollandic Dutch for these loanwords).

B.15 Sibilant merger

Example: /ʃanas̠a/ > /s̠anas̠a/

Mergers can occur from one PoA to another, from sibilant affricates to sibilant fricatives (see B.9) or from voiced sibilants to voiceless sibilants (see B.11). The motivation for sibilant merger is the perceptual confusion between two perceptually close categories, which itself can be caused by an articulatorily motivated reduction of the articulatory differences between two sounds: some speakers might not manage to produce sibilants perceptually different enough not to be confused, which results in perceptual confusion and thus reanalysis of the originally different sibilants as one and the same category.

Older Mauritian Creole exemplifies the merger from a 2s to a 1s due to the influence of the 1s (see 3.3.1.2), while modern French exemplifies the merger from a 3s to a 2s (/s̠/ and /s̪/ merging to /s/).

In a 3s inventory, the middle sibilant category (/s̠/ or /ɕ/) is the one to merge with another PoA. In a retroflex 3s, this middle category tends to merge with the back category (/ɕ/ with /ṣ/ from Sanskrit to modern Hindi and from standard Serbo-Croatian to Croatian, but no example of /ɕ/ with /s̪/ are known to me), while in a non-retroflex 3s, it tends to merge towards the front (/s̠/ > /s̪/ in Late Middle French, and /s̠/ > /s̪/ except /s̠/ > /ʃ̻/ after /r/ and before consonants in Late Middle High German, but no example of full merger /s̠/ > /ʃ̻/ are known to me).

B.16 Reanalysis of a non-native sibilant

Example: /as̡̠a/ > /aɕa/

Boersma and Hamann (2008), Seinhorst (2012) and Boersma et al. (2020) show how a child learning a language develops a categorising behaviour based on the input she gets; after perceiving e.g. the two different PoAs for sibilants of English a large number of times, the child has developed a knowledge of what acoustic input corresponds to which category (in BiPhon, by means of cue constraints). This learning process as 'perceptual tuning' to the sibilant inventory of the child's native language leads to the same categorisation mechanism when the adult listener perceives the sibilants of another language, something which is visible in loanwords: for example, a native speaker of a well-dispersed 2s will perceive and categorise both /ɕ/ and /ṣ/ from a retroflex 3s as /ʃ/ (e.g. French *babou*[ʃ]*ka* 'Russian old woman' from Russian *бабу*[ṣ]*ка* as much as *appara*[t͡ʃ]*ik* 'Communist bureaucrat' from Russian *аппара*[t͡ɕ]*ик*). In a 1s like Latin, as exemplified in 3.2.2.2, the nativisation of loanwords implies that any sibilant be mapped to the only PoA available.

There is thus no need to posit a perception grammar for non-native sibilants distinct from the perception grammar for sibilants in a language: the mapping of sibilant allophones to phonemes in the native language equals the mapping of sibilants in loanwords to native phonemes.

Two grammar-external factors can also play a role in the reanalysis of a non-native sibilant, leading to a phonemic output that would not otherwise be predicted by the perception grammar: historical correspondences between two related languages and orthography.

The first case is exemplified by transfers from Tuscan Italian to standard Italian; since historical singleton /t͡ʃ/ and /d͡ʒ/ are deaffricated intervocalically in Tuscan Italian (Ulfsbjorninn 2017), there is a correspondence between some intervocalic occurrences of [ʃ]/[ʒ] in Tuscan and [t͡ʃ]/[d͡ʒ] in standard Italian (e.g. *cu*[ʃ]*ina* 'kitchen' vs. *cu*[t͡ʃ]*ina*). This has resulted in the reanalysis of Tuscan words with intervocalic [ʃ] as having underlying /t͡ʃ/ in standard Italian, while it might have a different origin. In the case of e.g. the first name *Luigia* 'Louise', the French medieval pronunciation *Loui*[z̠]*e* was reanalysed in medieval Tuscan as *Lui*[ʒ]*a* (see B.18) before being transferred to standard Italian as *Lui*[d͡ʒ]*a* (Adams 1975: 285), as if it had had an underlying /d͡ʒ/ (which it never had), analogically with plenty of words with intervocalic [ʒ] in Tuscan that do come from a historical /d͡ʒ/. The same applies e.g. to the transfer from native Tuscan *ba*[ʃ]*o* 'kiss' to standard Italian *ba*[t͡ʃ]*o*, with a Tuscan [ʃ] that in fact originates in the palatalisation of Tuscan /sj/-clusters (< BASIU, cfr. Zampaulo 2019: 93) rather than a historical affricate.

Orthography can also influence the reanalysis of a non-native sibilant: this is exemplified by the Italian pronunciation of German words with /ʃp/ and /ʃt/ vs. /ʃl/ and /ʃ + other sonorants/. In the former case, the pronunciation [ʃ] is reflected as <s> in the German orthography (corresponding to Italian <s>) and thus pronounced [s] (e.g. *la scuola Steiner* 'the Steiner school' with [st] instead of [ʃt]). In the latter case, [ʃ] is reflected as <sch> (corresponding to German /ʃ/ and therefore to Italian /ʃ/) and thus pronounced [ʃ] (e.g. *lo schlager* 'the schlager' with [ʃ]).

B.17 Preconsonantal s-retraction

Example: /as̠pas̠/ > /aʃpas/

Some 2s languages exhibit a pattern /s/ > /ʃ/ in front of (certain or all) consonants. For example, English *stone* with [st] from historical **s̠t* corresponds to German *Stein* with [ʃt] from the same historical **s̠t* (Wiese 1991, 1996; Benware 1996; Hall and Scott 2007). In standard German, preconsonantal s-retraction only occurred word-initially, but in southern German varieties such as Tyrolean or Swiss German, it also applies to word-internal or word-final positions (e.g. *We*[ʃt] 'West', cfr. Bukmaier et al. 2014). In Italian dialects, it does not depend on the position in a word, but reveals interesting different patterns: /sp, st, sk/, /sp, ʃt, sk/, /ʃp, st, ʃk/ and /ʃp, ʃt, ʃk/ (Lorenzetti 2018).

Preconsonantal s-retraction as a sound shift is analysed among others in Alber et al. (2021), to which the reader is referred for a detailed account of its occurrence and its potential motivations. A possible motivation hypothesised in this paper is the perceptual similarity between a retracted alveolar [s̠] and [ʃ] in front of consonants, leading to a phoneme reanalysis /s̠/ > /ʃ/. This hypothesis would predict that preconsonantal s-retraction can only start to occur with a retracted alveolar [s̠] (although it could then be extended to front sibilants).

B.18 Asymmetrical reanalysis of (apical) retracted alveolar sibilants as postalveolar

Example: /azaz̠aʒa/ > /azaʒaʒa/

Alveolar retracted sibilants, when articulated apically (something which happens much more readily than for (denti-)alveolar sibilants, cfr. Recasens and Espinosa 2007: 146 and below), are more likely to be reanalysed acoustically as belonging to a posterior phoneme category rather than an anterior one. The perceptual motivation for this phenomenon is explained by Dart (1991: 29), citing Catford (1977: 157), as follows:

> The acoustic and aerodynamic differences between apical and laminal /S/ are more evident if they are alveolar than if they are dental. A retracted apical fricative opens up a large sublingual resonance cavity, which is characteristic of [ʃ] production and would presumably cause the /S/ to trespass on the acoustic space of this contrasting segment (...).

The larger sublingual resonance cavity of apical retracted alveolar sibilants makes them acoustically closer to postalveolar sibilants, and thus more likely to be perceived as posterior segments. This is confirmed independently by Zampaulo (2019: 93), citing Lloyd (1987: 266) and Ladefoged and Maddieson (1996: 145-150). Since there is a strong cross-linguistic tendency for dental segments to be laminal and for alveolar, in particular *retracted* alveolar segments, to be apical (Chomsky and Halle 1968: 312f.; Hall 1997*b*: 42), it results in a cross-linguistic tendency to reanalyse retracted alveolar segments as postalveolar segments.

Adams (1975) provides plenty of examples in European languages of the reanalysis of a non-native [s̠] in a well-dispersed 2s, and shows that in most cases, [s̠] is adapted as /ʃ/ rather than /s/ (e.g. *leash* from Middle Norman French *laisse* 'leash'). Even within a language, occurrences of apical /s̺/ can be reanalysed as belonging to /ʃ/, as illustrated e.g. by Old Spanish [ʃ]*abón* from SAPONEM or [ʃ]*ugo* from SUCUM (Zampaulo 2019: 94). Considering the absence of a common phonological environment (before vowels or consonants, be it high, low, front or back segments), this phenomenon cannot be explained otherwise as the manifestation of a perceptual bias favouring [s̺] > /ʃ/.

B.19 Independent affricate fronting/sociolinguistic markers

In 2.4, the fronting of standard Polish /ɕ/ to a sound approximating [sʲ] by young women was mentioned, as reported by Czaplicki et al. (2016). In this case, it seems to be a sociolinguistically motivated shift, because only young women are observed to do that. Such sociolinguistically motivated phenomena are more accurately describable as tendencies in pronunciation rather than sound shifts, since the fact that e.g. young women pronounce sibilants in a certain way does not necessarily mean that the entire population of speakers will also shift their pronunciation.

C. A unitary view on dessibilation

C.1 Rhotacism of voiced apical sibilants to coronal rhotics

Example: /az̠̺a/ > /ara/

The same way perceptual reanalysis can lead to rhotics becoming sibilants (see A.10), the reverse can happen so that a voiced sibilant fricative becomes a rhotic trill or tap-flap. The motivation for this phenomenon is articulatory markedness, which disfavours voiced sibilants and thus favours their perceptual reanalysis as a coronal rhotic that already exists in a given language. In simple words, rhotacism allows a language to 'get rid of voiced sibilants'.

This has happened in the transition from Proto-Germanic to modern Germanic languages, with the voiced retracted alveolar sibilant [z̠] (allophonically apical) becoming /r/ (Denton 2003; e.g. German *älter* 'older' from PG **alþizô*). It occurred in all Germanic languages except Gothic (compare *alþiza*). From Old Latin to Latin, rhotacism also transformed all occurrences of [z̠] into [r] (Roberts 2012; Pultrová 2013).

The hypothesis of unitary *des*sibilation, in this case, predicts that the phonetic nature (PoA, voicing etc.) of the direct output corresponds to that of the sibilant input. A retracted alveolar apical [z̠̺] thus yields a retracted alveolar apical [r], a postalveolar apical [ʒ̺] should yield a postalveolar apical [r̠], and so forth.

C.2 Interdentalisation of dental sibilants to [θ, θ̠]

Example: /as̪a/ > /aθa/

Whereas the reverse sound shift is motivated by the reduction of markedness, this sound shift is motivated by auditory dispersion (cfr. B.14). To increase the perceptual distance between a (denti-)alveolar sibilant and posterior sibilants, it is fronted and dessibilated to [θ, θ̠]. The hypothesis of unitary dessibilation predicts that [s̪] yields [θ] while [s] yields [θ̠] as direct phonetic outputs, since they correspond in PoA. However, if the resulting inventory has no (denti-)alveolar sibilants, the interdental fricative will be retracted to [θ̠] to avoid unnecessary articulatory effort.

This is exemplified by modern Spanish, in which the historical [s̪] became an interdental [θ] before being retracted to [θ̠], since only alveolar retracted sibilants were left in the sibilant inventory (Moore 2005; Allen 2002: 99). Another example is given by some varieties from Northern Veneto in Italy (described by Zamboni 1988; see also the literature mentioned in Alber and Rabanus 2017), in which historical [t͡s̪] and [d͡z̪] were shifted to interdental [θ], [ð] (Belloni 2006: 25). One must observe about these varieties that probably due to the markedness of the affricate [t͡θ], the sibilant affricate has become a deaffricated [θ], while the other affricates have not been deaffricated. Other languages in which sibilants have become interdental or alveolar fricatives [θ, θ̠] include Tahltan (Northern Athabaskan), (Galician) Spanish and Gosiute Shoshoni (Li 1996).

C.3 Dessibilation of sibilant fricatives to lateral fricatives

Example: /asa/ > /aɬa/

Ronén (2019) reports that some Sinitic varieties have experienced a change from historical Middle Chinese sibilants to the lateral fricative [ɬ]. This consitutes the opposite change from the one described in A.7. Considering the rarity of lateral fricatives, this sound shift is extremely rare in natural languages.

C.4 Dorsalisation or debuccalisation of sibilants to [ɧ∼x] or [h]

Example: /aʃa/ > /aha/

Several languages experience(d) a change that transforms sibilants into dorsals ([x], [ɧ]) or the segment [h]. It remains unclear whether dorsalisation and debuccalisation are two distinct sound shifts or two steps in one and the same development (e.g. [s] > [x] > [h] > [∅]). This sound shift can be understood as a lenition process (Ferguson 1990; Honeybone 2008) triggered by articulatory markedness (the space between the tongue and the roof of the mouth increases, with the tongue being less raised from its average position). It can also be motivated as an enhancement of auditory dispersion, when a sibilant is pushed further to the back to contrast better perceptually with more anterior sibilants.

This happened from Old to modern Spanish (Widdison 1997; Moore 2005; Allen 2002: 99), when the postalveolar sibilant [ʃ̺] was retracted to [x]. Some varieties of Spanish represent a further step towards lenition, since they realise this [x] as [h] (Canfield 1981). The enhancement of auditory dispersion is especially probable as a motivation for this change in the sense that the anterior sibilant was also shifted further to the front (to [θ]) in Castilian Spanish, so that only the middle sibilant [s̺] would remain in the sibilant inventory. An example of dispersion-motivated dorsalisation is given by standard Swedish, in which the posterior sibilant originating in /sj/-clusters (known as the *sje-ljud* 'sje-sound') was pushed further back to become [ɧ]. Had it not been so, Swedish would have a 4s, which is typologically very marked and rare. Another example of dorsalisation is given by Forest Nenets (Salminen 2007: 369; **s* > /x/ word-initially). Other examples of debuccalisation include Hawaiian (with /h/ from Proto-Polynesian **s*), some Québec French varieties (Morin 2002: 47, 62), St. Gallen Swiss dialects (Berger 1913: 132), Italian dialects around Brescia and Val Camonica (Cresci 2014), Proto-Finnic, some Eastern French dialects before consonants (Schmid 1956) and Proto-Greek (only word-initially, e.g. *helios* 'sun' from Late PIE **sāwélios*, compare Latin SOL, SOLIS). Ferguson (1990: 65) furthermore mentions Armenian, Yakut, Evenki, Tongan, Tahitian, Lomongo, Kpelle, varieties of Malay and Sanskrit. The interested reader is referred to Ferguson (1990) for a detailed analysis.

A recurring pattern is the fact that the most posterior sibilant is always the one to be debuccalised (e.g. in Late Middle Spanish, Swedish, Québec French and Proto-Finnic). In fact, dorsalisation or debuccalisation is not attested with (denti-)alveolar sibilants: its target is either retracted alveolar [s̠] (as in Proto-Greek or Late Medieval Eastern French dialects; Schmid 1956) or more posterior sounds (e.g. /ṣ > h/ in Iskonawa, while /s/ and /ʃ/ remained unaffected; Zariquiey 2015: 68f.). This can be explained by the fact "that [ʃ] and [x] are perceptually closer than [s] and [ʃ]" (Baroni 2014*a*: 97), as measured in the English perception of German by Kemp (2011: 18). Complementarily, fronting to [θ, θ̠] only targets (denti-)alveolar sibilants. The hypothesis of unitary dessibilation would predict that e.g. [ʃ] is dorsalised to [x̟] at first, [ɕ] as [ç] etc., although this cannot be tested here by lack of evidence.

Besides the hierarchical implication that debuccalisation triggers the most posterior sibilants, there is another implication that voiceless sibilants are more likely to be targeted. In Québec French varieties, both [ʃ] and [ʒ] become [h] and [ɦ] respectively (Morin 2002: 47, 62), but in Val Camonica dialects, only the voiceless [s̠] was debuccalised, which leads to paradigmatic alternations between [z̠] and [h] (Cresci 2014: 36). The gap at /s̠/ has been filled in the meanwhile with numerous loanwords from neighbouring dialects and standard Italian.

C.5 Stopping of sibilants

Example: /ad͡z̠a/ > /ada/

The shift from sibilants to stops is a rare phenomenon. The motivation for this sound shift is the markedness of fricatives with respect to stops.

In some varieties of Venetian Italian, one finds e.g. [d] corresponding to standard Italian [d͡ʒ] (e.g. *el piande* 'he cries' in Brazilian Veneto varieties, cfr. Mengarda 1996: 62), but it remains unclear whether this is an example of the stopping of sibilants or the stoppings of an earlier [ð] (which is found in other Venetian dialects). In Woleaian however, the change **s* > [t] is a certain one (Tawerilmang and Sohn 1984: 187). The reason for the sound shift **s* > [t] in Woleaian is undubitably linked to the sound shift [tʰ] > [t͡s] > [s] that created a gap at [t]; the latter must have preceded the former, otherwise **t* would have participated in the cycle also and be [t] in the modern language.

4. The universal principles that shape sibilant inventories

The overview of all attested (sub)types of sibilant inventories in the previous chapter (3.2.2) showed that the diversity of attested sibinvs is reduced to a limited set of possibilities:

- A number of places of articulation between 0 and 4, at most
- No voiced sibilant without a voiceless counterpart (except Arabic /d͡ʒ/)
- No inventory without an anterior (dental, alveolar or retracted alveolar) /s/
- Etc. (see 3.2.4 for a complete list)

This limited set contrasts with the large number of logically possible sibilant inventories, which leads to think that a set of principles might rule out a large number of logically possible inventories.[1] This chapter addresses and provides evidence for these principles.

A basic distinction can be made between two conceptions about the structure of sibilant inventories. The first is: "Sibinvs are what they are, and their possible structures cannot be (universally) predicted" (this either as a moderate relativism, claiming that the motivation for the structures of sibinvs is language-specific, or much more radically that there is no motivation at all and 'anything goes', as described by Hyman 2008: 87f., 123).

[1] Recall that 'logically possible' refers to the complete set of possibilities given a certain number of distinctions (4 PoAs, 2 voicing values etc.), while 'predicted' refers to a subset of these logically possible inventories after ruling out inventories that do not satisfy the restrictions observed and hypothesised as principles in natural languages. 'Attested' inventories are then a subset of the predicted ones.

The second states on the contrary that their possible structures *can* be predicted, and that the principles behind these structures are universally valid. In this chapter, I will demonstrate that the second hypothesis is right by showing that a number of observations can be made that are generalisable to all attested sibilant inventories, observations I call the *principles* that shape the structure of sibilant inventories.

The first section of this chapter gives several arguments against the 'null hypothesis' formulated as the first hypothesis here above. The subsequent sections describe each principle posited to exist and show, for each of them, that it is a universally valid principle both because of cross-linguistic evidence and because of an articulatory or perceptual motivation acknowledged in the literature.

4.1 Against a 'null hypothesis' of sibinv structure

In this section, I provide evidence against the 'null hypothesis' in the form of 1) statistical evidence, both from the observation of attested patterns and from its comparison with randomly generated inventories; 2) sensorimotor evidence and 3) the argument that, as this dissertation shows, the attested patterns are predicted by a restricted set of principles and their interaction as constraints in OT.

4.1.1 'Null hypothesis' as full or partial unpredictability

This subsection shows first to what extent the attested sibilant inventories are a few compared to the logically possible inventories, and thus to what extent the idea that sibinvs are fully random or unpredictable is unlikely to be true.

While the most radical 'sibinv-are-random-and-unpredictable' view has not (as far as I know) been defended in the rather recent literature on sibilant inventories, older views on phonology in general have held a rather radical relativistic position (Joos 1957: 96, underlining in the original):

> Trubetzkoy phonology tried to explain everything from articulatory acoustics and a minimum set of phonological laws taken as essentially valid for all languages alike, flatly contradicting the American (Boas) tradition that languages could differ from each other without limit and in unpredictable ways, and offering too much of a phonological explanation where a sober taxonomy would serve as well.

From the radical relativistic position, there is in fact a continuum towards the other radical 'sibinv-are-fully-predictable' end, a continuum on which a modern framework such as Evolutionary Phonology (Blevins 2004) places itself approximatively halfway by seeing inventory structure as a consequence of recurrent sound changes and perceptual reanalysis rather than as a direct consequence of phonetic-phonological principles encoded in the synchronic grammar. In the second part of this subsection (4.1.1.2), I argue that the possible structures of sibinvs are explained by universal synchronic principles, while the typological frequency of certain patterns is explained both by universal principles *and* the frequency of recurrent sound changes.

4.1.1.1 The 'non-randomness' of sibilant inventory patterns

Given a number of PoAs between 0 and 4 (inclusive), 2 possibilities for voicing contrasts (voiced vs. voiceless), 2 possibilities for manner contrasts (fricative vs. affricate) and the fact that a sibilant can either be present or absent in the inventory, there should *a priori* be 65536 logically possible inventories:

$$2existencevalues^{4PoAs\cdot 2voicingvalues\cdot 2mannervalues} = 2^{4\cdot 2\cdot 2} = 2^{16} = 65536$$

Despite the diversity of sibilant inventories (84 different attested inventories with 7 distinct PoAs in the SibInv, 75 with 4 distinct PoAs), the attested patterns present in the SibInv database represent only 0.11% of the 65536 logically possible inventories. The very limited number of different attested inventories, compared to the wide range of logically possible ones, hints that inventory structure is far from random (Flemming 2008*b*: 2; Mackie and Mielke 2011; Dunbar and Dupoux 2016), but rather follows a set of universal principles. Mackie and Mielke (2011) demonstrate in particular that the structure of phoneme inventories differs significantly from what would be expected from chance, and argue in favour of feature economy (cfr. Clements 2003) as a principle that explains why certain patterns are much more frequent than others.

The 'non-randomness' of sibilant inventories is clearly reflected in the SibInv database. It is true that since there are only 258 sibinvs in the database, this is necessarily a very small portion out of 65536 logically possible inventories, but one observes that the number of *different* attested sibinvs (84) constitutes only one third of *all* attested sibinvs in the database. This suggests that the attested sibinvs are clustered around a few frequent patterns, while many other patterns are rare or unattested. In other words, selecting 258 inventories randomly out of 65536 logically possible inventories would include extremely low chances of finding the same inventory twice (e.g. the 258^{th} inventory can have *at most* a 0.39% chance of being identical to another already selected inventory), but much more than finding the same inventory twice in natural languages, we in fact observe a few inventories existing over and over again, in totally unrelated languages. Figure LIV illustrates how a few sibilant inventories are very frequent, with

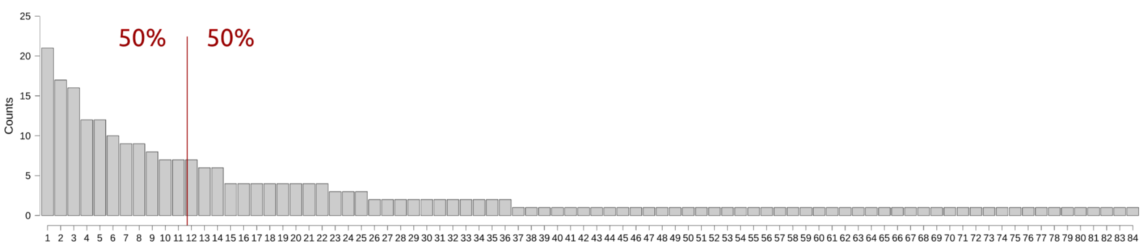


Figure LIV: *Frequency of each different attested pattern of sibilant inventories in the SibInv database (image generated in JASP Team 2021).*

e.g. the three most frequent sibinvs representing 20.93% of the entire database on their own, and the eleven most frequent representing 49.61%. The skewness in the distribution equals 1.394 (standard error: 0.152; a skewness of 0 indicates a symmetric normal distribution). Even more strikingly, one observes that the most frequent sibinv (21 counts) is perfectly symmetrical and well-dispersed:

(62)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ		
Sib. fri.		s z			ʃ ʒ		

The two next most frequent inventories (17 and 16 counts) consist of one with the same structure but without the front affricates /t͡s, d͡z/, and one with only a retracted alveolar voiceless sibilant fricative /s̠/. If there were no principles influencing the structure of sibilant inventories (e.g. symmetry and dispersion in the first inventory, affricate markedness and dispersion in the second, voicing and affricate markedness and centralisation in the third), the probability of finding with such high frequencies these well-dispersed, centralised and symmetrical inventories would be extremely low. Nothing would furthermore prevent e.g. the following inventory to exist, and to be (why not?) as frequent as the others:

(63)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.		t͡s d͡z		d͡ʑ			
Sib. fri.						ṣ	

To quantify precisely the 'non-randomness' of sibilant inventories, I replicated the study in Mackie and Mielke (2011) with the 236 sibinvs in the SibInv database,[2] generating 236 other random inventories and comparing how often one and the same inventory occurred in each set. The random inventories were generated based on exactly the same criteria as those for the SibInv inventories: 4 possible values for sibilant fricatives (no sibilant, a voiceless sibilant, a voiced sibilant, or both voiceless and voiced sibilants) at 7 PoAs (since phonetic detail includes up to 7 PoAs in the SibInv), multiplied by the same for sibilant affricates, i.e. a total of 268435456 (268 million) possibilities. Unsurprisingly given this large number, each of the 236 randomly generated inventories turned out to be a unique inventory. Comparing the frequencies of occurrence of each randomly generated and each SibInv inventory resulted in a highly significant difference between both sets ($F = 281.910$, $p < 0.001$): in the SibInv set, the mean frequency of occurrence of each inventory was higher (7.364) than the same number for the random set (1.000, i.e. only unique inventories).

[2] The statistical comparison of attested and random inventories was in fact performed at a stage of this dissertation at which there were 236 rather than 258 inventories in the database. Starting from here, the statistical comparisons are reported for 236 sibinvs, except that they were performed again for the 1 result (number of PoAs, in 4.2.1.1) whose p-value was not below 0.001 with 236 sibinvs.

Replicating this calculation by keeping, as the *only* distinction, 4 PoAs (i.e. getting rid of phonetic detail, voicing distinctions and fricative vs. affricate distinctions), the total number of logically possible inventories amounts to 16 instead of 268435456. With this criterion, there are 7 different types of sibilant inventories in the SibInv. Even with this calculation, the difference between the frequencies of random and SibInv inventories is statistically significant ($F = 435.161$, $p < 0.001$), again with recurrent patterns being much more likely to occur in the SibInv set (e.g. the well-dispersed 2s then occurs 129 times, i.e. in 50% of all SibInv inventories; mean number of identical inventories: 74.271) than in the random set (e.g. the most frequent pattern, an unattested 'overdispersed' 2s with /s/ - /s̺/, occurs 23 times, i.e. in 9.75% of all random inventories; mean number of identical inventories: 16.305).

The very asymmetrical statistical distribution of sibilant inventories in the SibInv database (a few inventories being excessively frequent), especially considering that the most frequent inventory is a well-dispersed, perfectly symmetrical and gapless 2s, evidences that a set of principles must be at work in shaping sibinv structure.

4.1.1.2 The need for universal principles rather than mere sound changes

Evolutionary Phonology (Blevins 2004: i) attempts to explain "the problem of how genetically unrelated languages, as far apart as Native American, Australian Aboriginal, Austronesian, and Indo-European, can often show similar sound patterns", i.e. to explain in this case the 'non-randomness' of inventory structure demonstrated previously, by means of "a formal model of sound change incorporating misperception and variable articulation [that] can account for attested sound systems without reference to markedness or naturalness within the synchronic grammar". According to Evolutionary Phonology (Blevins 2004: 281), there is thus no need for universal principles expressed as markedness, but these principles can in fact only be the observed consequences of *emergent* phenomena instead:

> [M]arkedness constraints play no role in determining the direction of sound change. In opposition to Grammont (1933), Martinet (1955), and more recent proposals within Optimality Theory (e.g. Cho 1998; Anttila and Cho 1998), sound changes which appear to be driven by functional or structural properties of sound systems are typically either illusory, accidental, or emergent.

Having ruled out illusion and accidentality as explanatory factors for the structures of sibilant inventories in 4.1.1.1, only emergence remains to shed light on the reasons behind the recurrent patterns observed.

Note that in this dissertation, as previously explained in footnote 16, I intend to use the terms 'teleological' and 'emergent' similarly to their use in Boersma and Hamann (2008). This means that 'teleological' refers to a principle encoded in the synchronic grammar of the speaker (in OT, in the form of a constraint), as opposed to an 'emergent' principle, which arises grammar-externally. My use of the term 'teleological' does thus certainly not imply that native speakers would be conscious at any moment of the principles that are posited to influence sibilant inventories.

Emergence *can* in fact explain certain principles, such as e.g. the symmetry between sibilant fricatives and affricates (see 4.7). There is indeed no need to posit a hypothetical OT constraint *Asymmetry, penalising sibilant affricates that do not have a corresponding fricative at the same PoA and inversely; recall from 3.2.3 that gaps, and thus asymmetries between fricatives and affricates, occur at any PoA (with very few unattested gap possibilities). To give an example of emergent, non-teleological symmetry, English /ʃ/ and /t͡ʃ/ are symmetrically aligned because they are the historical result of the palatalisation of /sk/-clusters to [ɕ] (e.g. Old English *Englisċ* > *English*), as explained in B.3, and of velars to [t͡ɕ] (e.g. Old English *ċīese* > *cheese*), as explained in A.1, respectively, followed by depalatalisation as detailed in B.6. To give an example with voicing, French [s] and [z] have the same PoA because the latter historically arose from the former in intervocalic contexts, and French [ʃ] and [ʒ] have the same PoA because they originate e.g. in the palatalisation of Vulgar Latin ca and ga respectively, i.e. from already symmetrical segments. Symmetry can emerge among others from the fact that segments undergo parallel sound shifts, and thus have predictably symmetrical outputs.

However, recurrent sound changes cannot explain everything about the structure of sibilant inventories and, as I will show here, a typology based on Evolutionary Phonology is in fact predicted to *overpredict*. Assuming that “recurrent synchronic sound patterns have their origins in recurrent phonetically motivated sound change” and that, as a result, “there is no need to directly encode the frequent occurrence of these patterns in synchronic grammars themselves” (Blevins 2004: 8), one struggles to explain e.g. why sibilant inventories with only 1 PoA have a central sibilant [s̠] rather than anything else, if nothing in the synchronic grammar favours central sibilants. The recurrent pattern observed cross-linguistically of having a single [s̠] (Martinet 1955; Adams 1975; Boersma and Hamann 2008; Vijūnas 2010) poses a problem to this theory, since it is not necessarily the consequence of a recurrent sound shift: in Finnish, it is the consequence of the dessibilation of historical *ʃ, so we observe [s] > [s̠] rather than [s] remaining alveolar despite being the only sibilant. Independently from recurrent sound shifts, centralisation operates visibly as soon as there is one single PoA in the inventory. Evolutionary Phonology would overgenerate, because it would e.g. allow a language like Finnish to have lost *ʃ and to continue having a single alveolar [s]. Applying recurrent sound shifts starting from a language like Eastern Norwegian, one could obtain the ‘crazy’ inventory with /t͡s, d͡z, d͡ʑ, ʂ/ in (63), which is unattested and very likely to be impossible: [s] could become voiced intervocalically, which is a frequent sound pattern; [s] and [z] could become [t͡s] and [d͡z] after coronals, which is also frequent; [s] and [z] could then become dessibilated to [θ] and [ð], and [g] could then become [d͡ʑ]. This sequence of changes is composed of probable and recurrent sound shifts, but would still be prohibited in the OT typology of this dissertation (for example, m.Periph(Sib) would have shifted [ʂ] to [s̠]). In the case of Finnish and other languages, we have to deal with a sound change that appears “to be driven by functional or structural properties of sound systems” (Blevins 2004: 281) *and* is neither accidental nor emergent, because it cannot be anything else than encoded in the synchronic grammar.

Turning now to frequency inside the inventory rather than its structure, one observes that certain frequencies are best explained by teleological principles expressed as OT constraints, while others are emergent and non-teleological. For example, the higher frequency of voiceless sibilants wrt. voiced ones is the result of voicing markedness (formalised as the constraint M.VOICE(SIB)), but the higher frequency of [t͡ʃ] and [d͡ʒ] wrt. other affricates is an emergent, non-teleological consequence of the fact that stop palatalisation is one of the most frequent sound changes (targeting nearly 50% of the languages in Bateman 2008: 26), and that its outputs [t͡ɕ] and [d͡ʑ] are most likely to be reanalysed as /t͡ʃ/ and /d͡ʒ/ in a well-dispersed 2s, the most frequent sibilant inventory (and also in a non-retroflex 3s or 4s). One thus needs both teleological principles and non-teleological explanations to account for statistical frequencies inside sibinvs.

4.1.2 The sensorimotor plausibility of the principles

Another argument in favour of the principles is that their existence is made plausible by the way human speech works. For example, it has been shown that gravity affects the way humans articulate vowels (Shamei and Gick 2019), since astronauts who return to earth after a considerable time in outer space appear to have lower, centralised vowels with respect to before their mission. Such evidence for ease of articulation is especially valuable to the extent that it shows how it is *synchronically active*, having direct effects in a short lifespan while most evidence traditionally came from sound change across generations.

Posited markedness constraints in OT that are grounded in speech production are thus likely to really exist in grammar, since they correspond to plausible articulatory phenomena (Hayes 1999; Hayes and Steriade 2004). Speakers themselves are aware that some sounds or sequences of sounds (e.g. [tkfrk]) are more effortful than others, and make use of this knowledge in speech production. The more a posited constraint 'makes sense' from an articulatory perspective, the more it is thus likely to exist in grammar.

4.1.3 The predictive power of the principles

Anticipating the results of this dissertation (see section 6.1), one can consider *a posteriori* that the fact that a typology based on the principles in this chapter yields a typology that matches the attested languages is an argument in favour of these very principles. A crucial aspect is furthermore that the predicted inventories match the attested ones *and* that the predicted inventories are still an extremely small proportion of all logically possible inventories. This makes it unlikely that the predicted typology would match the typological data only because it overgenerates in an 'everything is possible'-like manner, and thus that the match between predicted and attested would just be a consequence of excessively wide and vague predictions.

4.2 Number of PoAs and sibilant markedness

Now that the need for principles to explain the attested sibilant inventory structures has been evidenced, the rest of this chapter induces those principles starting from the observation of inventories in natural languages. Evidence for each principle is provided, based mainly on cross-linguistic as well as intra-linguistic statistical generalisations, error patterns in language acquisition, diachronic developments and the comparison with randomly generated sibilant inventories.

4.2.1 Cross-linguistic evidence

The largest attested number of places of articulation for sibilants is 4, with the Indian language Toda and the Caucasian language Ubykh exemplifying this largest possibility (Ladefoged and Maddieson 1996: 156-159, Hall 1997*b*). Interestingly, these PoAs correspond both in number and articulatory configurations to the four PoAs of the largest attested coronal inventories (Hall 1997*b*: 87, Bundgaard-Nielsen et al. 2015). These PoAs are: denti-alveolar, retracted alveolar, alveopalatal/palatoalveolar and subapical retroflex.[3]

(64)

s̪	s̠	ʃ	ʂ̠
t̪	t̠	c	ʈ̠
n̪	n̠	ɲ	ɳ̠

The results of the learnability experiment in section 2.3 provisorily confirm that four is the largest number of contrastable places of articulation for a 'normal' population (i.e. more than 4 is possible for certain individuals, but this is levelled out by the inability of the many others to master more than 4). There is thus a language universal about the possible PoAs of sibilants stating (Hall 1997*b*: 93):

(65) Languages have between 0 and 4 contrasting places of articulation for sibilants.

And indeed, all languages in the UPSID, the PHOIBLE and the SibInv databases fit into this 0-to-4-PoAs mould, as illustrated in figure LV. Nevertheless, there is a discrepancy between the logically predicted statistical frequencies of inventories according to number of PoAs and the attested frequencies. If all PoAs were equally likely to have sibilants or not, with no reason for any PoA to have more or less than a 50% probability to have a sibilant, there could be 16 inventory patterns, out of which 1 does not have sibilants, 4 have one sibilant, 6 have two, 4 have three and 1 has four. Instead of looking like LV, the expected statistical distribution per number of PoAs should thus look like figure LVI. There is thus an asymmetry between the two tails of the normal distribution: the left tail (i.e. 0s and 1s) are more frequent than the right tail (3s and 4s), totalling 39.47% on average in the three databases versus 8.13%, respectively. In all three databases, 1s are more frequent than 0s, 3s and 4s; in two out of three databases (UPSID and PHOIBLE), 0s are more frequent than 3s; in all three databases, 0s are more frequent than 4s; and in all three databases, 3s are more frequent than 4s. Statistically speaking, this results in a

[3] As elsewhere in this dissertation, the term *subapical retroflex* refers to a coronal segment articulated subapically with a bunched tongue at the most postalveolar PoA, rather than the plain retroflexes present in e.g. Russian, Polish or Chinese (see Hamann 2003: 18f. for a description of the difference).

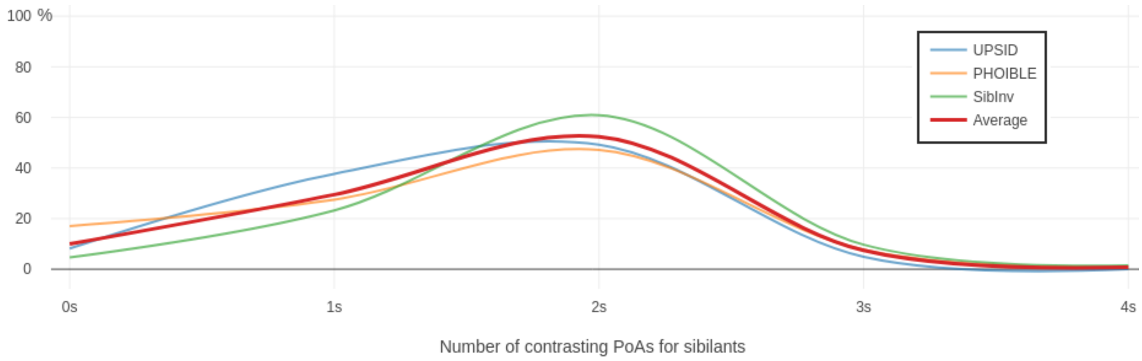


Figure LV: *A graphical representation of the frequency of sibilant inventories as a function of number of PoAs (with the numbers of fig. XL and additional curve smoothing).*

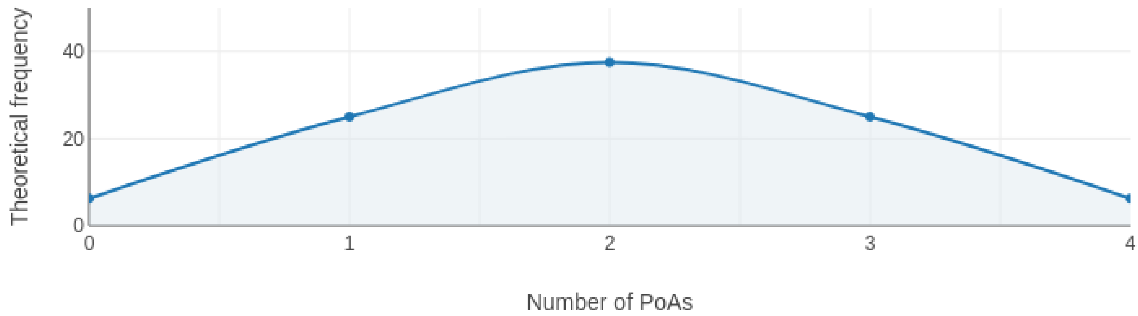


Figure LVI: *A graphical representation of the expected frequency of sibilant inventories as a function of number of PoAs, if no PoA had more reasons to have or not to have sibilants.*

skewness value of -0.144 for the distribution of the SibInv (standard error: 0.152). The right tail of the distribution is thus flatter than the left tail, meaning that having more PoAs than a 2s is 'worse' for a language than having fewer than 2.

Besides the universal stating that languages have between 0 and 4 PoAs for sibilants expressed in (65), there is thus a statistical tendency towards fewer PoAs:

(66) Languages are more likely to contrast fewer than 2 PoAs than to contrast more than 2 PoAs.

4.2.1.1 Comparison with random inventories

This preference for fewer PoAs in natural languages is confirmed by the comparison between randomly generated and SibInv inventories. If one assumes that 'anything goes' with respect to the number of PoAs in sibilant inventories and that there is no universal principle that influences the distribution of PoAs in sibinvs, one should: 1) find 5s, 6s or even larger inventories, which are unattested in natural languages; 2) be as likely to find 4s inventories as to find 0s if one assumes 4 possible PoAs, because there is logically 1 chance out of 16 to have no sibilant at all as well as 1 chance out of 16 to have 4 contrasting PoAs for sibilants. As explained above, a randomly generated set of sibinvs with maximum 4 PoAs is likely to have around 1/16 chance of having a 0s, 4/16 chances of having a 1s, 6/16 chances of having a 2s, 4/16 chances of having a 3s and 1/16 chance of having a 4s. I created a pseudo-random set consisting precisely of a proportion of 1/16 of 0s, 4/16 of 1s etc. predicted by *a priori* probabilities, with a total number of 256 rather than 258 inventories (because it had to be divisible by 16 to obtain integer numbers of inventories). Besides this, I also generated a truly random set of 258 sibinvs, each with a random number of PoAs between 0 and 4 according to the same probability. Comparing both sets with that of the SibInv inventories, it is indeed observed that the statistical distribution of PoAs in the SibInv database is significantly different both from the pseudo-random set (mean number of PoAs: 1.802 vs. 2.000; $t = 2.551$, $p = 0.011$) and from the randomly generated set (mean number of PoAs: 1.802 vs. 2.089; $t = 3.666$, $p < 0.001$). This difference is also highly significant in the UPSID and the PHOIBLE ($p < 0.001$ in both).

4.2.2 Motivations for sibilant PoA and markedness universals

There must be a principle favouring fewer places of articulation for sibilants, rather than this asymmetry towards fewer PoAs being a coincidence. If it were coincidental due to e.g. an accidental asymmetry in the chosen sample of languages, the distributions of PoAs would not be significantly left-tailed in *all three independent* typological databases, considering especially that one of them (the UPSID) is genealogically and areally balanced. In this subsection, I argue that this bias towards fewer PoAs is due to the markedness of sibilants (in accordance with e.g. Ueda 1996, Kubozono 2015: 683 and Sebregts 2014: 134). I furthermore posit that such a principle is encoded in grammar, in the form of a markedness constraint M.SIB. With respect to the reason why 4 is the largest possible number of PoAs, however, I argue that this emerges from perceptual acuity restrictions in normal populations and is not encoded in grammar.

4.2.2.1 Teleological or emergent?

The question whether certain linguistic phenomena are teleological or emergent has been the object of longstanding debate (Prince and Smolensky 1993/2004; Lindblom 1999; Blevins 2004; Boersma and Hamann 2008; van de Weijer 2017). *Teleological* phenomena are considered to reflect drives that language users directly experience when speaking (e.g. the articulatory complexity of certain segments wrt. others). They are encoded in grammar, more precisely as constraints in the case of Optimality Theory. By contrast, *emergent* phenomena are phenomena that arise during the interaction between speakers and listeners, are not a part of grammar and not something which language users aim to do at all (Boersma and Hamann 2008: 218).

In the case of the universal about number of PoAs formulated in (65) and the statistical tendency formulated in (66), I argue that the former is emergent, while the latter is teleological on the side of production and emergent on the side of perception. Speakers do not actively try to reduce the number of PoAs itself to less than 5, but this number tends instead to reduce itself diachronically because of mergers and dessibilations which are themselves phenomena encoded in the production grammar. This means that there are e.g. no constraints *5PoAs, *4PoAs etc. in the OT grammar of sibilant inventories, because such constraints directed towards the *system* rather than segments as candidates would imply that language users evaluate entire inventories when speaking, which is far from realistic. As will be shown in 5.1.1, "there is no need to posit a dedicated inventory grammar: inventories are the automatic result of the constraints and their rankings in the production grammar" (Boersma 1997*b*: 59).

The fact that natural languages have between 0 and 4 PoAs for sibilants is thus not a teleological result of something in grammar, but a consequence of the fact that 4 is the average largest number of PoAs that can be distinguished well enough to allow for successful communication in a human population. The fact that natural languages tend to have 2 or fewer PoAs for sibilants is a teleological consequence of the fact that markedness constraints in the production grammar tend to penalise the presence of sibilants, as explained in the two following subsubsections.

4.2.2.2 Articulatory motivations

Each place of articulation contrasted in a speaker's mother tongue represents an additional combination of tongue movements that has to be learned and mastered during language acquisition, and sibilants are particularly complex segments to articulate (Baroni 2014*b*: 34). It therefore represents a burden for speakers and results in the fact that the contrasts are preferably maximised at a smaller number of places of articulation (e.g. by also having affricates and voicing pairs, see 4.7), rather than having a few contrasts at many places of articulation. This is generally known in Distinctive Feature Theory as feature economy (Clements 2003). Errors in the production of the different PoA gestures and misperceptions in speech perception both result in the reanalysis of a sibilant segment as having another existing PoA or as being a non-sibilant segment, and this is the more likely to happen the more PoAs for sibilants are contrasted, because more categories implies a smaller permissible variation on the continuum (Boersma and Hamann 2008: 222; Sieber 2020). Mergers and dessibilation thus continuously reduce the number of places of articulation, but especially so in larger inventories, so that languages tend towards the typological '2s optimum' of figure LV.

Now, how does one account *formally* for the fact that not all languages have sibilants, and that they tend towards fewer PoAs for sibilants? In Optimality Theory (Prince and Smolensky 1993/2004), one would posit the existence of a constraint m.Sib that penalises any sibilant. Those who disagree with this would claim instead that synchronic grammars do not actively prohibit sibilants; however, languages like Hawaiian provide evidence for the opposite. Recall from 3.2.2.1 that besides the absence of sibilants, a foreign sibilant surfaces as /k/ in Hawaiian. There is thus something in the synchronic grammar of Hawaiian that prohibits sibilants and therefore, we need to posit such a teleological principle e.g. as the constraint m.Sib. This mapping of foreign sibilants to stops when sibilant markedness dominates is a manifestation of the universal that (sibilant) fricatives are more marked than stops (Cairns and Feinstein 1982: 208; Fletcher 1989: 746; Kirchner 2001; Boersma and Hamann 2009: 14f.; Maddieson 2013*a*).

Even within languages that do have sibilants, the asymmetry towards fewer PoAs persists: the statistical frequency of 1s inventories is still between two and seven times as large as the *combined* frequencies of 3s and 4s, depending on the database. This has to do with other markedness constraints that are detailed further in this chapter (centralisation and posterior markedness). Each of them penalises certain sibilants (peripheral and posterior ones, respectively), resulting in a low average number of PoAs per inventory. Applying the method of calculating the ranking volume of individual languages (Anttila 1995; Coetzee 2002; Riggle 2010) in the OT typology I propose in 5.2, it turns out that the average number of sibilant PoAs resulting from all possible constraint rankings is 1.367, i.e. even lower than the attested average number of PoAs. This however is due to an improbably high frequency of constraint rankings yielding a 0s, as witnessed by the frequency graph in LVII. The ranking volume thus gives a predicted frequency in

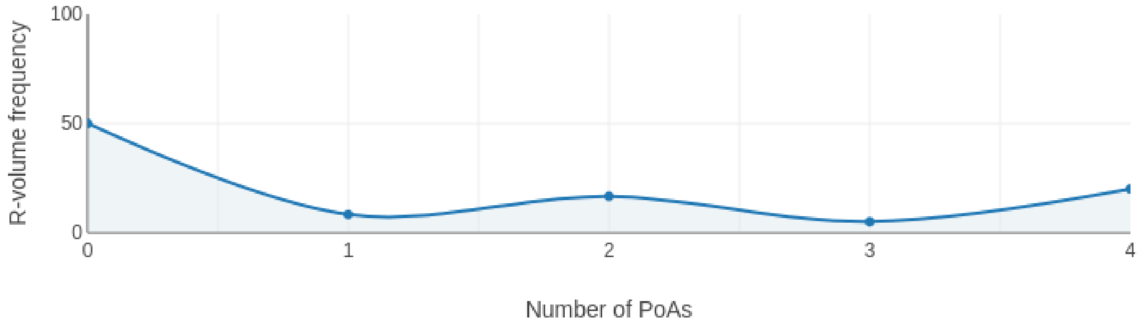


Figure LVII: *A graphical representation of the frequency of sibilant inventories as a function of number of PoAs according to the ranking volume of the OT typology in 5.2.*

favour of sibinvs with fewer than 2s, but with such a mismatch with respect to attested frequencies per number of PoAs that the validity of a comparison becomes questionable.

4.2.2.3 Perceptual motivations

As the learnability experiment in 2.3 has shown, more than four CoG distinctions (transposed to the ERB scale) results in an acoustic space that is too crowded to guarantee a normal (i.e. around 70-80%, cfr. Billings et al. 2016) correct phoneme identification rate. The excessively numerous confusions with neighbouring phonemes in a 5s would thus soon lead to a merger and thus, to at most a 4s. The fact that four is the maximum attested is an emergent consequence of the average human perceptual acuity: if everybody had a better hearing (such as e.g. the few participants in 2.3 with absolute pitch) on average, natural languages could potentially contrast more PoAs than four.

Besides the universal that languages have at most 4 sibilant PoAs, the fact that most languages contrast 2 or less PoAs for sibilants is due to the fact that the more categories are contrasted on a continuum, the smaller the permissible variation is and therefore, the more one and the same misarticulation is likely to result in a perceptually different category (Boersma and Hamann 2008: 222; Sieber 2020). In a 1s, any sibilant realisation is perceptually classified as the sole category /s/; in a well-dispersed 2s, a sibilant realisation must have been articulated very far from its prototypical realisation in order to result in misperception (because the auditory distance between [s] and [ʃ] is large). In a 3s, a smaller misarticulation (e.g. a slightly fronted [ʃ̻]) is able to result in misperception. The more PoAs a language thus contrasts, the more likely it will be to reduce these contrasts by means of misperception-induced mergers. One can thus summarise the motivations for this universal by saying that, so to say, 'less-skilled' speakers and 'less-skilled' listeners collaborate towards the reduction of PoA contrasts.

4.3 Centralisation

In this section and the next one, universals pertaining to specific places of articulation for sibilants will be addressed: one is centralisation, which favours centrally articulated sibilants (i.e. the least anterior or posterior possible), and the other is posterior markedness, which favours anterior sibilants.

4.3.1 Cross-linguistic evidence

A substantial part of the literature on sibilants, starting with remarks by Braune (1911) and followed by Joos (1952), Martinet (1955), Galmés De Fuentes (1962), Adams (1975), Boersma and Hamann (2008) and Vijūnas (2010), observed that the majority of the languages of medieval Western Europe had retracted alveolar sibilants, before observing in particular that this applies regularly to those which have only 1s, and finally extending this observation about 1s to the languages of the world in general. Evidence from historical linguistics and sound shifts (e.g. the Latin and North/West Germanic rhotacism, cfr. Kokkelmans 2017, 2020*a*) has lead to the insight that 1s languages like Latin, Proto-Germanic and Proto-Indo-European most likely had a retracted alveolar sibilant (Vijūnas 2010), like the modern languages with 1s (e.g. Icelandic). Besides Indo-European 1s languages with a retracted alveolar sibilant, Vijūnas (2010: 49) mentions Finnish, Old Korean, Samoan, Greenlandic and Canadian Inuit, as well as Kurux. However, he observes that a few languages do not stick to the pattern, for example Mexican Spanish (Vijūnas 2010: 43):

> The remaining two languages, Mexican Spanish and Catalan, indeed possess sibilant systems that do not conform well to Martinet's scheme (...). Both languages simply provide evidence that systems with not maximally distinguished phonemes or with a single dentialveolar sibilant can exist, at least for certain periods of time. It must be noted here that both systems are relatively recent, and the Mexican Spanish owes its alveolar *s* to the Andalusian dialects of Spanish (...).

I have shown in 3.2.2.3 that Mexican Spanish is in fact a 'hidden 2s', i.e. that it also has a (marginal but existing) phoneme /ʃ/, and thus that it does not constitute counterevidence for the centralisation universal. Catalan and Faroese are mentioned as examples of less-dispersed 2s inventories, and are thus counterevidence for the auditory dispersion universal (4.8). The universal about 1s inventories already observed in 3.2.2.2 thus still holds, and can be formulated more generally as:

(67) Sibilant inventories with 1 PoA have retracted alveolar/alveopalatal sibilants.

This universal is also reflected by the fact that in the attested types of sibilant inventories (summarised in table 2), the retracted alveolar/alveopalatal PoA slot is occupied in 8 out of 10 types of inventories with sibilants.

Furthermore, the very existence of the Catalan/Faroese type of 2s with a [s̠] - [ʃ] contrast is itself an argument in favour of centralisation, because it is an inventory that is ostensibly 'resistant' to both auditory dispersion (by having two acoustically close sibilant PoAs) and posterior markedness (by allowing [ʃ]). Without centralisation, there would thus be no reason for it to exist: attested 2s inventories would either be well-dispersed or have no postalveolars.

It is worth noting that centralisation does not favour only [s̠], but also [ʃ]. Both are anterior and posterior central sibilants, respectively, and the reason for which languages with 1s do not have [ʃ] as their sole sibilant is that [ʃ] is marked with respect to posterior markedness, as described in 4.4. A version of centralisation in which only [s̠] is favoured leads to the incorrect prediction that languages with [s̠] and [ʃ] like Catalan should not exist in the OT SibInv typology (cfr. 8.2.1.1). Additional evidence for the fact that centralisation also favours [ʃ] is provided by Hall (1997*b*: 95), who shows that in Caucasian languages with 4s and only two vowels, the vowels (which are subject to strong coarticulation depending on the surrounding consonants) are not realised as the least effortful central vowels [ə], [ɐ] when adjacent on both sides to anterior sibilants (but e.g. as an anterior /ɛ/ instead), which is however the case with [ʃ]. It thus seems that on the horizontal front - back axis, [ʃ] corresponds in centrality to [ə], which is an unmarked central vowel. The precise central PoA for sibilants is thus probably located right in between [s̠] and [ʃ], and as I show in the OT typology (8.2.1.1), this still predicts that [s̠] will be the single sibilant in a 1s. It is therefore unproblematic and even necessary (based on the Caucasian evidence) to assume that the least effortful place of articulation for sibilants on the horizontal dimension is located between [s̠] and [ʃ].

4.3.1.1 First language acquisition

Children acquiring languages with 2 or more PoAs are reported to produce tendentially central values for sibilants (closer to [s̠]) compared to adults (Fletcher 1989: 744; Zharkova et al. 2018: 1454, 1461), resulting in only "a small but stable ($p < .01$) difference between tongue contact positions for /s, z/ and /ʃ, ʃ', ʒ'/" (Fletcher 1989: 745). Furthermore, retracted alveolar and alveopalatal sibilants are acquired earlier than other sibilants: /ɕ/ is acquired earlier than /s̻/ in Japanese (Li et al. 2009) and earlier than the other sibilants in standard Chinese (Li and Munson 2016) as well as in Polish (Żygis et al. 2019). Ezeizabarrena and Alegría (2015: 99) mention mispronunciations of anterior /s̻/ in a Basque child, but I was not able to find any study for languages of the non-retroflex type of 3s. Very interestingly, figure 8 in Żygis et al. (2019: 1512)'s acoustic study shows how the three Polish sibilant PoAs are gradually distinguished from each other in production proportionally to the age of the children, all starting from a *middle* value located precisely at what approximatively corresponds to [s̠] and [ɕ]. When the age increases,

/s/ is pronounced with an increasingly higher CoG and /ṣ/ with an increasingly lower CoG. The difficulty that children experience in mastering both front and back sibilants wrt. central sibilants is thus an argument in favour of centralisation as an active phenomenon in the synchronic grammar.

4.3.1.2 Comparison with random inventories

The effect of centralisation on sibilant inventories in natural languages is visible in the difference with 236 randomly generated inventories. If one assigns one 'bad centralisation point' to an inventory for each sibilant that is *not* retracted alveolar or plain postalveolar (i.e. out of 4 PoAs, to the most front and the most back, irrespectively of voicing, affricates and phonetic detail), the mean 'bad centralisation score' is 0.695 for the inventories in the SibInv database and 1.068 for the same number of randomly generated sibinvs. This difference is highly significant ($t = 6.639$, $p < 0.001$). Given 4 PoAs, the likelihood of having 1 'bad centralisation point' is 8 out of 16 possibilities and that of having 2 'bad centralisation points' is 4 out of 16, while the most likely mean number of 'bad centralisation points' is 1. The fact that the random set generated is only 0.068 points distant from logical random probability but the SibInv set is statistically significantly lower in score (-0.305 points) demonstrates that centralisation is a principle that is actively shaping the structure of sibilant inventories. This is especially remarkable considering the frequency of well-dispersed 2s in the SibInv database (around 50% of the entire sample), i.e. inventories that have the frontmost PoA occupied by non-central sibilants and were therefore assigned 1 'bad centralisation point'.

4.3.2 Motivations for the centralisation universal

4.3.2.1 Teleological or emergent?

Centralisation is an articulatorily motivated principle that is teleological: speakers aim towards a lesser displacement of the tongue on the horizontal axis (from a central position to the front or the back, corresponding to an activation of the genioglossal or the hyoglossal muscle, respectively; Drake et al. 2009: 1039). In this dissertation, I posit that this sensorimotor knowledge is encoded in the synchronic grammar as a centralisation constraint penalising the activation of these two muscles in their respective direction.

4.3.2.2 Articulatory motivations

Peripheral sibilants (such as [s̪] and [s], as well as [ṣ] and [ṣ̠]) require more horizontal displacements of the tongue from a resting/neutral tongue position, proportionally to their peripherality. From early on (Martinet 1955), the observation that languages with 1s have a retracted alveolar sibilant has been explained as a manifestation of articulatory economy (also called 'ease of articulation'), a view that is still unchallenged and has been developed in the recent literature (Boersma and Hamann 2008: 231):

> Sibilants with central spectral mean values are easier to produce than those with peripheral values (...) Independent evidence for this comes from comparing the muscle activities required for producing central and peripheral sibilants. Sibilants with central spectral mean values like the Dutch flat laminal alveolar [s̻] or the Spanish apical alveolar [s̺] are produced by a simple raising of the articulators towards the roof of the mouth, without displacement or grooving of the tongue, and are thus as close as sibilants can get to the rest position of the tongue or to the average position of the tongue during vowels. At the high periphery of the spectrum, the deep grooved alveolar sibilant (as in English) has a similar position to the Dutch non-grooved sibilant but requires additional activity of the upper fibres of the transverse tongue muscle (Hardcastle 1976: 96, 100–106, 134–137); at the low periphery of the spectrum we note that the apical palatal (retroflex) sibilant (as in Toda) has a similar tongue shape to the Spanish anterior sibilant but requires a larger movement of the tongue tip from a schwa-like position towards the palate and a stronger involvement of the upper longitudinal tongue muscle.

This is in accordance with Hamann (2003: 4)'s observation that e.g. "retroflexes are more marked than apical alveolars (or dentals), since retroflexes involve a raising and displacement of the tongue tip towards the post-alveolar region, whereas an apical alveolar [JK: $\simeq$ retracted alveolar] involves only a tongue tip raising". While all sibilants require raising the tongue towards the palate or the alveola to some extent, the sibilants that require the least additional horizontal displacement are favoured. The genioglossus and hyoglossus, the main muscles responsible for pulling the tongue towards the front and the back of the mouth respectively, require an amount of energy proportional to the displacement, so that the more peripheral the sibilant, the more energy required and thus the more penalised by the centralisation principle.

As detailed in the following section, the articulatory economy principle related to the *raising* of the tongue itself is encoded as a distinct principle in the grammar (namely posterior markedness).

4.3.2.3 Perceptual motivations

Perception plays no role in favouring central sibilants. Since humans perceive lower frequencies with more accuracy (Moore and Glasberg 1983; Evers et al. 1998; Møller and Pedersen 2004; Ooijevaar 2011: 4), posterior sibilants should be favoured, but since anterior sibilants have more energy at higher frequencies (Jongman et al. 2000: 1253; Shinn 1986: 19), they are more easily distinguishable from the non-sibilants. There is thus no advantage of being a central sibilant with respect to perception.

4.4 Posterior markedness

This section argues in favour of posterior markedness, i.e. the fact that posterior sibilants are more marked than anterior ones, based on six types of evidence (cross-linguistic frequency, intra-linguistic frequency, first language acquisition, assimilation patterns, reduplication and diachronic fronting).

4.4.1 Cross-linguistic evidence

In the literature, several researchers observe that anterior sibilants are less marked than posterior sibilants (or: 'postalveolar', a term which I use here interchangeably with 'posterior'). For example, de Lacy (2002: 426, 461) states in his dissertation on markedness that "[-anterior] is a marked feature: all inventories with [-anterior] segments also have [+anterior] ones". More specifically, he remarks that every language in his sample that "has a [-anterior] coronal also has a [+anterior] coronal, regardless of the manner of articulation", something which is also observed by Hall (1997*b*: 86f.). Besides the observation that the presence of posterior sibilants implies that of anterior ones in a language, a relevant typological generalisation is that languages with 1s logically do not have posterior sibilants. Vijūnas (2010: 49), for example, states that "languages normally do not possess [ʃ] as their only sibilant".

As a side note, this preference for the front is also found in asymmetric vowel inventories: when asymmetries are found in vowel inventories, these inventories consist overwhelmingly more frequently of more anterior vowels than posterior ones (Tent 1993: 356ff. giving 7 examples, to which one could add Etruscan and Lardil). More generally in consonants, the "dominance of consonants that are articulated in the front of the mouth as against back consonants" (Franklin and Franklin 1962: 29) is found in many languages as well. It is however unclear whether this asymmetry in favour of anterior segments is triggered by similar mechanisms in the cases of vowels, consonants and sibilants (probably not for vowels, cfr. Tent 1993: 361).

Flynn (2012: 106) suggests that "[posterior] is especially marked in combination with [fricative], suggesting a contraint *[posterior, fricative]", as also recognised by Clements (2009: 34). Anterior sibilants are widely assumed to be "unmarked" (Ferguson 1990: 62) with respect to "retroflexes and palato-alveolars" (Hamann 2003: 140), although much of the literature does not attempt to formulate explanations for this observation.

4.4.1.1 Cross-linguistic frequency

In the languages of the world, posterior coronals are rarer than anterior coronals. For example, a search in the database of Eurasian phonological inventories (Nikolaev et al. 2015) yields 379 search results for the term 'alveolar', against 240 for 'postalveolar'. This predominance of anterior coronals is verified for sibilants in particular, and even more strongly in sibilant fricatives. In the PHOIBLE 2.0 database (Moran and McCloy 2019), /s/ (which is used for both dental and retracted alveolar voiceless sibilants in most of its language descriptions) is by far the most frequent sibilant in the world's languages (86%), followed by /t͡ʃ/ (57%), /ʃ/ (41%) and /z/ (40%). The anterior sibilant fricatives [s, z] are more than twice as frequent as their plain postalveolar counterparts [ʃ, ʒ], not to speak of the even much rarer other postalveolar fricatives. A notable exception to this preference for the front is the widespread plain postalveolar affricate /t͡ʃ/ (cfr. González 2003: 12),

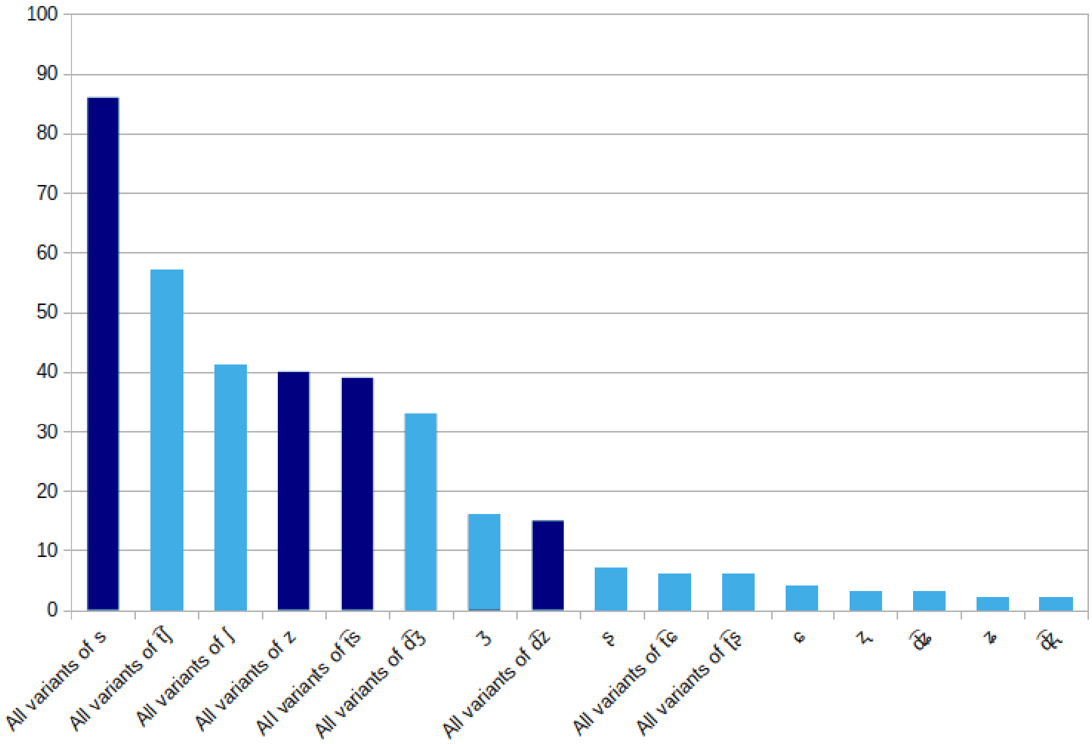


Figure LVIII: *Typological frequency of single sibilant phonemes in the PHOIBLE database (Moran and McCloy 2019). Sibilants with a frequency below 1% have been omitted. Dark blue corresponds to anterior sibilants and light blue to posterior sibilants. A detailed version of this figure with all variants below 1% frequency can be found in the appendix (CXXII).*

explained in the next paragraph. Regardless of this, the average frequency of all variants of alveolar sibilants in figure LVIII is 45%, compared to only 15% for postalveolar sibilants. Interestingly, the alveolar sibilants are much less numerous than the postalveolar ones (4 against 12 in fig. LVIII), but are still thrice more frequent.

In sibilant affricates, posterior markedness is not reflected as higher cross-linguistic frequencies; in fact, /t͡ʃ/ is even more frequent than /t͡s/ (Berns 2013: 82). This can be explained by the cross-linguistic commonness of stop palatalisation (Bateman 2008: 26): while a /t͡ʃ/ is very likely to appear over and over again diachronically from palatalised stops of any kind (stops and [i, j] being all very frequent), sound shifts that yield /t͡s/ are much less frequent (e.g. [tʰ] > [t͡s], considering that [tʰ] is a rarer sound). It is thus an emergent phenomenon rather than anything in the grammar favouring posterior sibilant affricates.

An argument in favour of this emergent frequency of /t͡ʃ/ is the fact that inventories with anterior affricates without posterior affricates do exist (e.g. Younger Malagasy). If something in the grammar prohibited anterior affricates, there would be a typological implication in the sense that either inventories with posterior ones only, or with both, would exist. Another argument is the fact that this statistical predominance of /t͡ʃ/ over /t͡s/ does not necessarily hold in Eurasia, where Nikolaev and Grossman (2018: 566) report 54% frequency for the former and 52% for the latter.

Another possibility is the fact that postalveolar segments are perceptually more distinct from anterior segments if they are [+strident], and thus that auditory dispersion favours posterior *sibilant* affricates (Flynn 2012: 105):

> As Clements (2009: 50) explains: "[±strident] can be understood as a feature that enhances the acoustic properties of (...) posterior sounds (such as palatoalveolars), in the sense of Stevens et al. (1986). It enhances ... posterior obstruents by making their characteristic lower frequency noise component in the region of the third formant more audible. ... The addition of [+strident] to a posterior stop increases its auditory distance from a nonstrident anterior stop such as /t/.
>
> As for [±strident], the high-frequency spectral energy of [+strident] sounds exceeds that of neighboring vowels, while the spectral energy of [−strident] sounds is lower than that of neighboring vowels at all frequencies (...)."

A combination of both explanations can be that posterior sibilant affricates are easily created by stop palatalisation, and that the systemic pressure in the inventory favours their maintenance as posterior affricates. In every case, it is crucially an emergent result rather than a teleological principle against anterior affricates, and cannot function as such within the production grammar.

One shall also observe that the perceptual advantage of posterior sibilant affricates over non-sibilant posterior stops, described in the quote by Clements (2009: 50) above, is not absolutely satisfied in the languages of the world: some languages have posterior stops but no posterior sibilant affricates, regardless of whether they allow posterior sibilants (e.g. Eastern Norwegian) or not (e.g. Australian languages). It might thus be an explanation for a statistical tendency, but not an absolute universal (e.g. [t͡ç] > /t͡ʃ/ more frequently than [t͡ç] > /t͡s/, but not necessarily always, as shown in A.1). It is thus an emergent tendency, formalised in the perception grammar presented in 5.2.3.

4.4.1.2 Intra-linguistic frequency

Cross-linguistically, front consonants (labials, anterior coronals) are more frequently used in speech than back consonants (posterior coronals, velars, uvulars, pharyngeals) in the same language (Everett 2018). For example, the former category is almost thrice as frequent as the latter category in French (Malécot 1974: 164). This tendency is confirmed in sibilant fricatives, as [s] is more frequent than [ʃ] in 28 out of 33 languages in the sample of Peust (2008: 126). A striking example in this sample is Basque, in which the reported frequency of a sibilant decreases linearily from front to back (namely 79 [s̻] > 40 [s̺] > 2 [ʃ], Peust 2008: 108f.). Everett (2018) reports similar results independently, with [s] being more than twice as frequent as the second most frequent sibilant, [t͡ʃ]. The summed frequency of anterior sibilants in the 34 most frequently used segments he reports is 0.0355 versus 0.0241 for posterior sibilants, i.e. 1.47 times higher.

4.4.1.3 First language acquisition

Children struggle to pronounce posterior sibilants for a longer period than that required to master anterior sibilants (Stemberger and Bernhardt 1997: 299f.; Flynn 2012: 106). For example, in standard Polish, [ʂ] is acquired latest with respect to [s] and [ɕ] (Bukmaier and Harrington 2016: 312). After having acquired the ability to pronounce sibilants in general, they universally experience a period of postalveolar fronting (e.g. saying *Englis* instead of *English*). This 'fronting universal' was originally formulated by Jakobson (1968) and has been reported e.g. in German (Fox and Dodd 1999), Jordanian Arabic (Yaseen and Mahadin 2018), Italian (Zmarich and Bonifacio 2005; Romani et al. 2017), French (Lemieux 2011) and English (Fletcher 1989; Goss-Grubbs 2007; Li et al. 2009). The same phenomenon has been observed in individuals with Down syndrome speaking Dutch, English and Kannada (van Borsel 1988; Rupela et al. 2010), although it is uncertain whether other physical factors might interfere in this case. An important caveat is that the order of acquisition is influenced by language-specific segment frequencies (Li and Munson 2016), so that a very frequent marked segment can be acquired earlier than predicted, interfering with universal sensorimotor constraints.

Children's sibilants are not only more central at first, as evidenced in 4.3.1.1, but tend to be fronted as soon as the front sibilants are acquired, something which is reflected in Li et al. (2009)'s observation that /ɕ/ is mastered earlier than /s̠/ in Japanese, while /s/ is mastered earlier than /ʃ/ in e.g. English and Dutch (van 't Veer 2015: 115). This is explained by both centralisation and posterior markedness: /ɕ/ is easier to acquire than /s̠/ and /s/ because it is central *and* not posterior, while /s̠/ and /s/ are easier to acquire than /ʃ/ because they are not posterior. Noteworthily, the reason for which /ɕ/ in particular is mastered earliest is that "children have greater tongue-palate contact in their speech than adults" (Richtsmeier 2010: 7): a lack of sensorimotor precision results in the fact that children produce tendentially more alveopalatal sibilants (Fletcher 1989), something which is independent from posterior markedness. The evidence is thus more reliable in the case of the /s/ > /ʃ/ asymmetry.

4.4.1.4 Assimilation patterns

As evidenced in B.12 and B.13, both long-distance and direct (or: local, contact) assimilation between sibilants present a bias in favour of posterior sibilants. Assimilation is either directional (progressive or regressive), or non-directional and then necessarily postalveolar-favouring (except for one language, cfr. Hansson 2001*a*: 192ff. and Kosa 2010). Regressive assimilation would for example yield /ʃasa/ > [sasa] and /saʃa/ > [ʃaʃa], while postalveolar-favouring assimilation would yield /ʃasa/ > [ʃaʃa] and /saʃa/ > [ʃaʃa]. This can suggest that the phenomenon at hand is the so-called *Emergence of the Unmarked* (*TETU*, cfr. McCarthy and Prince 1994; McCarthy 2001; Albright 2004; Mascaró 2004; Becker and Potts 2011), which would deem posterior sibilants to be *less* marked, but is in fact the opposite. It is indeed acknowledged in the literature that assimilation patterns tend to favour *more* marked segments, i.e. that less marked segments assimilate to more marked ones (Rice 1996: 494f.; Mohanan 1993, cited in Al-Deaibes 2016: 97;

Brown 2016: 417f.). This phenomenon is called the *Preservation of the Marked* (Lamont 2015: 34): since posterior sibilants are more difficult to pronounce, anticipatory and/or perseverative effort goes prioritarily into reaching the tongue configuration required for these more marked segments, while 'easier' segments are given the same configuration to preserve the difficult configuration attained. This results in the application of the articulatory gestures for the more marked sibilant to the less marked sibilant (Hansson 2010: 347f.).

It has been argued in this section that the higher cross-linguistic frequency of [t͡ʃ] compared to [t͡s] is no counterevidence for the posterior markedness universal, but is a consequence of the cross-linguistic frequency of stop palatalisation instead. An argument in favour of this view is provided by assimilation patterns: in languages in which sibilant affricates participate in long-distance or direct assimilation, they behave precisely like their sibilant fricative counterparts (Hansson 2010: 358). Therefore, [t͡ʃ] is more marked than [t͡s], since it patterns as a more marked segment in sibilant assimilation.

4.4.1.5 Reduplication and TETU

Contrarily to the *Emergence of the Marked* in assimilation patterns, I have found evidence of the *Emergence of the Unmarked* in reduplication patterns in the Canadian language Gitksan. The unmarked emerges in reduplication, since e.g. sibilant affricates surface as sibilant fricatives in the reduplicated prefix (Brown 2008: 148):

(68) m'ats mis-m'ats 'to hit, strike'

The unmarked also emerges with respect to PoA, namely as anterior: "When a base containing [JK: palatoalveolar] fricatives is reduplicated, the fricative surfaces as alveolar" (Brown 2008: 148).

(69) maʃxʷ mis-maʃxʷ 'white?'

Posterior markedness is thus supported both by evidence from patterns in which the unmarked emerges, and patterns in which the marked emerges.

4.4.1.6 Diachronic fronting

Contact between natural languages has yielded several cases in which speakers of a 1s, i.e. with retracted alveolar sibilants, are exposed to languages with well-dispersed 2s inventories and become such languages themselves. In such cases, /s̠/ could theoretically become /s/ or /ʃ/, but it only becomes /s/ because inventories with a single /ʃ/ are unattested and predicted to be impossible (Vijūnas 2010: 49). An example is provided by Afrikaans: it had originally a retracted sibilant, which the dialects in southern Holland (from where the language originates) still have today. Rather recently, /s/-fronting has been reported to be a phenomenon that is widespread among the younger (white) generations, so that retracted alveolar sibilants are only used by older (white) speakers (Wissing et al. 2015: 154). The Younger Afrikaans sibilant inventory has thus become one in which the vast majority of sibilants, in term of frequency inside the language, are a (denti-)alveolar /s/, besides a small portion of /ʃ/, mainly found in loanwords.

This generalisation favouring s-fronting does not only hold for 1s, but also for any sibinv with /s̠/. Rare are the languages (e.g. a Walser German dialect, cfr. Dal Negro 2004) in which a sibinv becoming a well-dispersed 2s has moved a historical /s̠/ predominantly to /ʃ/: in fact, when it happens, it only takes place in a few specific retraction-prone contexts. For example, Middle French shifted /s̠/ to /s/ in all contexts; in Middle High German, it happened in all contexts except certain preconsonantal contexts and after /r/, where it became /ʃ/; in Flemish Dutch, /s̠/ was shifted to /s/ in most contexts, except optionally in a few contexts such as /sj/. The same holds for Romanian and standard Italian, in which Latin */s̠/ is predominantly reflected by /s/. The same holds for e.g. Norwegian, in which Old Norse */s̠/ is reflected as /s/ in all contexts except in s-retraction contexts such as historical /skj/ and /rs/. The data is thus abundant to show that /s̠/ is much more likely to be fronted diachronically and backed in only a few contexts, rather than the opposite, thereby conforming to the posterior markedness universal.

4.4.1.7 Comparison with random inventories

The posterior markedness universal is supported by evidence from the comparison of the 236 inventories in the SibInv database with 236 randomly generated inventories. Assuming 4 PoAs and giving one 'postalveolar point' to sibinvs for every third or fourth sibilant they possess (i.e. for every postalveolar sibilant), a probability distribution is to be expected in which the chance of having 1 'postalveolar point' is 8/16 and that of having 0 or 2 'postalveolar points' 4/16 in both cases. This means that the most likely mean number of 'postalveolar points' is 1. Comparing randomly generated sibinvs with this probability distribution (4 PoAs, 0.5 chance of having a sibilant for each PoA), one obtains a highly significant difference ($t = 4.241$, $p < 0.001$) between the randomly generated and the SibInv inventories. The mean number of 'postalveolar points' is 0.894 for the random inventories and 0.653 for SibInv inventories, demonstrating that natural languages tend to have fewer postalveolar sibilants than what would be predicted by chance.

4.4.2 Motivations for the posterior markedness universal

4.4.2.1 Teleological or emergent?

The evidence from language acquisition suggests that the posterior markedness universal is a teleological phenomenon grounded in articulatory difficulty: children perceive the contrast between /s/ and /ʃ/, but do not manage to pronounce the latter and produce the former instead. This speaks in favour of a constraint in the synchronic grammar against posterior sibilants.

4.4.2.2 Articulatory motivations

As said above, evidence from language acquisition speaks in favour of an articulatory motivation. Although posterior markedness is usually mentioned as such in the literature without any articulatory or perceptual motivation, as explained in 4.4.1, one can posit that more tongue muscles need to be activated and/or larger displacements of the tongue are required to produce posterior sibilants. I hypothesise that this effort is related to a vertical dimension: when producing posterior sibilants, the tongue must be raised more because the palate is further away from it (Green and Wang 2003: 2831), as illustrated on the sketch in fig. LIX. The form of the palate, going from a higher posterior zone to a lower prepalatal/alveolar (and thus anterior) zone, plays in disfavour of articulatory economy for posterior sibilants (and coronal segments in general, I hypothesise). Anterior and posterior sibilants all require some muscle activation for raising the tongue

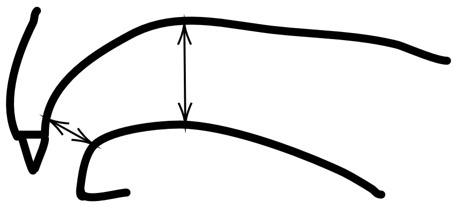


Figure LIX: *Sketch of the distance between tongue and palate for anterior and posterior sibilants.*

towards the alveola and the palate, respectively, but the latter imply a stronger activation because of the larger distance. Keating (1991: 41) confirms that palatoalveolars are produced:

> [L]ike alveolars, but with the whole tongue moved *back and up* just a little; [JK: plain] retroflexes can also look like alveolars, but with the blade curled back just a little more. (Sublaminal retroflexes can also be made well behind the [JK: alveolar] corner, of course.) [my emphasis]

For example, Kluender and Walsh (1992: 329) mention that "[i]n the articulation of /t͡ʃ/, the tongue moves both downward and backward from the alveolar ridge toward the more alveopalatal position for /ʃ/ (Jones 1956)", i.e. with a backward movement. Palatoalveolars thus require additional raising of the blade and dorsum, and retroflexes a raising targeting predominantly the tongue tip (Hamann 2003: 4). This difference between anterior and posterior sibilants is encoded in the synchronic grammar as an additional constraint against raising the tongue, because it is determined by a different muscle, the styloglossus (Green and Wang 2003: 2831), which shifts the tongue on a different axis (namely the vertical one). Producing a [s̺], for example, requires the activation of the hyoglossus to pull the tongue downwards and backwards, of the styloglossus to raise and retract it, and of the upper longitudinal muscle to raise the tongue tip, making it an apical sound (Boersma and Hamann 2008: 231; Drake et al. 2009: 1039ff.).

4.4.2.3 Perceptual motivations

An alternative explanation could be that the perceptual properties of anterior sibilants favour them with respect to posterior sibilants. Flemming (2008*a*: 158) and Clements (2009: 29f.) emphasise the fact that anterior sibilants have a higher CoG and are thus more distinct from other sounds than non-sibilant fricatives and affricates. This explains, according to Flemming (2008*a*), the cross-linguistic preference for sibilant fricatives and affricates rather than non-sibilants. This would imply that the sibilants with the highest CoG, thus the most anterior sibilants, are preferred. An objection to this is that such an emergent, perception-based principle would not be encoded in the production grammar and thus fail to explain why children perceive [ʃ] (Cristià et al. 2011) but cannot produce it at first, despite being encouraged to do so. Furthermore, it would allow the existence of 1s languages with excessively anterior sibilants (e.g. with only [s̪]), inventories which would reflect an absolute proportional preference for the front sibilants. Posterior markedness as a perceptually motivated is thus not supported by the evidence, contrarily to the articulatory motivations provided here.

4.5 Voicing markedness

It is widely acknowledged in the literature that voiced obstruents, and thus voiced sibilant fricatives and affricates, are more marked than their voiceless counterparts (cfr. e.g. Rice 2007: 81, 85). This section summarises the evidence and discusses its implications for phonological theory.

4.5.1 Cross-linguistic evidence

In the languages of the world, the presence of a voiced sibilant generally implies that of a voiceless sibilant at the same PoA with the same manner, as shown in 3.2.1.3. Among others, Maddieson (1984: 47) states that "generally, the existence of a given voiced fricative in the inventory implies the presence of a voiceless counterpart in the inventory". The same implication holds for sibilant affricates (Żygis, Fuchs and Koenig 2012: 301), and is observed in obstruent consonant inventories in general (Rice 2007: 94).

4.5.1.1 Cross-linguistic frequency

In 3.2.1.3, it has been evidenced that voiced sibilants are cross-linguistically less frequent than voiceless sibilants (Johnson 2003: 124). For instance, if there is only one fricative in a language, it is most likely "a voiceless sibilant of the type usually represented by the letter *s*" (Maddieson 2013*a*). This frequency asymmetry also holds for (sibilant) affricates, which are more frequent as voiceless rather than voiced (Berns 2013: 68f.). In the 258 inventories of the SibInv database, there are 446 voiceless sibilant fricatives and 241 voiced sibilant fricatives, 303 voiceless sibilant affricates and 168 voiced sibilant affricates. The ratio voiceless vs. voiced is 1.85 for the fricatives and 1.80 for the affricates.

This implicational universal (/Z/ -> /S/) is thus reflected in the statistical distribution of sibilants in the world's languages.

4.5.1.2 Intra-linguistic frequency

In his sample of 50 languages, out of which 23 possess a voicing contrast between [s] and [z], Peust (2008: 124) found [s] to be more frequent than [z] in all but three languages (i.e. 86.96%). All other languages have no such voicing contrast and only [s], while no language with [z] lacks [s].[4] The same voicing asymmetry holds for postalveolar sibilants, with no language having [ʒ] without [ʃ] and the latter being more frequent than the former in 13 out of 16 languages (81.25%) in Peust (2008).[5]

With an independent sample, Everett (2018: 8) confirms this statistical trend in the frequency of use of fricatives: "[t]he combined rate of occurrence of [f], [s], and [ʃ] in the data, based on the cross-family averages of each sound type, is 0.039. In contrast, the combined rate for [v], [z], and [ʒ] is 0.0066. In other words, there are about six times more voiceless fricatives in the database". Inside a language, voiceless sibilants are thus much more frequent than their respective voiced counterparts, with a few exceptions.

4.5.1.3 First language acquisition

Also in language acquisition, voiced sibilants are disfavoured. Benet et al. (2012: 400f.) report that Catalan-speaking children produce fewer target-like tokens of /z/ than adults, as observed independently by Bosch Galcerán (1987).

> [A]ccording to Grunwell (1982) for English, children produce [s] at age 3;0 – 3;6, before they produce [z] at ages 3;6 – 4;0 and 4;0 – 4;6. According to Bosch Galcerán (1987) for Catalan, /s/ is acquired at 4 years and /z/ at 5 years. (…) [I]t is plausible that this voiced sibilant has not been (completely) acquired by some of the younger children [JK: in their study on Catalan] yet.

Ball et al. (2014) report the implicational scale of acquisition /s > z > ʃ > t͡ʃ, d͡ʒ/ (from earliest to latest acquired) in English. The same is reported for Jordanian Arabic by Yaseen and Mahadin (2018: 11), who mention that "different studies on language acquisition (…) concluded that voiced sibilants are of the latest consonant [sic] to be acquired in the world's languages (Ohala 1983, Goldman et al. 2000, Amayreh and Dyson 1998)". As an orthophonist, Lemieux (2011) reports that French [z] is mastered slightly later than [s], around 2.5 years of age. The same is documented for Japanese /z/ wrt. /s/ (Kubozono 2015: 683) and Arabic (Alamer 2018: 90). More generally, Beers (2003) observes the frequent devoicing of obstruents in Dutch-speaking children, while van 't Veer (2015: 121,

[4] He quotes Nartey (1979: 8) mentioning Chukchee as having [z] without [s], but this is actually not the case in Chukchee (Dunn 1999).

[5] Even in the three languages in which [ʒ] is more frequent than [ʃ], e.g. standard French (Malécot 1974), historical reasons can be advanced to explain this: French [ʃ] in native vocabulary comes from a historical affricated /ka/, while [ʒ] comes from a corresponding affricated /ga/ *but* also from a historical affricated /gj, gi, ge/ (Berns 2013) and an assibilated /j/ in onset position, whereas there was e.g. no corresponding /ç/ to shift to [ʃ].

124) illustrates the earlier acquisition of voiceless obstruents by Dutch-speaking children. Five out of seven children in his study acquired Dutch /s/ earlier than /z/, one acquired them simultaneously and one acquired /z/ earlier than /s/ (van 't Veer 2015: 190-203).

In the more specific context of word- or utterance-final positions, Smit (1993), cited in Richtsmeier (2010: 3), notes that English-speaking children devoice word-final obstruents at an early age (around 2 years old). Among the English obstruents, /z/ is the most frequently devoiced with a 15-30% devoicing frequency, against 5-15% for /v, d͡ʒ/ and voiced stops. Song et al. (2013: 2932) similarly report that very young English-speaking children devoice final obstruents. The markedness of voiced obstruents is thus reflected in child phonology, irrespectively of the native language.

4.5.1.4 Comparison with random inventories

Comparing the 236 inventories in the SibInv database with the same number of randomly generated inventories, a statistically highly significant difference ($t = 18.337$, $p < 0.001$) arises with respect to the mean number of voiced sibilants per sibinv. Given 4 PoAs, a fricative - affricate distinction and 4 possibilities for each sibilant (no sibilant, a voiceless sibilant, a voiced sibilant or both voiced and voiceless sibilants), the probability of being a voiced sibilant equals 0.5 in randomly generated inventories; with 4 PoAs and affricates, this means that the mean number of voiced sibilants in a randomly generated inventory should be located around 4 (because there are 8 sibilants). The mean number of voiced sibilants indeed turns out to be 4.089 in random inventories, compared to 1.597 for SibInv inventories. Even when replicating this calculation without the possibility of having a voiced sibilant without a voiceless sibilant at the same place and manner, despite the logical probability of having voiced sibilants being reduced to 1/3 (i.e. 2.666 out of 8 sibilants considering 4 PoAs and affricates), this difference is still highly significant ($t = 8.046$, $p < 0.001$; mean number of voiced sibilants: 2.682 vs. 1.597).

4.5.2 Motivations for the voicing markedness universal

4.5.2.1 Teleological or emergent?

The markedness of voiced obstruents is encoded in grammar as a teleological principle (cfr. e.g. the OT constraints *Lar in Lombardi 1991, *[αvoice] in Bradley and Delforge 2006, *[cont][voice] in van 't Veer 2015: 30 and specifically for sibilants, *Voiced-Sib in Yaseen and Mahadin 2018: 16). It is articulatorily based, since voicing conflicts aerodynamically with frication, as explained in the following subsubsection. However, voicing markedness in sibilants is grounded in perception as well, as detailed in a further subsubsection.

4.5.2.2 Articulatory motivations

Ohala (1983) and Żygis, Fuchs and Koenig (2012: 310) describe the articulatory constraints on the combination of voicing and obstruency as follows:

> It has long been known that vocal-fold vibration requires a pressure differential across the glottis (…) [JK: that] ensures sufficient airflow for a transfer of energy to the vocal fold tissues.

To realise this in a 'homemade' experimental way, the interested reader can try to pronounce a continuous [m]-sound as long as possible with the mouth fully closed but an unobstructed nose. (S)he can last as long with this continuous [m] as (s)he has air in the lungs (e.g. 46 seconds, my own result). Doing the same while totally obstructing the nose also by pinching one's nostrils, (s)he should be able to produce one continuous [m] for a much shorter time period (e.g. 2.5 seconds, my own result), and should feel increasingly the incapacity of vibrating the vocal folds proportionally with the air getting stuck in the mouth and the nose. This is because when the nose was open, air could flow out and thus the pressure could be higher below the glottis than above it, but when the air cannot escape, this difference in pressure can only last a very brief moment, after which the air pressure becomes equal everywhere because the air cannot escape.

Żygis, Fuchs and Koenig (2012: 310) pursue their explanation as follows:

> In obstruents, the closure or constriction in the upper vocal tract leads to an increase in supraglottal or intraoral pressure (P_{io}). Assuming that the subglottal pressure (P_{sub}) remains approximately constant, increasing P_{io} will lead to a decreasing transglottal pressure difference (P_{trans}). When P_{trans} falls below the threshold value, voicing will cease.

This thus means that the stronger the obstruction is in the mouth cavity, the harder it will be to create a stronger pressure below the glottis. In our radical example with the continuous [m] and closed nostrils, the obstruction was total and therefore, voicing became rapidly impossible; when the obstruction is not as radical but still considerable (e.g. when producing a [s]), the air escapes less easily than when pronouncing e.g. [a], so that the pressure after the glottis and before the obstruction (at the back of the mouth) will be closer to the pressure below the glottis when producing [s] rather than [a]. This decreased pressure difference makes it hard to maintain voicing for an obstruent, since the risk of having an equal pressure below and above the glottis will be higher.

> [A]erodynamic modeling (Müller and Brown 1980; Westbury 1983) has suggested that voicing can be sustained throughout the duration of a typical stop closure only if speakers perform compensatory actions to slow the rate of P_{io} increase (i.e., maintain a higher P_{trans}). Such actions may include altering the compliance of the vocal tract walls (Westbury 1983), expanding the supraglottal volume via movements of the larynx and upper articulators (Kent and Moll 1969; Perkell 1969; Bell-Berti 1975; Riordan 1980; Svirsky

> et al. 1997; Fuchs 2005), and allowing airflow leakage through the velopharyngeal port (Bell-Berti 1975). These studies also suggest that speakers may use combinations of such volume-compensating maneuvers. These considerations provide an explanation for the typological infrequency of voiced obstruents: These sounds require compensatory maneuvers, or greater aerodynamic control, compared to voiceless obstruents, since an oral closure will make continued voicing difficult. Ohala (2011) refers to this effect as the Aerodynamic Voicing Constraint. (Żygis, Fuchs and Koenig 2012: 310)

They further explain how articulatory constraints on voicing and frication conflict: an excessively high pressure in the oral cavity reduces the possibility of voicing, while an excessively low pressure in the oral cavity prevents frication, so that only a precise 'dosage' of pressure can successfully produce a voiced fricative (Żygis, Fuchs and Koenig 2012: 311). The complex articulatory combination of voicing and frication is also documented by e.g. Proctor et al. (2010). Żygis, Fuchs and Koenig (2012: 312f.) then detail how this conflict with voicing applies even more for affricates.

4.5.2.3 Perceptual motivations

Moreover, the markedness of voiced sibilants is also explained by perceptual factors (Chuang and Fon 2017: 2). As already reported above (e.g. 2.1.2 and 4.3.2.3), sibilants are perceptually distinguished from non-sibilants by their energy at high frequencies (i.e. their high CoG). Ohala and Solé (2010: 42) explain that voiced obstruents are disfavoured in perception because this acoustic cue is less salient as a consequence of voicing. Other perceptual motivations are also given:

> [A] lower oral pressure for voiced obstruents will result in a lower intensity of high frequency noise during the fricative constriction or at stop release vis-à-vis voiceless obstruents. In addition, due to the reduced transglottal flow, voiced obstruents take longer to build up oral pressure behind the oral constriction, which results in a delayed onset of audible frication for fricatives (Solé 2002) and a weaker burst for stops compared to their voiceless counterparts. Thus, the characteristic cues for obstruency – abrupt amplitude discontinuities and high intensity noise cues – are enhanced in voiceless obstruents due to the larger rate of flow through the glottis. In sum, for aerodynamic and auditory-acoustic reasons voicelessness favors or enhances obstruency (i.e., high intensity frication and release burst).

They thus shed light on the fact that the acoustic cues that distinguish obstruents are better perceived in voiceless rather than voiced obstruents. This has been corroborated by perceptual studies such as that of Balise and Diehl (1994), who report lower correct identification rates for voiced sibilants compared to voiceless sibilants.

One can thus conclude, together with Holt (2007: 389), that "voiced sibilants are marked both articulatorily and perceptually".

4.6 Affricate markedness

4.6.1 Cross-linguistic evidence

Affricates are more marked in the world's languages than fricatives. This has been documented in detail in 3.2.1.4, where the posited universal implicational hierarchy between sibilant fricatives and affricates has been relativised. It nevertheless still holds as a strong statistical tendency that "an alveolar affricate such as /ts/ [JK: *generally*, and only generally] implies the presence of /s/ in a language but not vice versa" (Greenberg 1963: 3), something which applies to (sibilant) affricates in general.

4.6.1.1 Cross-linguistic frequency

Nikolaev and Grossman (2018: 563) observe that only 6.4% of the languages in the UPSID sample (Maddieson 1984) have 5 or more affricates, while 39.9% of the languages in the sample have 5 or more fricatives. As reported in detail in 3.2.1.4, sibilant affricates are cross-linguistically rarer than their fricative counterparts, and only a few languages have more sibilant affricates than fricatives (4.48% in the SibInv database).

4.6.1.2 Intra-linguistic frequency

Unexpectedly, sibilant affricates are generally more frequently used inside one and the same language than their fricative counterpart, as calculated in Peust (2008: 108-113)'s corpora. I counted the sibilant affricate - fricative ratio for each pair corresponding in PoA and voicing in his 50 languages, and arrived at the surprising result of a 1.37 time higher frequency of use for sibilant affricates compared to their fricative counterpart. However, there is a large disparity between anterior and posterior sibilants: the affricate - fricative ratio is 0.39 for anterior sibilants but 2.09 for posterior sibilants, with a statistically significant difference (t = -2.724, p = 0.008). This is likely to reflect the cross-linguistic frequency of stop palatalisation, as explained in 4.4.1.1, rather than a preference for sibilant affricates over fricatives. The markedness of sibilant affricates is thus reflected in the statistical frequency of anterior sibilants, but not in posterior sibilants due to this emergent bias in favour of [t͡ʃ].

4.6.1.3 First language acquisition

Sibilant affricates are cross-linguistically acquired later than fricatives. Fletcher (1989: 746) reports the following in his phonetic study on the acquisition of (among others) the English sibilant affricates /t͡ʃ, d͡ʒ/:

> A further comment might be made with respect to the apparent complexity of the maneuvers in stop, sibilant, and affricate articulatory gestures. (...) Generating the combined stop/sibilant maneuvers of the affricates would presume prior acquisition of both stop and sibilant gestures. This, combined

> with the observation that the number of sensors contacted in the initial affricate constriction movement was consistently greater than in /t, d/ sound production, suggests that considerable effort as well as skill may be required for affricate production. These physical demand and phonetic contrast interpretations are consistent with the commonly observed order of stops, sibilants, then affricates in normal speech development.

The later acquisition of sibilant affricates wrt. fricatives has been observed e.g. in English (Dyson 1988: 92; Dinnsen and Farris-Trimble 2009; Brosseau-Lapré and Rvachew 2018: 66), Italian (Zanobini et al. 2012), Arabic (Alamer 2018: 93), Mandarin and Cantonese Chinese as well as Russian (Hua and Dodd 2006: 66f.). Hua and Dodd (2006) mention Japanese as an exception, since /t͡ɕ/ and /d͡ʑ/ are apparently acquired earlier than predicted. However, Ota (2015: 684) shows that this is only true before /i/ (which is not surprising, since it is a very palatalisation-friendly context), but not before other vowels.

> [T]his putatively universal order of development is disrupted in Japanese by the sibilant affricates [t͡ʃ] and [d͡ʒ]. Contrary to the crosslinguistically prevalent pattern in which sibilant affricates are acquired much later than corresponding stops, [t͡ʃ] and [d͡ʒ] in Japanese typically reach production accuracy criteria around the same time as their stop counterparts [t] and [d]. However, a closer analysis by Edwards and Beckman (2008) reveals that this non-compliance of the universal order is conditioned by the subsequent vowel. In most contexts, [t͡ʃ] actually lags behind [t] in development. But [t͡ʃ] has a higher production accuracy than [t] before /i/, making the average timing of acquisition similar for the two sounds. The pre-/i/ context is where the contrast between /t/ and /t͡ʃ/ neutralizes to [t͡ʃ] in Japanese except in some loanwords (e.g., [pa:ti:] 'party'), and consequently presents substantially more instances of [t͡ʃ] than [t] in the input. The relatively early acquisition of sibilant affricates in Japanese, therefore, can be seen as a case where a language-specific frequency effect is "overlaid on a universal articulatory ease effect" (Edwards and Beckman 2008: 146).

Similar results have been reported for children with speech disorders, although a few exceptions can be found (e.g. English /t͡ʃ, d͡ʒ/ being acquired earlier than sibilant fricatives for 1 out of 40 children in Dinnsen et al. 1990: 32 and for the only child in Miccio and Ingrisano 2000). The late acquisition of sibilant affricates thus holds generally, as evidenced by the literature. The fact that often, /t͡ʃ, d͡ʒ/ are acquired rather early (as compared to /t͡s, d͡z/, e.g. in Italian, cfr. Zanobini et al. 2012: 22 and Hua and Dodd 2006: 67) is linked to the fact that children are likely to realise affricates with a large tongue contact (Fletcher 1989) and centrally, thus producing [t͡ɕ, d͡ʑ] in more precise phonetic terms. This acoustic output is likely to be perceived by adults as /t͡ʃ, d͡ʒ/, while its early acquisition is in fact a consequence of the centrality and tongue contact pattern of [t͡ɕ, d͡ʑ], not of [t͡ʃ, d͡ʒ].

4.6.1.4 Comparison with random inventories

Comparing the number of affricates in 236 randomly generated and SibInv inventories, one observes a statistically highly significant difference ($t = 10.585$, $p < 0.001$) between randomly generated inventories (mean number of affricates: 2.072, i.e. close to the logical probability of 2/4 PoAs) and SibInv inventories (mean number of affricates: 1.127). The same calculation expressed as an affricate-per-fricative ratio, after ruling out what would be a division by 0 (e.g. in the case of sibilantless languages), yields a highly significant difference ($t = 7.776$, $p < 0.001$) with a mean ratio of 1.178 for random inventories (close to the logical mean probability of 1) and 0.675 for SibInv inventories.

4.6.2 Motivations for the affricate markedness universal

4.6.2.1 Teleological or emergent?

The motivation for affricate markedness resides mainly in the articulatory complexity of sibilant affricates, as detailed in the next subsubsection. It is therefore a teleological principle, because speakers have a (not necessarily conscious) knowledge of the difficulty of their production, in the form of sensorimotor constraints (Boersma 2009). Nevertheless, an additional emergent perceptual motivation has also been found, as explained below.

4.6.2.2 Articulatory motivations

The articulatory gestures required to produce sibilant affricates are more complex than those for stops and fricatives. As Żygis, Fuchs and Koenig (2012: 312) explain:

> Since affricates involve precise temporal control over two constriction regions, as well as changes in the coronal configuration of the tongue, one can hypothesize that they are phonetically more complex than singleton stops or voiced fricatives. (…)
> [A]ffricates seem to require not only that speakers produce a sequence of a stop and a fricative, but that they produce rather brief stop and fricative components. One can hypothesize that producing such temporally short sequences represents an additional form of motoric complexity. In sum, these factors may provide an explanation for why affricates are less common, cross-linguistically, than simple fricatives or stops.

A similar account of the complexity of sibilant affricates is given by Fletcher (1989: 746) as described in 4.6.1.3 above.

4.6.2.3 Perceptual motivations

Sibilant fricatives and affricates are distinguished from each other on the basis of duration: a short frication duration is perceived as an affricate, a long one as a fricative (Kluender and Walsh 1992). The same is reported in Żygis, Fuchs and Koenig (2012:

312)'s quote in the previous subsubsection. On the continuum of duration, the shortest segments are thus stops, followed by affricates and then fricatives (Shinn 1986: 27).

Since affricates are 'clamped' in the middle between stops and fricatives, they are likely to be misperceived as a stop if they are too short and as a fricative if they are too long, while stops and fricatives only risk to be misperceived if they are too long or short, respectively. This higher likelihood of being misperceived is reflected in the higher misperception rate found by perceptual studies (e.g. in German, Mandarin Chinese and Czech, cfr. Shinn 1986: 260). In Shinn (1986)'s study, stops were correctly identified around 90% of the time in Mandarin and Czech and 80% in German. The misperceptions consisted exclusively of affricates in the two first languages, while it consisted for 17% of affricates and less than 3% of fricatives in German. About the same was found for the perception of fricatives: correct identification ranged between 72% and 87%, misperceptions consisting almost exclusively of affricates. However, the perception of affricates was less successful: 78% (against 91% for stops and 87% for fricatives) in Mandarin, 69% (against 90% for stops and 80% for fricatives) in Czech, and 65% (against 80% for stops and 72% for fricatives) in German. Across these languages, the correct perception of affricates lags behind by around 10%, and misperceptions go, as predicted, in both directions (stop and fricative).[6]

The higher misperception rate of affricates as fricatives or stops can thus be understood as an emergent tendency towards the perceptual reanalysis of affricates as another segment and thus a diminution in their cross-linguistic frequency. This, however, is not encoded in grammar but the result of the interaction between speakers and listeners; it can be understood as a kind of auditory dispersion effect for durational cues (cfr. Boersma and Hamann 2008: 224 and section 4.8).

4.7 Fricative - affricate and voicing symmetry

It can be observed in natural languages that very frequently, sibilant affricates have a corresponding fricative at their same PoA (Berns 2013: 27), as well as vice-versa though to a smaller extent (cfr. 5.4.2). Even more strongly than this high frequency of symmetrical phoneme pairs distinguished only by manner, symmetrical voiceless and voiced pairs only distinguished by voicing constitute a near-universal already evidenced in 3.2.1.3 and 3.2.3.3. As exemplified by the most frequent sibilant inventory in the SibInv database in (62), fricatives and affricates as well as voiceless and voiced sibilants are so often well-aligned in terms of PoA that it appears likely that a principle (be it teleological or emergent) favours the symmetry between them. This lead e.g. Martinet (1968: 483) to note that "for the same total of phonemes, [JK: symmetrical inventories] require fewer articulations to keep distinct". Similarly, Bybee (2001: 54) observes that there is a "strong tendency for speakers to reuse a single set of highly entrenched neuromotor patterns and to substitute members of this set for novel or less common configurations". The

[6] Interestingly, the perceptual confusion between fricatives and affricates is larger than any other (Shinn 1986: 260), something which coincides with the cross-linguistic frequency of deaffrication to fricatives (see B.9) and the rarity of deaffrication to stops (see C.5).

topic of symmetry has been addressed mainly in phonetic studies on sibilant fricatives and affricates inside one and the same language, on the phonetic side, and in typological studies on feature economy (mainly) in the DFT literature, on the phonological side.

4.7.1 Cross-linguistic evidence

4.7.1.1 In phonetics

Several fine-grained phonetic studies on the production of sibilant fricatives and affricates point out that they typically coincide in terms of PoA and general articulatory configuration, even if not with maximal precision. The fact that the average area of contact during the fricative portion of affricates tends to be slightly more towards the alveolar zone is a consequence of the preceding stop portion (Dixit and Hoffman 2004: 155; Żygis 2008: 33), an effect which is not sufficiently strong to have consequences for phonological patterning. Such "small but systematic differences" (Żygis 2008: 33) are also found between voiceless and voiced sibilants, but these fine-grained articulatory nuances are most frequently not relevant for phonology and the structure of the inventory (although see the next paragraph for an exception). When it comes to voicing, symmetry is an extremely widespread phenomenon, to such an extent that it is normally taken for granted in phonology under the term "voiceless/voiced pairs" (e.g. Trask 2004: 165; Dannenbring 1980: 980), and will not need to be further demonstrated here.

An exception to the parallel behaviour of voiceless and voiced sibilants in phonology is the fact that a voiceless stop can be palatalised at the same time as its voiced counterpart doesn't. In this case, the fine-grained phonetic differences in articulatory configuration play a role for phonology. Voiceless stops have a longer friction phase before /i, j/ than their voiced counterpart, and since "a longer friction duration after a coronal stop is more likely to be interpreted as affricate than a shorter one" (Hall et al. 2006: 62, 75), voiceless stops can be palatalised without their voiced counterpart. This explains for example why the reflex of Latin CE, CI in historical standard French was palatalised early and then fronted to [t͡s] when the third palatalisation (that of CA, GA) occurred (e.g. [t͡s]*ent* 'hundred' from CENTUM), but its voiced counterpart was palatalised later, at the same time as CA, GA, and thus merged with the result of GA to become [d͡ʒ] (e.g. [d͡ʒ]*ens* from GENTES; cfr. Tifrit and Voeltzel 2016: 2).

Evidence for similar articulatory place and configurations for sibilant fricatives and affricates has been found among others in English, in which the posterior affricates "fit rather neatly into the stop system as counterparts of the /ʃ, ʒ/ in the fricative system" (Fletcher 1989: 746). The same is reported for Catalan (Recasens and Espinosa 2007: 146f.), with an explicit use of the term 'symmetry' to denote the similar articulations. Dixit and Hoffman (2004: 155) confirm this trend in the articulation of a speaker of Hindi. Similarly, Lee and Zee (2010: 336) report that "[b]oth Cantonese fricative [s] and affricate [ts] [sic] are articulated with the blade of the tongue making contact on the alveolar ridge. The affricate [ts] has a close affinity in articulation with the fricative [s] rather than the stop [t]". Ladefoged and Wu (1984), cited in Żygis (2008: 34), show that "[a]part from (...) small disparities [JK: cfr. the previous paragraph], the affricates and the corresponding fricatives share the same place of articulation" in Pekingese. Finally, Avesani et al. (2017: 285) note that in Italian dialects of *Veneto*, /s/ and /z/ are "perceptually similar to the postalveolar fricatives [ʃ] and [ʒ]", while the postalveolar affricates [t͡ʃ] and [d͡ʒ] lack the typically posterior lip protrusion, making the PoA of all these segments very close to each other (and thus suggesting symmetry, without however stating explicitly if they share precisely the same PoA). Impressionistically, I found no perceptual difference between the fricatives and the fricative parts of the affricates in the Veronese regiolect (e.g. [t͡s̺ao] for standard Italian [t͡ʃ]*ao* 'ciao'). This view is supported by spectral measurements in Trentino dialects in Alber and Kokkelmans (in prep.).

4.7.1.2 In phonology

It has been observed by several researchers that most sibilant affricates "have *basically the same* (or fewer) possibilities of place of articulation as do the sibilant fricatives" (Smith 2000: 249; my emphasis). Smith (2000) exemplifies this with English, which has two PoAs for sibilant fricatives and a /t͡ʃ, d͡ʒ/-pair coinciding with the postalveolar sibilants, and Bžedukh, whose three PoAs for sibilant fricatives precisely match those for sibilant affricates. van de Weijer and Hinskens (2004: 218), cited in Berns (2013: 101), also mentions that affricates "typically occur at the same places of articulation as (coronal) fricatives". "[A]ffricates indeed rarely have a corresponding plosive at the same place of articulation, while in the majority of cases there is a corresponding fricative" (Berns 2013: 102); this is a consequence of the fact that in a stop + fricative combination, "the stop is influenced by the fricative rather than the other way around" (Recasens and Espinosa 2007: 146).

About phonological segments in general, Żygis, Fuchs and Koenig (2012: 300) observe that:

> [P]honological systems tend towards symmetry, such that contrastive features will tend to be used in parallel across the inventory. The general expectation of symmetry in phonological systems can be observed, for example, in classical distinctive feature theory (e.g., Jakobson et al. 1952; Jakobson and Halle 1956; Chomsky and Halle 1968; Clements 1985; Hall 2001), which placed a premium on maximizing the use of a small set of distinctive features (e.g., Clements 2003).

In Distinctive Feature Theory, symmetry in fricative - affricate and voicing pairs is explained as feature economy (Clements 2003). Feature economy can be defined as the tendency of natural languages to have a number of phonemes matching as closely as possible the maximal number of combinations that can be made with the features distinguishing these phonemes. To efficiently measure to what extent an inventory is economical, Hall (2007), cited in Mackie and Mielke (2011: 46), provides the following formula (called *exploitation score*):

$$E = \frac{\text{Number of segments present in an inventory}}{2^{\text{Number of features required to characterise it}}}$$

For example, the perfectly symmetrical and economical inventory in (62) with /s, z, t͡s, d͡z, ʃ, ʒ, t͡ʃ, d͡ʒ/ has an exploitation score of 1, because its 8 segments are generated with 3 features ([±anterior], [±continuant] and [±voice]; 2 x 2 x 2 = 8). Considering that English lacks two of these segments (/t͡s, d͡z/), its exploitation score for sibilants is 6/8 = 0.75; if one considers that standard French lacks phonemic affricates, its exploitation score for sibilants would be 4/4 = 1. Feature economy is thus best understood concretely as the tendency of languages to have an exploitation score as close as possible to 1.

The inventory in (62) happens to be perfectly *symmetrical* and *economical*, but this does not mean that symmetry and economy are the same. The difference between both becomes apparent if we consider the following inventory:

(70)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.
Sib. aff.		t͡s d͡z			t͡ʃ		
Sib. fri.		s z			ʃ		

Clements (2003: 291f.) explains that such an inventory might be technically symmetrical, but it is not perfectly economical, because it lacks /ʒ, d͡ʒ/. There are gaps, which makes it not perfectly economical, but these gaps preserve symmetry by occurring symmetrically (as opposed to e.g. gaps at /z, d͡ʒ/).

In the case of sibilant fricatives and affricates, we are interested in symmetry especially, i.e. in the fact that sibinvs might not exploit all possible contrasts but still tend to use the same articulatory gestures for fricatives and affricates (i.e. 'gesture reuse'). As Clements (2003: 295) puts it, what I call fricative - affricate symmetry is "a principle analogous to feature economy operating at the phonetic level". Maddieson (1995: 574), quoted in Clements (2003: 295), notes the following:

> There is [a tendency] to be economical in the number and nature of the distinct articulatory gestures used to construct an inventory of contrastive sounds, and it is this (rather than a more abstract featural analysis) that underlies the observed system symmetry.

This echoes Bybee (2001: 54)'s observation that there is a "strong tendency for speakers to reuse a single set of highly entrenched neuromotor patterns and to substitute members of this set for novel or less common configurations".

The 236 languages in the SibInv database confirm this tendency: their mean exploitation score is 0.802 (including both voicing and affricates), and their mean symmetry score is 0.706 (including only fricative - affricate contrasts). The former is calculated as in the formula above (number of segments present in the sibinv / number of segments that would be present if existing contrasts were maximally exploited), with the "2 exp number of features required to characterise it" simplified as the equivalent "number of segments that would be present if existing contrasts were maximally exploited". The latter is calculated as the number of sibilant fricatives that have an affricate counterpart and vice-versa, divided by the number of segments that would be present if all fricatives and affricates had a manner counterpart (excluding inventories without sibilants or without affricates). For example, the perfectly symmetrical inventory in (62) would have a symmetry score of 4/4 = 1; English would have a symmetry score of 2/4 = 0.5, and standard French would have no symmetry score (because there are no affricates to compare to fricatives). Interestingly, the fact that the exploitation score (which takes voicing into account) is higher than the symmetry score (which doesn't) confirms the observation made above that voicing symmetry is more strongly established in the languages of the world than fricative - affricate symmetry.

4.7.1.3 Comparison with random inventories

Comparing the 236 SibInv inventories with randomly generated ones, a statistically highly significant difference ($t = 15.407$, $p < 0.001$) arises. If one gives one 'asymmetry point' to every sibinv for every sibilant fricative that does not have a corresponding sibilant affricate at the same PoA and for every sibilant affricate that does not have a corresponding sibilant fricative at the same PoA, the logical mean 'asymmetry score' should be 0.5 (i.e. 2/4 for 4 PoAs, because every sibilant has 0.5 chance to exist and 0.5 chance to have a corresponding fricative or affricate). In randomly generated inventories, this 'asymmetry score' turns out to be close to this logical mean probability (2.047), while the 'asymmetry score' in SibInv inventories is significantly much lower (0.754). The fact that the 'asymmetry score' in the SibInv database is 2.71 times lower than what is expected from the logical probability provides strong evidence for the fricative - affricate symmetry universal.

Moreover, I also compared the mean exploitation score of the languages in the SibInv database (0.802) with that of randomly generated languages (0.491). This difference is statistically highly significant ($t = 22.156$, $p < 0.001$). Natural languages are thus not only much more symmetrical wrt. their sibilant fricatives and affricates, but also much more economical in their entire sibilant inventory, i.e. when also taking voicing into account. This corroborates what Mackie and Mielke (2011) observed for segment inventories in general. The mean number of fricative - affricate gaps was 0.651 in the SibInv, and 3.610 in randomly generated inventories ($t = -27.439$, $p < 0.001$).

4.7.2 Motivations for the fricative - affricate and voicing symmetry universal

4.7.2.1 Teleological or emergent?

The fact that the symmetry universal is emergent has already been explained in 4.1.1.2. There, it was observed that one does not need to posit the existence of e.g. fricative - affricate symmetry constraints, because symmetry emerges among others from parallel sound shifts (e.g. a language palatalises /ti/ and /si/ at the same time, which will yield symmetrical /t͡ɕ/ and /ɕ/). If a sound shift phonemicises a voiced sibilant from a voiceless one, it is expected that both will have the same PoA.

An argument against symmetry as a teleological principle is that it would imply that speakers (in an OT perspective) evaluate not only the segments in their native language, but also their paradigmatic relation to each other. Evaluating segments simply requires articulatory/sensorimotor knowledge of what articulatory configuration is harder or easier, while evaluating the relation between segments (in contrasts, or here: in their articulatory symmetry) would require a metalinguistic knowledge of how one's native language is structured. Boersma and Hamann (2008: 225f.) formulate this argument in a similar way when talking about auditory dispersion in sibilant inventories.

From a diachronically oriented perspective, the non-teleologicity of symmetry has also been noted by Blevins (2004: 281f.), who states among others that there is "no evidence of goal-oriented change in the service of symmetrical inventories". However, "[d]espite the conclusion that sound change is not driven by gap-filling imperatives whose ultimate goal is symmetrical inventories, certain symmetrical tendencies in the structure of segment inventories clearly exist" (Blevins 2004: 284). In the case of sibilant inventories, the explanation resides both in "gestural efficiency" and the direct "result from sound change" (Blevins 2004: 284). Result from sound change has already been explained above in 4.1.1.2, with examples from English and French. The following subsubsection explains how symmetry arises as a direct consequence of the OT grammar of a given sibilant inventory.

4.7.2.2 Articulatory and perceptual motivations

As the sibilant OT typology in chapter 5 will show, structural/markedness constraints against certain sibilant PoAs, against voiced and against affricate sibilants interact with a faithfulness constraint. Under a specific constraint ranking, foreign sibilants (with PoAs, voicing or manner ruled out by the native constraint ranking) for example will end up being produced as a native category, independently of the question whether a speaker perceives it as a foreign category at first or directly perceives it as a native category (see 5.2.3 for tableaux). Similarly, a new segment entering the sibilant inventory from within the native language is subjected to the same evaluation process. For example, stop palatalisation in certain modern regional varieties of standard French (e.g. *voiture* 'car') yields a phonetic output [t͡ɕ] that is not present from the start in the standard French inventory. Furthermore, modern standard French generally prohibits palatal segments (e.g. historical [ʎ] > [j], [ɲ] > [nj]; also, native French speakers tend not to distinguish e.g. Italian [ʎ]*i* 'them (dat.)' and [l]*i* 'them (acc.)' both in perception and production, confusing for example *gli do* 'I give [something] to them' and *li do* 'I give them [to someone]'). Palatal segments being prohibited in modern standard French (with a constraint like *PalatalSegment in Herrick 2002: 79), and [t͡s/d͡z] or [t͡ʃ/d͡ʒ] being the only sibilant affricates that the French sibilant inventory grammar allows, the result of palatalisation [t͡ɕ] is reanalysed by the grammar as /t͡s/ or /t͡ʃ/. In some varieties (e.g. Québec French), it becomes /t͡s/, while in others (e.g. Cajun French) it becomes /t͡ʃ/. It might thus have been heard as /t͡ɕ/ in the perception grammar, but the production grammar will prevent this mapping in any case. In such a way, a new sibilant affricate without a fricative counterpart is avoided non-teleologically.

Besides the interaction of production and perception described previously, there is nothing in perception that favours symmetry (e.g. whether an inventory has /s, t͡s̠/ or /s, t͡ʂ/ does not make it better or worse in perception, if not in terms of auditory dispersion, centralisation etc.). Martinet (1968: 483), cited in Clements (2003: 293f.), explains that an advantage and motivation of symmetry in inventories can be that reducing the number of different gestures reduces the probabilities of misperception, i.e. that "a potential feature combination will be disfavoured if the resulting sound cannot be easily distinguished from another sound in the system" (Clements 2003: 293). This pertains in fact to auditory dispersion, and will thus be addressed in the next section.

4.8 Auditory dispersion

With auditory dispersion as a principle, it is meant that languages tend to have contrasting PoAs for sibilants with preferably distinct CoGs, corresponding to more distinct articulatory realisations. For example, English and French are considered to have well-dispersed sibinvs (/s - ʃ/), Catalan and Trentino dialects to have less-dispersed sibinvs (/s̠ - ʃ/ and /s̪ - s̠/, respectively). Auditory dispersion as a kind of sound shift has already been detailed in B.14.

4.8.1 Cross-linguistic evidence

Boersma and Hamann (2008: 217) state that "[i]t has often been observed that sound systems are structured in a way that minimises the perceptual confusion between its elements", a view already held by e.g. Passy (1890) (Seinhorst 2012: 3) and Martinet (1955: 236f.) (Vijūnas 2010: 42). For example, sibilant inventories tend to space their PoAs equally in perceptual dimensions; the more crowded the perceptual space, the smaller the distance between its elements, because the perceptual space itself is finite, as Boersma and Hamann (2008: 230) illustrate in fig. LX. The vast majority of the sibilant inventories of natural languages conforms to this universal of auditory dispersion, to the extent that it is rather worth counting the exceptions rather than the languages that conform to it (see 3.2.2.3). As noted in B.14, 129 out of 157 2s inventories in the SibInv database are well-dispersed, i.e. 82.17%. As mentioned there, auditory dispersion can be observed diachronically in the transition from Vulgar Latin to standard Italian, from Proto-Germanic to modern English or from Common Slavic to Serbo-Croatian. In these examples, already existing sibilants were moved away from the centre of the phonetic continuum to make space for new sibilants.

The interested reader is referred to e.g. Flemming (1996), Boersma and Hamann (2006), Padgett and Żygis (2007), Boersma and Hamann (2008), Flemming (2018) for models of this phenomenon in Optimality Theory with dispersion constraints (Flemming, Padgett and Żygis) as well as without (Boersma and Hamann). Noteworthily, van Leussen and Vondenhoff (2008) modelled the emergence of auditory dispersion as in Boersma and Hamann (2008) with pseudo-'real speakers' (i.e. simulated populations of speakers-listeners, mirroring real human populations) rather than a single speaker-listener. Using Neural Networks, auditory dispersion has been modelled by Seinhorst

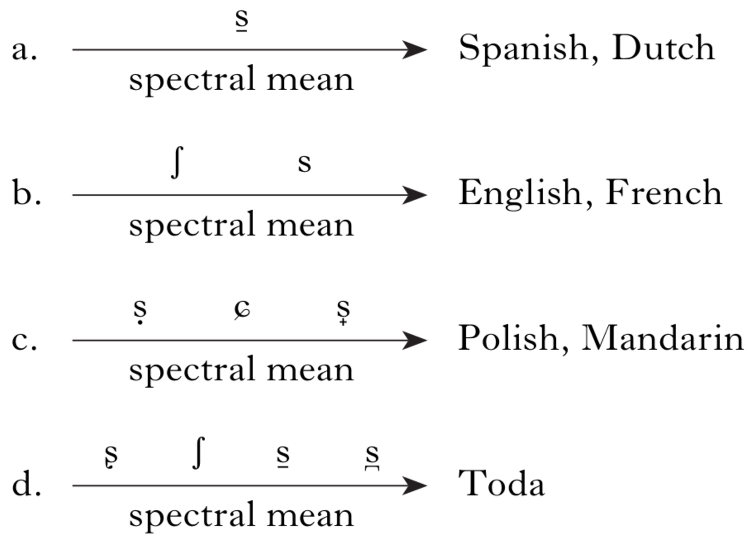


Figure LX: *Auditorily dispersed sibinvs along the perceptual scale from low to high CoG, taken from Boersma and Hamann (2008: 230).*

(2012) and Boersma et al. (2020).

4.8.1.1 Comparison with random inventories

Evidence in favour of auditory dispersion as a principle shaping the structure of sibilant inventories is provided by the comparison between randomly generated inventories and inventories in the SibInv database. If one assigns one 'bad dispersion point' to every sequence of two adjacent sibilants, assuming 4 PoAs and affricates, a maximal inventory with 4 sibilant fricatives and 4 sibilant affricates would have a total of 6 'bad dispersion points'. This is because a sequence 'SSSS' (i.e. an inventory with sibilants at 4 out of 4 PoAs) contains 'SS' three times. The logical probabilities in random inventories are: 8/16 of having 0 'bad dispersion point' (one 0s, four 1s and three well-dispersed 2s), 5/16 of having 1 'bad dispersion point', 2/16 of having 2 and 1/16 of having 3. Ruling out the 5/16 cases in which auditory dispersion is not applicable (i.e. ruling out 0s and 1s), this means that the average number of 'bad dispersion points' in random inventories must approximate 1.0909 given 4 PoAs, and 2.1818 given 4 PoAs and affricates. The 236 randomly generated inventories indeed have an average 'bad dispersion score' of 1.610, whereas this score is 0.607 in the SibInv set, i.e. 2.65 times less. This difference is highly significant ($t = 8.110$, $p < 0.001$), demonstrating the influence of auditory dispersion on sibilant inventory structures in natural languages.

4.8.2 Motivations for the auditory dispersion universal

4.8.2.1 Teleological or emergent?

Several approaches (Flemming 1995; Padgett and Żygis 2007; Flemming 2018) have formalised auditory dispersion with dispersion constraints in OT. This implies a teleologicity with which it can be dispensed: "language users do not aim at improving their sound systems" (Boersma and Hamann 2008: 217) and "there is no need to posit a dedicated inventory grammar: inventories are the automatic result of the constraints and their rankings in the production [JK: and perception] grammar" (Boersma 1997*b*: 59). Boersma and Hamann (2008) and Boersma et al. (2020) show convincingly how it can be assumed that auditory dispersion is emergent: without the help of improbable metalinguistic, system-oriented constraints, language learners start to discern the different categories of their native language and when reusing this ranking (in OT) or these connection weights (in NN) in production, well-dispersed sibilants emerge.

4.8.2.2 Articulatory motivations

Auditory dispersion, as the name itself already hints, is certainly not articulatorily-based. As shown in the previous sections, articulatorily ideal sibilants are central and anterior. From the perspective of mere articulation, an inventory like /s̺ - s - s̻ - ɕ/ is optimal; from the perspective of mere perception, it is extremely ill-structured. Therefore, languages find "an optimal balance between minimisation of articulatory effort and minimisation of perceptual confusion" (Boersma and Hamann 2008: 228).

4.8.2.3 Perceptual motivations

Every phonetic realisation of a phoneme is slightly different from other phonetic realisations of the same phoneme by the same speaker (e.g. with CoGs of 5101.32 Hz, 5012.14 Hz, 4698.07 Hz etc. for /s/). In a sibilant inventory with one PoA, this variation is not particularly relevant: any sibilant is perceived as belonging to one and the same category anyway. In sibinvs with more than one PoA, this variation becomes important: sibilant realisations close to another acoustic category are potentially ambiguous and prone to misperception, while realisations towards the extremities of the continuum are much less prone to misperception. For example, in a 2s with /s - ʃ/, the sibilant with a CoG of 4698.07 Hz is much more likely to be misperceived as /ʃ/ than the one with 5101.32 Hz. Under this approach, Boersma and Hamann (2008: 235f.) show that a child learning English would rerank cue constraints in her perception grammar when a mismatch between a perceived token and a lexical category occur (e.g. when hearing 'I love Bashque' instead of 'I love Basque'). The innate ability of very young children to perceive fine-grained phonetic nuances (Lassettre and Donegan 1998: 354), combined with reranking of cue constraints (modelled with a Gradual Learning Algorithm in Boersma and Hamann 2008), leads to an effect known as the prototype effect: speakers will tend to consider as the most prototypical example of a category a more distinct example than what they normally use, i.e. particularly peripheral sibilants. The effect of auditory dispersion is then counterbalanced in production by articulatory markedness constraints. Boersma and Hamann (2008) concentrate on auditory dispersion on one continuum, namely the one of spectral mean (CoG), but it is also reflected in distributedness distinctions: Basque contrasts laminal - apical - laminal sibilants, while Polish does so with apical - laminal - apical, as shown in 3.2.2.4.

4.9 Conservatism and contact

A rather obvious observation about natural languages is that children aim to pronounce the sibilants of their native language as closely as possible to the 'acoustic exemplars' to which they are exposed. Rather than redesigning the entire language as they please (which would make it unintelligible from one generation to another), speakers of a language aim to understand and, most frequently, to be understood: therefore, it is necessary to align one's grammar and pronunciation on those of other speakers. Besides intelligibility, there is a high social pressure to pronounce phonemes 'correctly', and sibilants being particularly acquired late and prone to speech defects like lisping encourage children to master sibilants identically to the model given by the surrounding (Nutterville 1934; Anjos et al. 2018), thereby reducing the chances of e.g. "being excluded from social groups for 'sounding different'" (Krueger 2019: 89).

This has implications inside one and the same language (conservatism), with speakers striving to have the same sibilant pronunciations on the surface, and between languages (contact), with languages in contact trying to have the same sibilants on the surface, though not necessarily the same sibinv grammar. The former is modelled in OT as faithfulness constraints, and the latter emerges from imperfect L2 acquisition (similarly to the phonemic reanalysis detailed in 4.7.2.2).

4.9.1 Cross-linguistic evidence

4.9.1.1 Diachronic stability

Boersma and Hamann (2008: 245-250) show that sibilant inventories change progressively rather than abruptly, across several generations. This is observed e.g. in Icelandic, Faroese and Norwegian, which reflect three diachronic stages spreading over at least 1000 years, from a 1s across a less-dispersed 2s with /s̺ - ʃ/ to a well-dispersed 2s (Vijūnas 2010: 47f.). Even stronger evidence for conservatism is provided by the *absence* of change: the Proto-Indo-European */s̺/ has maintained itself stably over the course of +/- 3500 years in Danish, Icelandic, Central Greek, Venetian Italian or Hollandic Dutch (Adams 1975; Vijūnas 2010).

4.9.1.2 Areal and contact phenomena

The influence of language contact on sibilant inventories can be observed universally: in Basque, bilingual native speakers of Castilian Spanish provoke or accelerate the merger of sibilants to their single PoA (Alvarez Enparantza 1982: 19; Muxika-Loitzate 2017); Castilian Spanish influence also contributes to sibilant devoicing in Catalan (Benet et al. 2012); Afrikaans was a 1s at first, but became a well-dispersed 2s through contact with French, German, English and Bantu languages; Flemish dialects differ from Hollandic Dutch in having /s, z, ʃ, ʒ/, precisely like French (e.g. Verhoeven 2006 for West Limburgish, de Vriendt 2006 for Brussels Flemish). This areal convergence through contact is illustrated on the three maps in LXI - LXIII, which attempt to represent sibilant inventories in Europe accurately, around the years 200, 1200 and 2020.

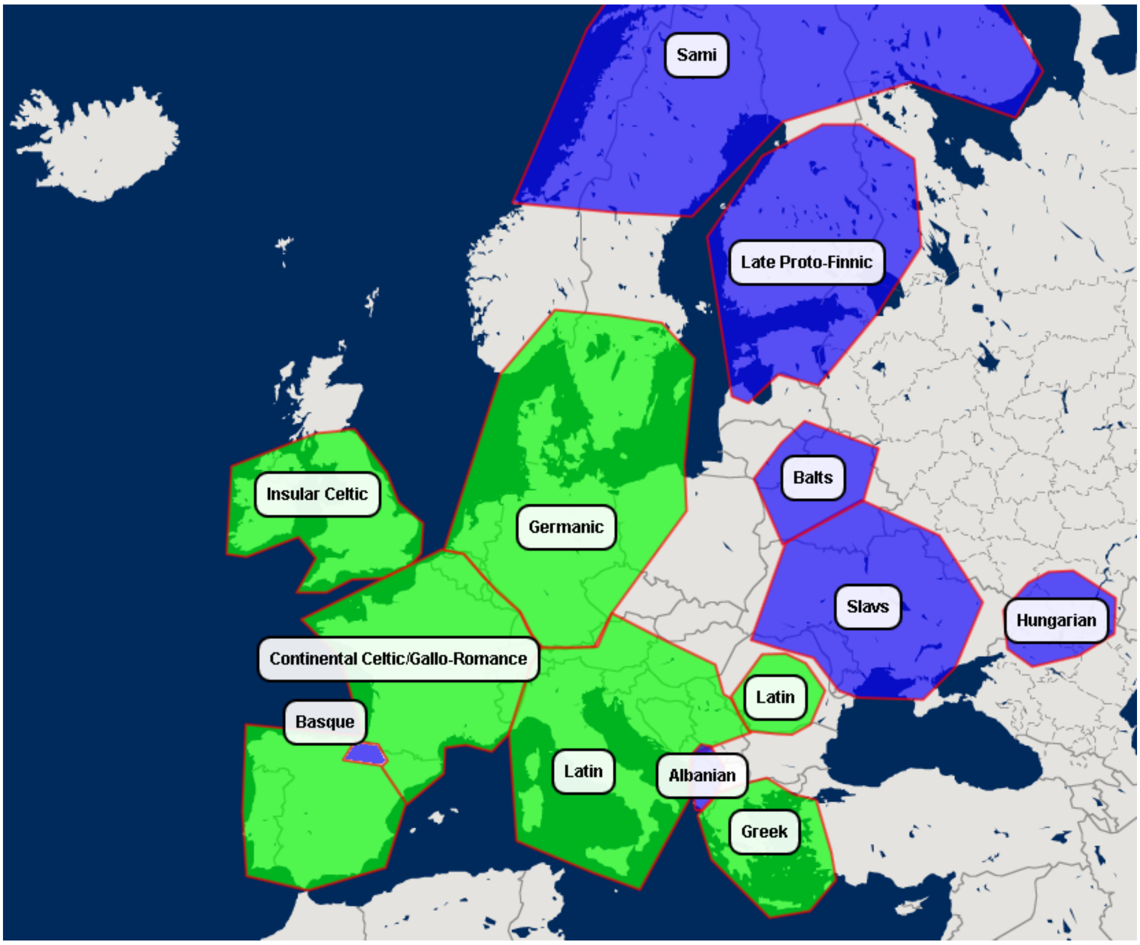


Figure LXI: *Approximate map of sibilant inventories in Europe around the year 200. Green indicates a 1s, blue a well-dispersed 2s.*

4.9.2 Motivations for the conservatism and contact universal

4.9.2.1 Teleological or emergent?

Inside one and the same language, conservatism is a teleological principle: speakers aim deliberately at producing precisely the sibilants that occur in their native language. In contact situations, speakers in fact do aim at producing precisely the sibilants that occur in the *non-native* language. If this succeeded perfectly everywhere, contact phenomena such as the ones reported in 4.9.1.2 would not occur at all, because bilingual speakers would keep the two phonologies distinct (which only skilled bilinguals can do) and not import one's processes and pronunciation in the other. However, since very few speakers are perfectly bilingual, L1 phonology is applied to the L2 language by reusing the constraint ranking of the L1 (Boersma and Hamann 2009).

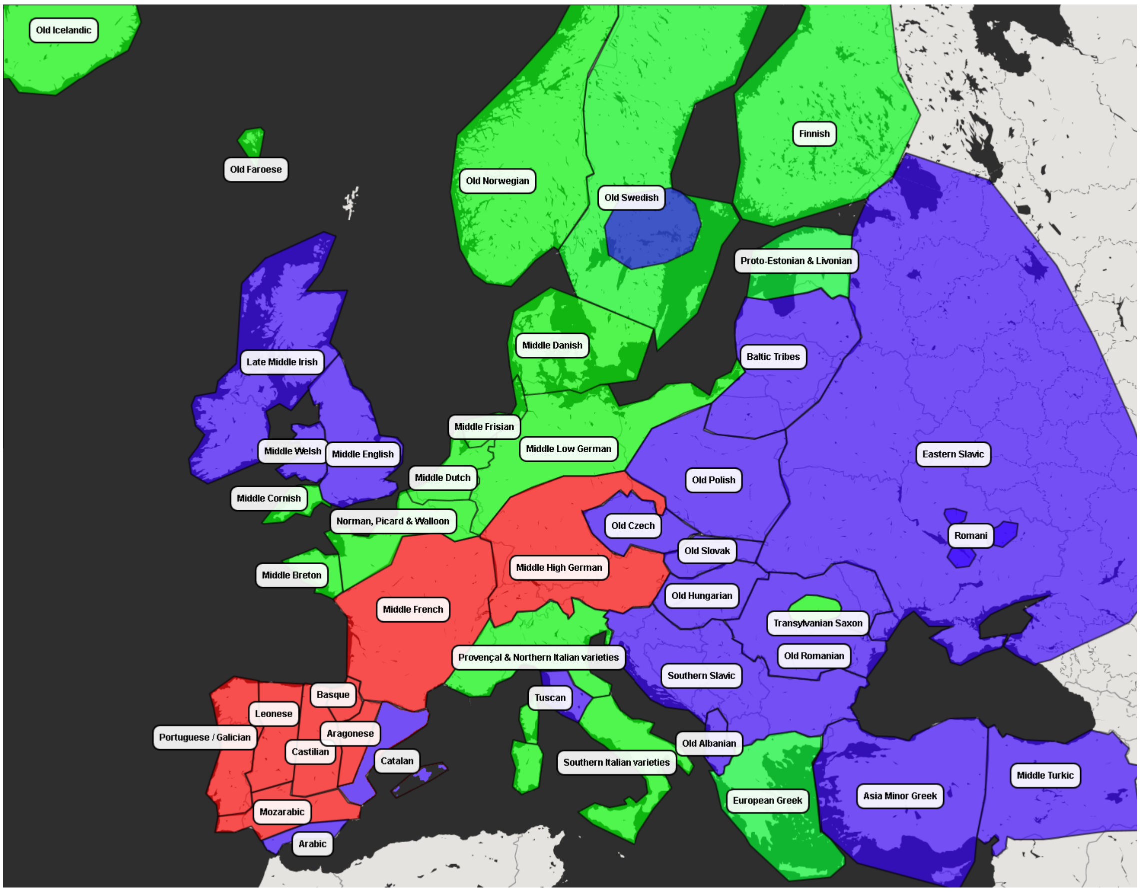


Figure LXII: *Approximate map of sibilant inventories in Europe around the year 1200. Green indicates a 1s, blue a well-dispersed 2s, red a non-retroflex 3s.*

4.9.2.2 Motivations

The motivations for this principle are not really articulatory or perceptual; instead, they are grounded in intelligibility and social pressure that trigger faithfulness between input and output, as explained at the beginning of this section.

4.10 Summary

To summarise the principles and their interaction in sibilant inventories, one can say that there is a force 'pulling' towards the centre (centralisation), one towards the front (posterior markedness), one pulling sibilants away from each other (auditory dispersion) and one striving to maintain sibilants where they are. Independently from this interaction of principles at the level of PoA, there is a markedness principle against sibilants in general, one against voiced sibilants and one against sibilant affricates. Finally, the principle of fricative - affricate and voicing symmetry emerges as a consequence of the bidirectional production - perception grammar and parallel sound changes.

Since a table says more than many words, the principles are illustrated in tables 3 and 4, with a link to each principle's section and statistical results of the comparison

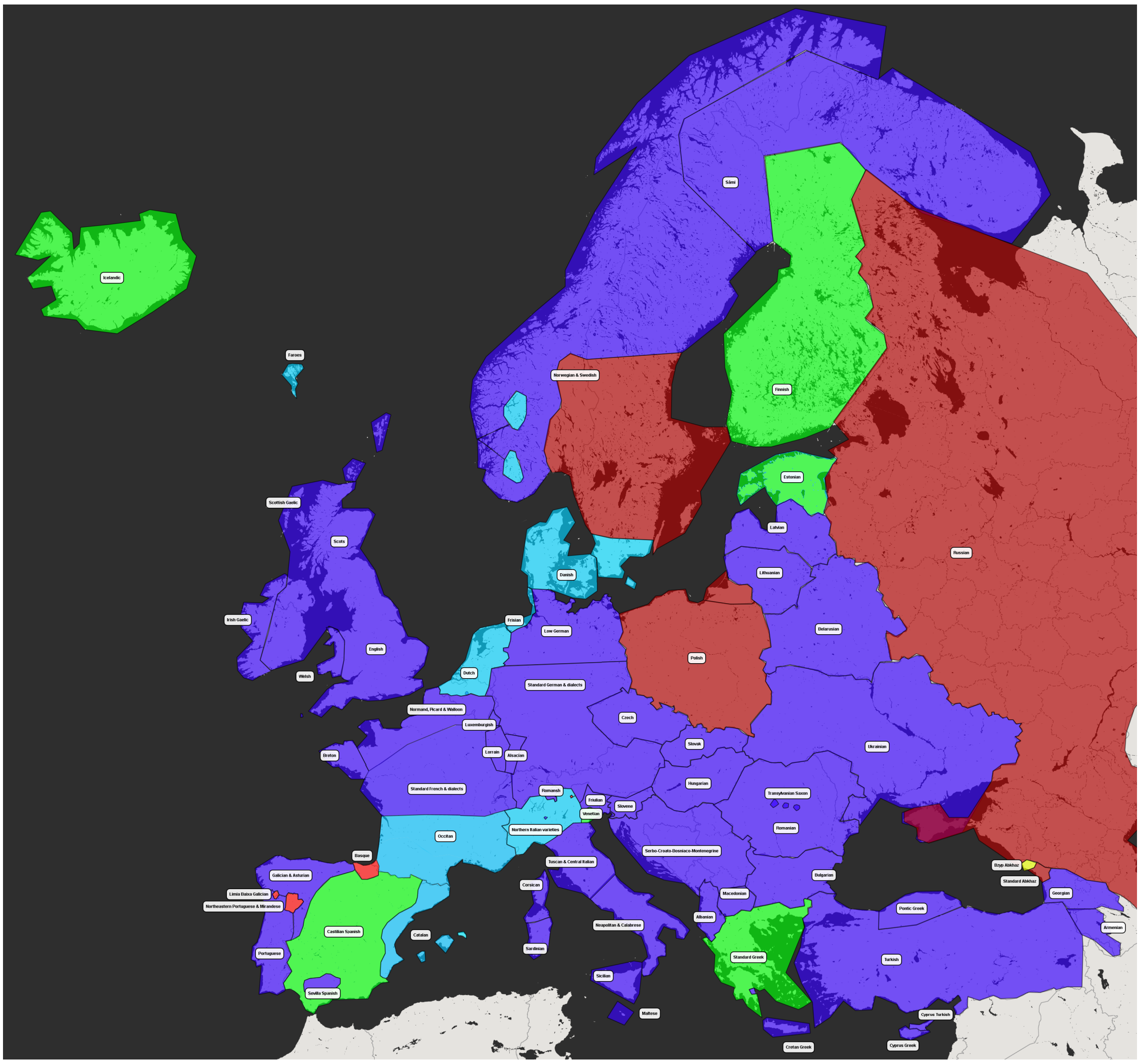


Figure LXIII: *Approximate map of sibilant inventories in modern-day Europe. Green indicates a 1s, blue a well-dispersed 2s, light blue a less-dispersed 2s, plain red a non-retroflex 3s, Bordeaux red a retroflex 3s and yellow a non-retroflex 4s (in the Caucasus).*

with randomly generated inventories.

In several cases, the motivations for a principle are both teleological and emergent and/or both articulatory and perceptual, like e.g. voicing markedness: it is then hard to distinguish between both motivations, since their effects are the same. In the end, the only consequence of having both motivations might be that the effects are simply reinforced. These principles with both kinds of motivations are: number of PoAs (sibilant markedness), potentially posterior markedness (depending on how one considers the frequency of [t͡ʃ]), voicing markedness and affricate markedness.

1.	Number of PoAs and sibilant markedness	favours disfavours	0-4 PoA contrasts **sibilants**
2.	Centralisation	favours	**central sibilants**
3.	Posterior markedness	favours	**anterior sibilants**
4.	Voicing markedness	favours	**voiceless sibilants**
5.	Affricate markedness	favours	**sibilant fricatives**
6.	Fricative - affricate and voicing symmetry	favours	sibilant affricates with the same PoAs as sibilant fricatives
7.	Auditory dispersion	favours	perceptually evenly spaced sibilants
8.	Conservatism and contact	favours favours	**faithful mappings** having PoAs identical to those of neighbouring varieties

Table 3: *Summary of the principles that shape sibilant inventories. Principles directly encoded in grammar are in bold.*

	Principle	***t*-value**	***p*-value**
4.2	Number of PoAs and sibilant markedness	2.551	0.011
4.3	Centralisation	6.639	<0.001
4.4	Posterior markedness	4.241	<0.001
4.5	Voicing markedness	18.337	<0.001
4.6	Affricate markedness	10.585	<0.001
4.7	Fricative - affricate and voicing symmetry	15.407	<0.001
4.8	Auditory dispersion	8.110	<0.001
4.9	Conservatism and contact	n.a.	n.a.

Table 4: *Summary of the comparison between attested and random sibilant inventories for each principle (Student's* t*-test), with* p*-values.*

5. Typological models of sibilant inventories

Chapter 3 shed light on the diversity of attested sibilant inventories and thereby especially on recurrent patterns, suggesting that several principles are responsible for these recurrences. Chapter 4 then proposed a list of eight principles and described them, arguing for their influence on sibilant inventory structure based on various types of evidence. Now that it has been shown what the structures of sibilant inventories are in natural languages and that we have posited what principles determine these structures, it remains to be seen how one can model the relation between the sibinvs observed on the surface and the principles posited to influence them by allowing certain inventories to exist while ruling out other inventories. By doing so, one can finally gain a full insight into *how language works* with respect to sibilant inventories: that is, how changes in the underlying grammar are reflected on the surface, and how sibilants on the surface are (re)interpreted by the grammar. For this, in the end, is the goal of typology and phonological models: a "search to define the limits, patterns, and explanations that characterize cross-linguistic variation" (Epps 2010: 635). While chapter 3 has mainly been about finding the limits and patterns and chapter 4 about finding the explanations, this chapter aims to model their interaction explicitly from the perspective of phonological typology. Like phonological typology in general, its objective will thus be "to model and explain the human knowledge of and capacity to acquire phonological systems" (Gordon 2016: 1), in this case to model the typology of underlying grammars that yield the attested possible sibilant inventories.

In the first section of this chapter, the existing studies that provide a typological overview or model of sibilant inventories are mentioned, and the strengths and weaknesses of these approaches are reviewed in order to make clear which aspects of previous models are to be integrated in the typological model presented here (i.e. legitimating the choice of frameworks and assumptions). A typology of sibilant inventories is then modelled successively in different frameworks: Optimality Theory (Prince and Smolensky 1993/2004) in section 5.2, bidirectional Neural Networks (Boersma et al. 2020) in section 5.3, a combination of OT and NN in section 5.4 and finally, Distinctive Feature Theory (Hall 2001; Clements 2009) in section 5.5, including an integration of DFT in OT.

Among these frameworks, it will be shown that the combination of OT and NN yields the typology that I consider to be an 'accurate' typology, namely a typology that limits as much as possible the risks of overpredicting or underpredicting (i.e. the criterion of *observational* adequacy applied as typological matching on the surface, in the sense of accurately determining what sibilants are allowed or not in all possible grammars, cfr. Green 2006: 49, Chomsky 1965: 24 and Boersma 1997*b*: 67) *and* that accurately explains the interaction of the principles in grammar (i.e. the criteria of *descriptive* and *explanatory* adequacy, in the sense of explaining *how* the underlying grammars are structured universally, cfr. Chomsky 1965 and Green 2006). As importantly as observational adequacy (especially, not failing to predict any attested inventory), the typology must indeed show us the workings that underlie the universal typology and these should be grounded in cognitively, articulatorily and/or perceptually motivated principles (e.g. the difficulty of producing voiced obstruents). Put in negative terms, a typological model that accurately matches the attested languages but is made of opaque and arbitrary *ad hoc* rules or constraints is as unlikely to be correct as a transparent and explanatory model based on independent evidence that does not match all attested languages. With evidence firmly grounded in cognition/articulation/perception, "we can seek a level of explanation deeper than explanatory adequacy, asking not only *what* the properties of language are, but *why* they are that way" (Chomsky 2001: 2).

The 'accurate' typology proposed in this chapter consists of an OT typology with sibilant segments as candidates and with several of the principles of chapter 4 translated as constraints, whose interaction yields six predicted sibilant inventories. Their acquisition and evolution is then modelled in Neural Networks (Boersma et al. 2020) in Praat (Boersma and Weenink 2019), where the predicted inventories turn out to be stable patterns in NN and unpredicted ones to be unstable, becoming predicted ones after several generations. The well-dispersed /s - ʃ/ sibilant inventory is (expectably) not predicted by the OT typology by lack of a dispersion constraint, but it is shown how it nevertheless emerges as a stable pattern in NN. To shed light on how generations of language users learn this pattern as a stable inventory, also in OT, I provide two possible OT accounts of its acquisition (in BiPhon-OT and with an ungrounded constraint, respectively).

5.1 Previous typological models of sibilant inventories

Sibilants have received a significant amount of attention in the literature on sound shifts and particular phonological processes linked to them (e.g. sibilant harmony, stop palatalisation or to a lesser extent, s-retraction), as well as in the phonetic literature describing their articulatory and/or acoustic characteristics in a given language. However, this contrasts strongly with the scarcity of typological studies who attempt to determine which sibilants can occur in one and the same inventory and which cannot, i.e. comprehensive research on all sibilant inventories in the languages of the world. To my knowledge, this was done by Boersma and Hamann (2008)/Seinhorst (2012)/Boersma et al. (2020) in BiPhon, Flemming (2018) in ADT and Hall (1997*b*)/Clements (2009) in DFT. To some extent, Berns (2013) provides a typology of sibilant affricates, based on the UPSID database (Maddieson 1984), although it includes non-sibilant affricates as well. The focus is thereby set on affricates in particular, and on the issue of underlying representations rather than on explanations for the typological distribution of the sibilant affricate inventories (partly addressed in Berns 2013: 82-85).

Boersma and Hamann (2008: 229ff.) illustrate attested types of sibilant inventories in detail, thereby providing a rich description of possible sibinv structures (although they do not explicitly claim to describe *all* attested types of sibinvs). Their contribution highlights the role of centralisation and auditory dispersion in sibilant inventories. They mention the 4 sibilant inventories illustrated in figure LX, showing that these tend to have both as central and as well-dispersed sibilant PoAs as possible. This means that they mention 4 of the 7 inventories (with 4 PoAs) found in section 3.2 in this dissertation. Inventories absent from their model are 0s (which is logical, since they address sibilant-related phenomena that are irrelevant in sibilantless languages), both types of less-dispersed 2s presented in 3.2.2.3 and the non-retroflex type of 3s. They state that "[i]f a language has two sibilants, neither of them has a central spectral mean" (Boersma and Hamann 2008: 230), something which excludes the possibility of the existence of less-dispersed 2s such as Faroese (/s̠ - ʃ/) and Veronese regional standard Italian (/(t͡)s̠ - s̠/). Nevertheless, such languages are predicted by their model to exist as *unstable* stages in the diachronic development towards stable inventories (somewhat like the initial distribution in their 'skewed and confusing English' simulation, cfr. Boersma and Hamann 2008: 249). Less-dispersed inventories are indeed rare and less stable (e.g. Hollandic Dutch with its /s̠ - ɕ/ contrast neutralised in Amsterdam, cfr. Ooijevaar 2011), but I intend to show that they are more stable diachronically than expected (see 3.1.4).

In subsequent work on auditory dispersion with sibilants, Boersma et al. (2020) and Seinhorst (2012) use Bidirectional Neural Networks to show that auditory dispersion also emerges when feeding a neural network with acoustic tokens taken from normal distributions (corresponding to sibilant phoneme categories) and letting the network produce sibilants afterwards. The types of sibinvs mentioned are 1s, 2s (both the well-dispersed and the less dispersed /s̠ - ʃ/ type) and 3s in Seinhorst (2012), while Boersma et al. (2020) focus mainly on the retroflex 3s. Their recent work within bidirectional Neural Networks does not necessarily aim at providing a detailed typological overview of sibilant inventories, and rather concentrates on explanations for and simulations of auditory dispersion and centralisation. Sibilant affricates are not part of their study. In total, they mention 5 out of 10 attested types of sibilant inventories (with 7 PoAs, as in table 2), and contribute to a better understanding of the interaction of centralisation and auditory dispersion.

Based on another theoretical perspective, namely that of Adaptive Dispersion Theory (Liljencrants and Lindblom 1972; Flemming 1996), Flemming (2018) recently proposed a typology of sibilant inventories also focusing on auditory dispersion, but contributing furthermore by incorporating posterior markedness into the model. It lists attested sibilant fricative contrasts and explains them by means of dispersion constraints in OT. It does not include retracted alveolar sibilants or sibilant affricates. Therefore, 0s, 1s, less-dispersed 2s and non-retroflex 3s, all having a retracted alveolar /s̠/ (or no sibilant at all), are absent from the typology proposed. Flemming (2018) does not distinguish plain from subapical retroflexes (e.g. Mandarin Chinese and NW Caucasian languages are said to have the same '/ʂ/') and one of the inventories he mentions (/s - ʂ/) does not exist as such (as /s̪ - s̠/ or /s - ʃ/ instead). Some languages he names do in fact have another sibinv (e.g. Burushaski is a typical retroflex 3s, cfr. Berger 1998: 13, 21f. and Munshi 2006: 58). The inventories in Flemming (2018) are thus not more complete than those in Boersma and Hamann (2008); in total, he mentions 4 out of 10 attested types of sibilant inventories and highlights the role of auditory dispersion as well as posterior markedness.

Clements (2009: 25ff.) does not propose a typology precisely of sibilant inventories, but he explains that inventories of coronal sounds universally behave as follows, something which naturally applies to sibilants (both fricatives and affricates) as coronal consonants. In his DFT analysis, there are two features ([±anterior] and [±distributed]) which, when combined, yield the four possible segments [+ant, +dist] = [s̪], [+ant, -dist] = [s̺], [-ant, +dist] = [ʃ̻], [-ant, -dist] = [s̠]. The occurrence of all four corresponds approximatively to a 4s like Toda. When not using all four feature combinations, this also predicts the existence of a 0s (no [+strident] segment at all), a 1s (with [+strident] but unspecified for [±anterior] and [±distributed]), a 2s with [±anterior], a 2s with [±distributed] (which is unattested), two 3s that are attested and two 3s that are unattested, besides other predictions detailed in 5.5.1. Clements illustrates the possible contrasts using voiceless stops (2009: 26), similarly to the [t̪ - t̺ - c - t̠] contrast illustrated in (64). Hall (1997*b*: 90-96) provides a detailed DFT typology of sibilant inventories in his monograph on coronal segments. He mentions 0s and 1s inventories, but omits retracted alveolar sibilants and reports the contestable existence of inventories with a single /ʃ/ or a single /ʂ/ (see Vijūnas 2010: 49 for evidence against 1s with /ʃ/ and Boersma and Hamann 2008: 230 for evidence against 1s with /ʂ/). When it comes to the 2s, he does not mention the /s̪ - ɕ/ nor the /s̪ - ʃ/ type, but mentions the /s̪ - s̠/ type (though without mentioning it as having a retracted alveolar sibilant) and the well-dispersed /s - ʃ/, besides an unattested /s - ʂ/ (depending on the analysis in 3.2.2.3) and a definitely unattested /ʃ - ʂ/ type. The Basque (non-retroflex) and the Polish (retroflex) types of 3s are listed, together with an unattested /s - ʃ - ʂ/ inventory (which might be intended as a /s - ʃ∼ʂ/ 2s like Eastern Norwegian). The 4s inventory is mentioned. To sum up, Hall (1997*b*) provides an extensive typology (6 out of 10 attested sibinvs, though with many additional unattested types), which he motivates as possible combinations of the two features [±anterior] and [±distributed], similarly to Clements. The Distinctive Feature model thus makes accurate predictions for sibilant inventories, as I show in 5.5.1.

The bidirectional OT/NN approach and the ADT approach have in common that they base their typologies on two conflicting forces: articulatory markedness (i.e. ease of articulation) and perceptual distinctiveness (emergent in the first, teleological in the second). The DFT approach does not explicitly appeal to these two forces, although Hall (2011) argues that one obtains dispersion effects in DFT without needing any teleological device. Many of the approaches mentioned do not incorporate posterior markedness (with the exception of Flemming 2018) or sibilant markedness as a principle, and some of them include voiced or affricate sibilants (and thus voicing markedness or affricate markedness). Approaches in DFT typically mention voicing symmetry or fricative - affricate symmetry, but the other approaches do not (as is expected for those who do not address voiced or affricate sibilants). Most of the eight principles evidenced in this dissertation are thus already present in the literature, although in a scattered way rather than all incorporated into one model. The typology I will propose aims to synthesise the existing accounts into a comprehensive model, taking so to say “the best from each model”, to provide a typology of sibilant inventories that satisfies all of these criteria:

- Mentioning 0s, 1s, the three types of 2s, 3s and 4s (i.e. not undergenerating)
- Not overgenerating (e.g. /s - s̡̠/, /s̠̻ - s̠̺/)
- Including voicing distinctions
- Including sibilant affricates
- Motivating the forces behind the typology and their interaction empirically

5.1.1 Which approaches are discarded and why

Among the existing models that aim at providing an insightful and accurate typology of sibilant inventories, certain aspects can be ruled out as e.g. cognitively unrealistic. In this subsection, I first argue that it can be dispensed with dispersion constraints (in 5.1.1.1) and with contrast maximisation constraints (in 5.1.1.2) in Adaptive Dispersion Theory. Then, I argue in 5.1.1.3 that approaches of learning a native sibilant inventory that require pre-existing knowledge of the sibilant phoneme categories should be avoided, similarly to what Boersma and Hamann (2008: 238, 258, 263) note about their own model. Besides this, approaches within Distinctive Feature Theory also have some weaknesses, but they compensate for them with some advantages and are thus detailed as a separate analysis in the dedicated section (5.5).

Hall (2011) proposes a DFT-based model of auditory dispersion. Since he provides a detailed description and evaluation of the previous approaches regarding auditory dispersion (in his section 2) that can partly complement the one given in this section, the interested reader is referred to this paper for similar and different arguments to the ones provided below.

5.1.1.1 Discarding dispersion constraints

In applications of Adaptive Dispersion Theory in OT (called "the dispersion theory of contrast" by Flemming 2002: 4), auditory dispersion is explained as the result of the interaction of three types of constraints. These three conflicting constraints are markedness constraints, which favour less effortful articulations, dispersion constraints, which favour more distinct contrasts, and contrast maximisation constraints, which favour inventories with more contrasts (Flemming 1995: 4). I argue against the use of the two last mentioned types of constraints (dispersion and contrast maximisation); markedness constraints (or in Boersma and Hamann 2006's terms, articulatory constraints or articulatory-based structural constraints), on the contrary, are to be kept because one cannot account for phenomena such as centralisation without them.

How dispersion constraints and contrast maximisation constraints work is exemplified by Itō and Mester (2003: 667f.), who illustrate this on the high vowel continuum (from [u] to [i]). Their candidates and constraints are defined (in a modified version, since they e.g. do not include the contrast maximisation constraint) in (71).

(71)

GEN: An *empty* input that can be mapped to four outputs, which are high vowel inventories with 1, 2, 3 or 4 phonemes, respectively.

CON: Constraints assigning 1 violation for the acoustic space being occupied by at least 2, 3 or 4 phonemes respectively, and a constraint assigning 1 violation for each phoneme absent from the inventory.

These candidates and constraints are shown in table 5. The empty input is not displayed. Every dispersion constraint (Space) assigns a violation mark based on whether the au-

					$\text{Space}_{VCol} \geq \frac{1}{3}$	$\text{Space}_{VCol} \geq \frac{1}{2}$	$\text{Space}_{VCol} \geq= 1$	MaxContrast
a.	i	ɨ̟	ɨ̠	u	*	*	*	
b.	i	ɨ		u		*	*	*
c.	i		u				*	**
d.	ɨ							***

Table 5: *Dispersion and contrast maximisation constraints applied to the F2 vowel space continuum for high vowels, taken and modified from Itō and Mester (2003: 667f.).*

ditory space available is greater or equal to $\frac{1}{x}$, with x being all integers between 1 and the number of segments -1. For example, an inventory with 2 phonemes can fit each in $\frac{1}{2}$ of the full auditory space between the frontmost [i] and the backmost [u], and this inventory therefore does not violate $\text{Space}_{VCol} \geq \frac{1}{2}$. The constraint Maximise Contrasts (short: MaxContrast) assigns one violation for every possible segment that is absent from the largest possible inventory. Languages in which the constraint MaxContrast is dominated by all dispersion constraints (as is the case in table 5) have a single /ɨ/, which is the case of the (default realisation of the) phoneme /ɨ/ in many NW Caucasian languages with only one phonemic high vowel. If MaxContrast is situated above $\text{Space}_{VCol} \geq = 1$ but below the other Space constraints, the winning candidate is a /i - u/ contrast, like in e.g. French; and so forth for inventories with three and four high vowels. Markedness constraints furthermore make sure that winning candidates are as central as possible (e.g. /ɨ/ in an inventory with one high vowel), but are not shown here in order to concentrate on the interaction of dispersion constraints and contrast maximisation constraints. This interaction is summarised in the screenshots from OT-Workplace (Prince et al. 2007-2021) in fig. LXIV, which account for the typology of the Violation Tableau (VT) given in table 5. One observes that such a typology is signifi-

Dispersion	VT.ini		*4	*3	*2	*0
input	output		Space1/3	Space1/2	Space1	MaxContrast
	[i ɨ̟ ɨ̠ u]	1_ 2_ 3_ 4	1	1	1	0
	[i ɨ u]	1_ 2_ 3_ 0	0	1	1	1
	[i u]	1_ 2_ 0_ 0	0	0	1	2
	[ɨ]	1_ 0_ 0_ 0	0	0	0	3

Dispersion	FacTyp
Inputs->	
AnAttestedLanguage	1_ 2_ 3_ 4
Russian	1_ 2_ 3_ 0
French	1_ 2_ 0_ 0
NWCaucasian	1_ 0_ 0_ 0

Figure LXIV: *Violation Tableau (left) and Factorial Typology (right) in OTWorkplace (Prince et al. 2007-2021) of the high vowel dispersion phenomenon presented in table 5.*

cantly different from 'classical' Optimality Theory in the sense that there is no input - output correspondence, and that entire inventories are evaluated as outputs.

Dispersion constraints will not be included in the typology of sibilant inventories for two main reasons: 1) their existence is difficult to motivate from a cognitive perspective and 2) a theory of auditory dispersion can work without them. The first argument, already mentioned in 4.2.2.1, 4.7.2.1 and 4.8.2.1 when arguing against the PoA, symmetry and dispersion universals as teleological principles, rests on the fact that it is highly unlikely that a speaker evaluates contrasts rather than phonological forms whenever (s)he speaks. A speaker of French would for example evaluate the contrast between /i/ and /u/ on the F2 dimension as illustrated in the example above, but also e.g. the contrast between /i/ and lower vowels, at any moment in which the speaker utters a word with /i/. This is not only unrealistic because of the burden it sets on linguistic production and perception, increasing drastically the number of evaluated forms, but also because it is inherently teleological, as if speakers were trying to actively increase their number of contrasts and as if they would evaluate not the segment itself but its relation to the inventory. This is also the criticism expressed by Boersma and Hamann (2008: 225):

> [Dispersion constraints] are explicitly teleological with respect to auditory dispersion. Furthermore, an empirical problem is that these constraints evaluate multiple inputs at a time: they can be said to evaluate whole inventories (Flemming 1995: 33ff., Boersma 1998: 361, McCarthy 2001: 226f.) or even entire languages (Padgett 2003: 311, Flemming 2004: 268). These constraints are therefore hard to reconcile with the single-input constraints introduced by Prince and Smolensky (...).

As Boersma (1997*b*: 75) observes, a typology with dispersion constraints without input - output correspondence is equivalent to any 'trivial' formulation of a real-life decision (such as e.g. choosing a means of transportation to go somewhere). This misses a central aspect of OT and Generative Phonology in general, namely the relation between input and output forms. As Hall (2011: 2) puts it, "the constraint hierarchy ceases to be a device by which languages select optimal surface forms, and becomes a device for selecting optimal languages".

The second important argument against dispersion constraints is that it is possible to model auditory dispersion without them, as Boersma (1997*b*: 74f.) and Boersma and Hamann (2008: 225) observe. If a model can work without an additional mechanism such as dispersion constraints, the mechanism in question is superfluous and it is preferable to eliminate it, as van 't Veer (2015: 47) also explains:

> What is important about the notion of inventory espoused in Boersma and Hamann (2008), is that it strikes a balance between the functional goals of 'Ease of Articulation' and 'Perceptual Clarity', without using teleological or imprecisely defined constraints. Following Boersma and Hamann (2008) in their application of Occam's Razor, I propose that Parallel Bidirectional Phonology and Phonetics is preferable to Dispersion Theory.

Ockham's principle indeed states that a theory must be preferred to another if both explain the same phenomenon with shared theoretical assumptions or requirements, but the first theory has additional assumptions to the very same purpose. Since auditory dispersion has been successfully modelled without inventory-oriented constraints (Boersma and Hamann 2008; Hall 2011), these constraints are unnecessary. Furthermore, eliminating dispersion constraints is also a good decision because "the dispersion constraints do nothing else beside explaining inventories" (Boersma and Hamann 2006: 3). By avoiding using them, one avoids appealing to constraints with one single function in the entire grammar.

5.1.1.2 Discarding contrast maximisation constraints

Another problematic aspect of Flemming's Dispersion Theory is that it uses the constraint MAXIMISE CONTRASTS. This constraint is necessary for analyses in this model because without it, inventories with only 1 (or even better: 0) segment would win regardless of the constraint ranking, as can be seen in table 5. The constraint works in a similar fashion to faithfulness constraints, in the sense that it prevents contrasts to be reduced. However, it is different in two major ways. Firstly, faithfulness evaluates the correspondence between an input and an output form; since there is typically no input - output correspondence in Dispersion Theory (e.g. Flemming 1995: 35, Bradley and Delforge 2006 or Padgett and Żygis 2007: 310; cfr. Hall 2011: 9), MAXCONTRAST does not behave as (I-O) faithfulness. Secondly, faithfulness is about preserving contrasts, i.e. *not reducing* them, while MAXCONTRAST – as the name clearly reveals – actively tries to increase the number of contrasts by favouring more and more contrasts. It is hardly realistic for a theory of grammar to have a principle that teleologically aims to increase the *number* of contrasts, especially considering that aiming to increase the *distinctiveness* of contrasts has already been shown to be improbable. Faithfulness works in such a way that a speaker uttering a sentence tries to articulate the segments of an underlying form faithfully, because the more (s)he alters the surface form wrt. the underlying form, the more likely it becomes that (s)he will be misunderstood. This is 'trying to stick to the conventional pattern', i.e. much the opposite of 'trying to change one's native language to include more contrasts', which goes against both faithfulness and markedness. Instead of this, new segments constantly appear in an inventory diachronically because satisfying markedness constraints (such as e.g. pronouncing /sj/ as [ɕ]) constantly creates new forms that can be phonemicised and thereby added to the underlying representation. In such a system in which markedness constantly creates new segments in specific contexts and deletes marked segments in general while faithfulness tries to keep everything in its right place, positing the existence of a constraint favouring more contrasts is superfluous, and should be dispreferred according to Ockham's logic.

Hall (2011: 8) adds that in principle, the fact that MaxContrast is a positive constraint (i.e. not penalising but favouring things in the output, namely contrasts) is problematic, because "if this constraint were to outrank all of the Mindist constraints, then the optimal output would be one containing infinitely many contrasting forms". Indeed, in the example given in fig. LXIV, I have formulated this constraint negatively, i.e. as penalising the absence of contrasts (which coincides logically with favouring the presence of contrasts). With this formulation, given the assumption of Richness of the Base (Prince and Smolensky 1993/2004), there should be additional candidates to the ones I listed in LXIV, for example '1_2_3_4_5' dominated by MaxContrast. This candidate should be preferred to what should then be '1_2_3_4_0', which was the winning candidate when MaxContrast dominated all others in the initial typology. '1_2_3_4_5', however, is not as good as '1_2_3_4_5_6', and so on. Therefore, MaxContrast is problematic as it elects infinitely better candidates if it dominates the Mindist constraints.

5.1.1.3 Discarding predefined categories in learning perception

Boersma and Hamann (2008: 238, 258, 263) note about their own model that "it is unrealistic to assume an initial state with fully random prelexical perception together with perfect lexical representations". They model the acquisition of a language with a /s - ʃ/ contrast by a hypothetical child who does not know at all whether certain ERB values (i.e. a perceived CoG) correspond to one or the other category (/s/ or /ʃ/), but who knows what the underlying form of the word (s)he hears should be (e.g. [s/ʃ]*eep* should be 'sheep'). This is anachronic, in the sense that the child cannot have a perfect lexical knowledge (e.g. /sɪp/, /ʃiːp/) before having learnt the perceptual distinction between /s/ and /ʃ/: having tuned one's identification of perceptual cues is a prerequisite for hearing the difference between two categories, and thus for storing these categories in the lexicon. Perfect lexical representations before accurate perception is thus so to say "putting the cart before the s(h)eep". This anachrony is a consequence of the fact that they use cue constraints that already contain the categories in their formulation itself (for example: the constraint *[26.6 Erb]/s/, meaning "a perceptual spectral value of 26.6 ERB should not be understood as a /s/ token"). In subsequent work with bidirectional Neural Networks, Boersma et al. (2020) modelled the auditory dispersion of sibilants in a similar fashion to the 2008 article *but* without needing to specify the categories beforehand. This echoes progress made in the field of the unsupervised learning of OT grammars, where "the problem of learning the ranking and lexicon simultaneously" has been tackled (Jarosz 2006: 50). The NN model in Boersma et al. (2020) is thus preferable to the OT model with the Gradual Learning Algorithm used in Boersma and Hamann (2008).

5.1.2 Which approaches are included and why

The typology proposed in this chapter, for the Optimality Theory part, is based on the technique of mapping each segment of the universal inventory into each segment of that same inventory, a technique already found in e.g. Prince and Smolensky (1993/2004) (inspired by Stampe 1969, 1979). This means that input - output candidate pairs are individual segments as they occur in the lexicon and at the surface (e.g. the word *Sasha*: <saʃa> → /saʃa/ → [saʃa], [ʃaʃa] or [s̠as̠a] etc.), rather than entire inventories. It has already been argued in detail above in favour of the segment-oriented approach, which does not require the assumption that speakers evaluate phonemic contrasts.

For the Neural Networks part, the typology is based on the approach in Boersma et al. (2020) for several reasons. Firstly, the assumption of bidirectionality (i.e. that speakers use the same constraint ranking/connection weights in perception and production) lets auditory dispersion emerge without assuming that speakers evaluate phonemic contrasts. Auditory dispersion emerges because acoustic values in 'ambiguous' perceptual zones are not associated with a specific pattern as strongly as acoustic values in peripheral perceptual zones. Furthermore, as the authors put it, their "model is able to handle two seemingly disparate phenomena at the same time: gradual category creation and auditory dispersion" (Boersma et al. 2020: 103f.). It can also account at the same time for "the behaviour of the speaker, the listener, and the learner" (Boersma et al. 2020: 106) and is thus efficient because of its ability to tackle several problems within one model.

Besides this, the NN model is biologically plausible, in the sense that the way in which it works parallels the way in which human (and animal, cfr. Hennequin 2017) cognitive behaviour works: categories as mental representations are created gradually through experience in a way which reflects "complex dynamics previously observed in animal cortex (...) during both learning and performance of flexible behavior" (Miconi 2017). This is explained in detail in Seinhorst (2012: 15-17): different phonological patterns correspond to different neural activation patterns.

A third advantage is that it reconciles discrete, categorical representations with the gradient realisations or changes observed on the phonetic surface: neural networks "are capable of modelling phenomena that reveal gradient properties of phonological categories" (Seinhorst 2012: 18). It integrates the stochastic variation present in the articulation of sibilants by native speakers (recall that typically, a speaker does not produce precisely the same sibilant twice in a row in terms of articulatory and acoustic parameters, cfr. 2.3.1) and variation in the transmission noise, leading to category creation as a normal probability distribution. This is an advantage over analyses that do not manage to reconcile the categorical and the gradient aspects of phonology and phonetics.

In the typology I propose in 5.4, combining OT and NN, I model a typology of all possible sibilant inventories that combines the evaluation of phonemic categories on the phonological side (OT) and the concrete realisations of these categories on the phonetic side (NN). It is minimalistic among others in representing acquisition, production and perception with two speakers, one 'adult' and one 'child learner'. As I show in 5.4, it accurately mirrors observed acquisition patterns and integrates the stronger aspects of existing models (minimal assumptions, no teleologicity in perception and inventory-optimisation, no predefined categories for the learning child).

5.2 A sibilant typology in Optimality Theory

In this section, I propose an OT typology of sibilant inventories, named SibInv, based on the evaluation of single sibilant segments. The proposal is formulated in two parts: the basic typology, named bSibInv, which includes only principles related to place of articulation contrasts (the constraints f.Ident(Sib) for faithfulness, m.Sib for sibilant markedness, m.Periph(Sib) for centralisation, and m.PostCD(Sib) as well as m.PostD(Sib) for posterior markedness), and the extended typology, named eSibInv, which adds affricate markedness (m.Affr(Sib)) and voicing markedness (m.Voice(Sib)) to the constraints of the basic typology. As I will show, the gist of the typology, i.e. its most fundamental, complex and interesting constraint interactions, is concentrated in bSibInv. The two additional constraints of eSibInv let it match the attested inventories of the SibInv database, but work in a way that is fairly common in the OT literature (namely each of the two interacting individually with f.Ident(Sib) only).

The language with a well-dispersed /s - ʃ/ contrast (named AC in the typologies) is not predicted, because none of the typologies presented here include any dispersion constraint. This is a direct consequence of the decision made above to avoid the use of dispersion constraints (see 5.1.1.1). As will be shown in detail in 5.2.2.6, the AC language that should have been predicted but wasn't is in fact structurally identical to the predicted ABC language, except that underlying forms with retracted alveolar sibilants (B) do not surface as such (B) but either as A or C. It is shown how BiPhon-OT and an ungrounded constraint can shed a light on the acquisition of an AC language, each in its own way.

In the next subsection (5.2.1), I show how the principles of chapter 4 are 'translated' to OT, defining bSibInv.GEN, bSibInv.CON, eSibInvGEN and eSibInvCON. In subsection 5.2.2, I then describe in detail the OT typologies (bSibInv and eSibInv) that follow from the definition of GEN and CON, providing insights into their inner workings with the help of a Property Analysis and typohedron. In subsection 5.2.3, aspects of how the SibInv typologies work in production and perception are addressed.

5.2.1 From principles to practice

In this subsection, it is detailed how the principles will be included in the OT typologies of 5.2.2. Note that I define GEN and CON for the SibInv typologies here in only *one* way, to the exclusion of many other potential typologies resulting from many other potential definitions of GEN and CON. In the appendix (8.2), I motivate in detail the choice and definition of constraints and candidates made here and show why alternative typologies resulting from alternative GEN and CON definitions are to be discarded (e.g. because of under- or overgeneration). Readers can thus so to say find the essential *what* here and the detailed *why* in the appendix.

Recall that in Optimality Theory (Prince and Smolensky 1993/2004), a typology is the outcome of three interacting parts, namely GEN, CON and EVAL:

(72) GEN, CON, EVAL

GEN: A set of input - output candidate pairs.

CON: A set of constraints that assign a certain number of violations to outputs, or to the relation between inputs and outputs.

EVAL: The evaluation of the candidate pairs by the constraints when ordered in a certain ranking, and the selection of the winning candidate(s).

• **bSibInv.GEN**: In the basic sibilant typology, the candidates are input - output mappings between segments articulated at or perceived as being articulated at the 4 places of articulation for sibilants distinguished by natural languages, in accordance with the A-B-C-D nomenclature already used in table 2 in chapter 3. This is shown in (73).

(73) Inventory elements

A = a dental or alveolar sibilant

B = a retracted alveolar or alveopalatal sibilant

C = a palatoalveolar or plain retroflex sibilant

D = a subapical retroflex sibilant

n = a non-sibilant segment (articulated at any PoA)

The abstract non-sibilant candidate (n) can be thought of as any non-sibilant segment, although it typically is a non-sibilant coronal obstruent (e.g. /t/). The universal inventory of bSibInv thus consists of sibilant segments with these 4 sibilant PoAs and a non-sibilant segment whose articulatory/perceptual specification is not crucial to our argument. The universal inventory is 'mapped into itself' in bSibInv.GEN as illustrated in (74).

(74) Input - output mappings

UI = {A, B, C, D, n}

Input = any element of UI

Output = any element of UI

Candidate = $x \rightarrow y$ where $x, y \in$ UI

In other words, each of the 5 segments is a possible input for each of the 5 segments as an output, meaning that there are 25 possible mappings as listed in table 6. • **bSibInv.CON**: The constraint set of bSibInv is given in (75).

input → output	input → output	input → output	input → output	input → output
A → A	B → A	C → A	D → A	n → A
A → B	B → B	C → B	D → B	n → B
A → C	B → C	C → C	D → C	n → C
A → D	B → D	C → D	D → D	n → D
A → n	B → n	C → n	D → n	n → n

Table 6: *Possible input - output candidate mappings of the basic sibilant typology.*

(75) bSibInv.CON

m.Sib= *A, B, C, D — Not violated by: n

m.Periph(Sib)= *A, D — Not violated by: B, C

m.PostCD(Sib)= *C, D — Not violated by: A, B

m.PostD(Sib)= *D — Not violated by: A, B, C

f.Ident(Sib): Custom function (see (76) below).

m.Sib corresponds to the principle of sibilant markedness (4.2), m.Periph(Sib) to centralisation (4.3), m.PostCD(Sib) and m.PostD(Sib) to posterior markedness (4.4) and f.Ident(Sib) to conservatism (4.9). The markedness constraints, prefixed with m, are defined respectively as assigning a violation mark for every *sibilant* in the output, for every *non-central sibilant* in the output, for every *posterior sibilant* in the output and for every *subapical posterior sibilant* in the output. The posterior markedness constraints m.PostCD(Sib) and m.PostD(Sib) form a stringency hierarchy (Prince 1999; de Lacy 2002; Merchant and Krämer 2018), where one constraint penalises a subset of what another penalises. The faithfulness constraint f.Ident(Sib), prefixed with f, is defined in (76).

(76) Custom faithfulness function

s.UI = {A, B, C, D}

S[A] = 1, S[B] = 2, S[C] = 3, S[D] = 4

When both input and output are sibilants:

Candidate = x → y where x, y ∈ s.UI

f.Ident(Sib)(x → y) = |S[x]−S[y]|

Between sibilants and non-sibilants:

Candidate = x → y where x ∈ s.UI and y ∉ s.UI or x ∉ s.UI and y ∈ s.UI

f.Ident(Sib)(x → y) = #s.UI = 4

This function penalises mappings between sibilants with different PoAs, proportionally to the distance between the two PoAs on the scale A > B > C > D: for example, a mapping from A to A obtains 0 violation mark, a mapping from C to D 1 violation mark, a mapping from B to D 2 violation marks and a mapping from A to D 3 violation marks. The faithfulness violations are thus assigned *gradually*, proportionally to distance on the

PoA scale from A to D. If a sibilant is mapped to n or inversely, the mapping obtains a number of violations equal to the number of sibilants in the set, i.e. 1 more than the mapping between sibilants with the largest distance (in this case, 4). The custom VBA function that assigns F.IDENT(SIB) violations in OTWorkplace (Prince et al. 2007-2021) is provided in 8.6.1 in the appendix. It converts the alphabetic characters of the input and output strings to their ASCII number, calculates the difference in ASCII numbers between the characters at place N in the input and output strings, and returns the sum of differences (or 4 if the input or output is a non-sibilant and the other is a sibilant).

• **eSibInv.GEN**: In the extended sibilant typology, the inputs and outputs of GEN replicate the structure of bSibInv.GEN (example (74)), except for the addition of voicing and fricative - affricate distinctions, as shown in (77).

(77) Input - output mappings

‘t’ = affricate, ‘s’ = fricative, ‘u’ = voiceless, ‘v’ = voiced

UI = {sAu, sAv, tAu, tAv,
sBu, sBv, tBu, tBv,
sCu, sCv, tCu, tCv,
sDu, sDv, tDu, tDv, n}

Input = any element of UI

Output = any element of UI

Candidate = x $\rightarrow$ y where x, y $\in$ UI

Every sibilant segment of bSibInv is thus articulated here into four variants: a voiceless fricative, a voiced fricative, a voiceless affricate and a voiced affricate. This brings the number of sibilants from 4 to 16, the number of segments of the universal inventory to 17 (i.e. the 16 sibilants + n) and the number of candidate mappings to 289, i.e. 17 x 17.

• **eSibInv.CON**: The constraint set of eSibInv is given in (78).

(78) eSibInv.CON

M.SIB= *A, B, C, D

M.PERIPH(SIB)= *A, D

M.POSTCD(SIB)= *C, D

M.POSTD(SIB)= *D

M.AFFR(SIB)= *t

M.VOICE(SIB)= *v

F.IDENT(SIB): Custom function (see (76) above).

Parallel to the addition of affricates and voicing distinctions with respect to the basic sibilant typology, eSibInv.CON adds two constraints against affricate sibilants and voiced sibilants: M.AFFR(SIB) corresponds to the principle of affricate markedness and M.VOICE(SIB) to voicing markedness. The constraint F.IDENT(SIB), in the case of a difference in manner or voicing (‘s’ vs. ‘t’, ‘u’ vs. ‘v’), assigns one violation per manner or voicing difference. This is because the letters ‘s’ and ‘t’ are 1 step distant from each other in the alphabet, as is the case for ‘u’ and ‘v’ (see the ASCII-converting function in 8.6.1). For example, the mapping from sAv to tCu incurs 1 violation for manner, 2 for PoA and 1 for voicing (i.e. 4 in total).

5.2.1.1 Summary

Table 7 provides an overview of how each principle is included in the model of this dissertation, and whether it is incorporated in the OT sibilant typologies or in the NN part.

1.	Number of PoAs and sibilant markedness	GEN: 4 candidates A, B, C, D + non-sibilant (.) CON: M.SIB against sibilants
2.	Centralisation	CON: M.PERIPH(SIB) against peripheral sibilants
3.	Posterior markedness	CON: M.POSTCD(SIB) against posterior sibilants, M.POSTD(SIB) against subapical sibilants
4.	Voicing markedness	GEN: 4 candidates x 2 = 8 (voiceless and voiced) CON: M.VOICE(SIB) against voiced sibilants
5.	Affricate markedness	GEN: 8 candidates x 2 = 16 (fricative and affricate) CON: M.AFFR(SIB) against sibilant affricates
6.	Fricative - affricate and voicing symmetry	Fricatives and affricates are derived from one grammar; manner and voicing gaps are possible, grammar-external and accidental (cfr. 5.4.2)
7.	Auditory dispersion	Emergent; modelled in Neural Networks
8.	Conservatism and contact	CON: F.IDENT(SIB) Language contact modelled in Neural Networks

Table 7: *Conversion of the principles from chapter 4 into OT candidates and constraints (in the case of teleological principles). See table 3 for a verbose phrasing of the principles.*

5.2.2 The SibInv OT typologies

What follows in this subsection is the gist of the OT typology in this dissertation, namely the basic sibilant typology (bSibInv). The reader is first introduced to the Violation Tableau (VT), which is the starting point of a thorough analysis of the typology and its workings. After this analysis, the extended typology (eSibInv) is presented, with a focus on its differences wrt. bSibInv. The software used for the evaluation of candidates and the creation of factorial typologies etc. is OTWorkplace (Prince et al. 2007-2021).

5.2.2.1 Violation Tableau

Violation Table(au) 8 reports the candidates and the constraints of the basic sibilant typology together with their respective violations. In the case of every markedness con-

Constraint definition:		*A, B, C, D	*A, D	*C, D	*D	[custom]
input	output	M.SIB	M.PERIPH(SIB)	M.POSTCD(SIB)	M.POSTD(SIB)	F.IDENT(SIB)
A	A	1	1	0	0	0
	B	1	0	0	0	1
	C	1	0	1	0	2
	D	1	1	1	1	3
	n	0	0	0	0	4
B	A	1	1	0	0	1
	B	1	0	0	0	0
	C	1	0	1	0	1
	D	1	1	1	1	2
	n	0	0	0	0	4
C	A	1	1	0	0	2
	B	1	0	0	0	1
	C	1	0	1	0	0
	D	1	1	1	1	1
	n	0	0	0	0	4
D	A	1	1	0	0	3
	B	1	0	0	0	2
	C	1	0	1	0	1
	D	1	1	1	1	0
	n	0	0	0	0	4
n	A	1	1	0	0	4
	B	1	0	0	0	4
	C	1	0	1	0	4
	D	1	1	1	1	4
	n	0	0	0	0	0

Table 8: *Violation Tableau of the basic sibilant typology. The percentage of grey is proportional to the number of violations.*

straint, the violation pattern is the same irrespectively of the input, but the faithfulness constraint F.IDENT(SIB) assigns violations depending on the input - output relation and therefore exhibits varying numbers of violations throughout the table(au).

5.2.2.2 The typology of bSibInv

After EVAL selected the optimal candidate pairs for the 5 inputs of bSibInv under all constraint rankings, the Factorial Typology yields 6 possible languages, displayed in table 9. For example, the first language maps A → A and all other sibilants to B, and can

Languages	**Inputs**				
	A	**B**	**C**	**D**	**n**
AB	A	B	B	B	n
ABC	A	B	C	C	n
ABCD	A	B	C	D	n
B	B	B	B	B	n
BC	B	B	C	C	n
n	n	n	n	n	n

Table 9: *Factorial Typology of the basic sibilant typology.*

be called 'the AB language' accordingly.

A first observation is that all languages except the sibilantless language have the mapping B → B. This is because it is the least marked sibilant: it is neither penalised by centralisation nor by posterior markedness. A second observation is that if a language allows for at least one sibilant, all sibilant inputs have a sibilant rather than a non-sibilant output. This is a consequence of the fact that F.IDENT(SIB) is formulated as a gradual constraint, assigning 4 violations to any mapping between sibilants and non-sibilants; it will thus always be preferable for a sibilant, if unfaithful, to be mapped to another sibilant rather than to a non-sibilant. In the version with categorical faithfulness (see 8.2.1.6 in the appendix), all faithfulness violations are equally assigned 1 violation and all unfaithful mappings are thus made towards the non-sibilant n, because it has 0 violation for all other constraints than F.IDENT(SIB).

A closer look at the 6 languages reveals a subdivision into three patterns:

- A **non-central** type, characterised by the presence of A (the first three languages in table 9)
- A **central** type, characterised by the absence of both A and D (the two subsequent languages)
- The **sibilantless** type (the last language)

The stringency hierarchy of posterior markedness (m.PostCD(Sib) and m.PostD(Sib)) is reflected in the non-central type: the AB language does not allow C and D, the ABC language allows C but not D, and the ABCD language allows both C and D. The hierarchy is also partly mirrored in the central type. Since central languages exclude the possibility of having D, we find only the reflection of the two first patterns: the B language which does not allow C and D, like the AB language; and the BC language which allows C but not D, like the ABC language.

Recalling the attested types of sibilant inventories in table 2 in chapter 3, the 6 languages of bSibInv look similar to the 7 attested sibinv types when reduced to 4 PoAs. An exception is that the well-dispersed 2s type (AC) is not predicted to exist, a point addressed in 5.2.2.6. Attested and bSibInv languages are listed in table 10. Importantly,

# PoAs	Example inventory	Reduced to 4 PoAs				Matching OT inventory
0	Hawaiian	_	_	_	_	n (sibilantless)
1	Athens (standard) Greek, Older Elfdalian	_	B	_	_	B (central)
2	Xhosa	A	_	C	_	none
	Faroese	_	B	C	_	BC (central)
	Évolène dialect, Japanese	A	B	_	_	AB (non-central)
3	Standard Basque (Euskara Batua), Polish	A	B	C	_	ABC (non-central)
4	Toda	A	B	C	D	ABCD (non-central)

Table 10: *Shortened summary of the attested sibilant inventories, not taking into account possible gaps and voicing distinctions (based on table 2).*

the 6 languages exhibit a behaviour of perceptual reanalysis of foreign sibilants that reflects the behaviour of attested sibilant inventories. As is visible in table 9, for example, the B language maps all sibilants to B, irrespectively of their place of articulation. This coincides with the observed loanword adaptation pattern in 1s inventories (see 3.2.2.2 with examples from Latin and Wolof). The same holds for e.g. A which is adapted as B in BC languages, and so forth. An advantage of the formulation of f.Ident(Sib) as a single gradual constraint is thus that it accurately mirrors the observed behaviour of natural languages, which is not the case of categorical faithfulness (see 8.2.1.6).

The domination hierarchies of constraints that yield each language are illustrated with Hasse diagrams in the following 3 figures (LXV - LXVII), with the non-central, central and sibilantless languages listed successively. Note that in a Hasse diagram, crucial domination is indicated by a line connecting a constraint to another, and freedom to be dominated or dominate by the absence of such a line. Domination is transitive, in the sense that if a constraint X dominates a constraint Y and Y dominates a constraint Z, X dominates Z.

All languages with sibilants, both non-central and central, have f.Ident(Sib) dom-

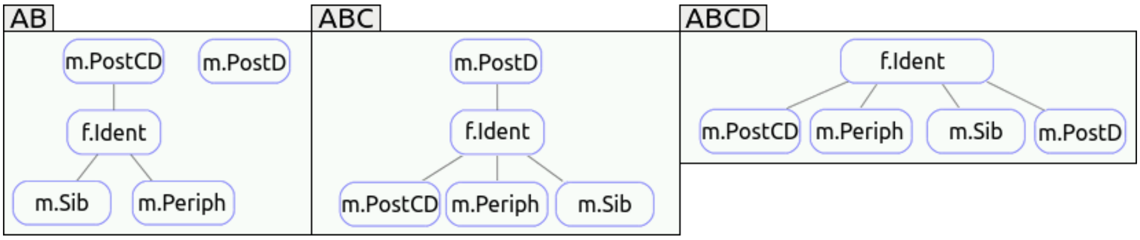


Figure LXV: *Hasse diagrams, generated in OTWorkplace (Prince et al. 2007-2021), showing the domination relations yielding the non-central languages (AB, ABC and ABCD).*

inating m.Sib, because otherwise the language necessarily becomes a sibilantless language. The Hasse diagram on the left of fig. LXV shows the domination relations yielding the AB language. In this language, m.PostCD(Sib) dominates f.Ident(Sib), something which prevents C and D from surfacing faithfully. Then, f.Ident(Sib) itself dominates m.Periph(Sib), something which ensures that A will surface faithfully. m.PostD(Sib) can be ranked anywhere, because since m.PostCD(Sib) dominates f.Ident(Sib), D will not be allowed to surface anyway.

The domination relations of the ABC language only differ from those of the AB language to the extent that in the ABC language, m.PostCD(Sib) is dominated by f.Ident(Sib), and m.PostD(Sib) must now be ranked above f.Ident(Sib) to prevent D from surfacing together with C. If m.PostD(Sib) is dominated by f.Ident(Sib) instead, one obtains the ABCD language. This language is totally faithful, in the sense that every input form surfaces without any modification.

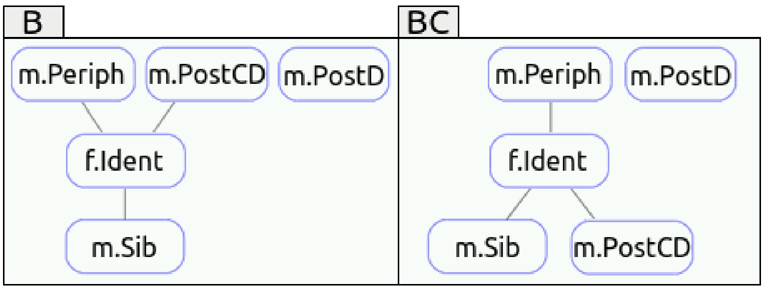


Figure LXVI: *Hasse diagrams, generated in OTWorkplace (Prince et al. 2007-2021), showing the domination relations yielding the central languages (B and BC).*

The central languages, in which m.Periph(Sib) dominates f.Ident(Sib), rule out A and D from the start. They are further distinguished by whether m.PostCD(Sib) also dominates f.Ident(Sib), which ensures that only B surfaces faithfully; if f.Ident(Sib) dominates m.PostCD(Sib), both B and C surface. m.PostD(Sib) can always be ranked anywhere, because D is eliminated by m.Periph(Sib) dominating f.Ident(Sib) from the start.

The sibilantless language is fairly straightforward: since m.Sib dominates f.Ident(Sib), no sibilant is allowed to surface faithfully and sibilants are mapped to non-sibilants instead. Centralisation and posterior markedness constraints can be ranked anywhere, because they penalise sibilants while there cannot be any sibilant.

It is worth observing that in all languages, markedness constraints interact indi-

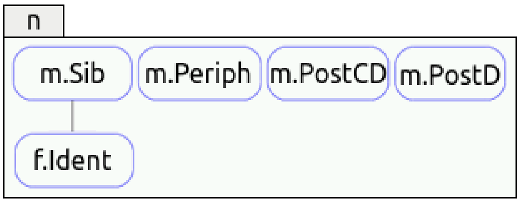


Figure LXVII: *Hasse diagram, generated in OTWorkplace (Prince et al. 2007-2021), showing the domination relation yielding the sibilantless language.*

vidually with faithfulness (F.IDENT(SIB)) and not among themselves. Indeed, the lines in the Hasse diagrams of LXV - LXVII always exclusively connect F.IDENT(SIB) with markedness constraints. The markedness constraints themselves do not interact: only the placement of F.IDENT(SIB) wrt. the markedness constraints will determine what the resulting language looks like. This typology is thus a typical example of the interaction of markedness and faithfulness constraints as two competing forces (Jarosz 2006: 53): markedness prefers certain less marked sibilants, and faithfulness counteracts this by preferring the phonetically closest sibilant to that of the underlying form.

To summarise what characterises the 6 languages predicted by the bSibInv typology, one can answer a series of questions as illustrated in table 11. This way of analysing the

Question	OT language					
	AB	ABC	ABCD	B	BC	n
Does M.SIB dominate F.IDENT(SIB)?	N	N	N	N	N	Y
Does M.PERIPH(SIB) dominate F.IDENT(SIB)?	N	N	N	Y	Y	n.a.
Does M.POSTCD(SIB) dominate F.IDENT(SIB)?	Y	N	N	Y	N	n.a.
Does M.POSTD(SIB) dominate F.IDENT(SIB)?	n.a.	Y	N	n.a.	n.a.	n.a.

Table 11: *Overview of what questions characterise the 6 languages of the OT typology.*

typology is typical of analysis in Property Theory (Alber and Prince 2016; Alber et al. 2016; McManus 2016; DelBusso 2018; Bennett and DelBusso 2018; Merchant and Krämer 2018), detailed in what follows.

5.2.2.3 Property Analysis

The reasons behind the structure of the 6 sibilant inventories described here are best explained as *properties*, a concept which is equivalent to the 'questions' asked in table 11. Properties are "ranking conditions that have mutually exclusive values" (Alber et al. 2016: e88): they compare one constraint or a group of constraints to another constraint or groups of constraints, asking which one dominates the other in each language or group of languages. Whereas a Factorial Typology only lists the predicted languages, a Property Analysis allows us to understand what fundamental traits distinguish the languages and thus to dig deeply into the structure of the typology. DelBusso (2018: ii) explains the relevance of Property Theory as follows:

> The list of languages generated in an OT factorial typology shows what the theory predicts, but not why it does so nor how it organizes the languages in the typological space. Property analysis answers these questions, finding the core structure that emerges directly from the logic of OT.

To each 'question' asked by a property, a language can 'answer' yes or no, but also 'n.a.', like in table 11. In Property Theory, this last possibility is called 'mootness' (meaning that both constraints can dominate one another without yielding another language). Table 12 lists the four properties of the Property Analysis of the basic sibilant typology, and table 13 the values for each property. These 4 properties are presented in the same order as in

Property name	**a <> b**	**Property description**
Sibilance	f.Ident(Sib) <> m.Sib Scope: all languages	Determines whether **sibilants** are mapped to *sibilants* (f.Ident(Sib)) or to *non-sibilants* (m.Sib).
Peripherality	f.Ident(Sib) <> m.Periph(Sib) Scope: languages with sibilants	... whether **non-central sibilants** are mapped *faithfully* (f.Ident(Sib)) or to *central sibilants* (m.Periph(Sib)).
PosteriorityCD	f.Ident(Sib) <> m.PostCD(Sib) Scope: languages with sibilants	... whether **posterior sibilants** are mapped *faithfully* (f.Ident(Sib)) or to *anterior sibilants* (m.PostCD(Sib)).
PosteriorityD	f.Ident(Sib) <> m.PostD(Sib) Scope: ABC and ABCD languages	... whether **subapical sibilants** are mapped *faithfully* (f.Ident(Sib)) or to *more anterior sibilants* (m.PostD(Sib)).

Table 12: *The four properties of the basic sibilant typology.*

PA table	Sibilance	Peripherality	PosteriorityCD	PosteriorityD
AB	f.Ident(Sib)	f.Ident(Sib)	m.PostCD(Sib)	moot
ABC	f.Ident(Sib)	f.Ident(Sib)	f.Ident(Sib)	m.PostD(Sib)
ABCD	f.Ident(Sib)	f.Ident(Sib)	f.Ident(Sib)	f.Ident(Sib)
B	f.Ident(Sib)	m.Periph(Sib)	m.PostCD(Sib)	moot
BC	f.Ident(Sib)	m.Periph(Sib)	f.Ident(Sib)	moot
n	m.Sib	moot	moot	moot

Table 13: *Property values of the basic sibilant typology.*

table 11, i.e. ordered by the width of their scope. Noteworthily, where there is mootness in table 13, there is a freely ranked constraint in the Hasse diagrams (LXV - LXVII). Given 4 properties, there are 16 logically possible combinations of property values, but only 6 languages: this is because 10 combinations yield a language that is the same as one of the 6 languages. For example, the AB language can have f.Ident(Sib) dominate or be dominated by m.PostD(Sib) for the property PosteriorityD, something which explains why there is one language for two possibilities.

An important merit of Property Theory is that it introduces the concept of scope: a property is relevant/distinctive for a certain number of languages, but not necessarily for all languages, where it can e.g. be moot. For example, PosteriorityD is only relevant to distinguish the ABC and the ABCD language. It thus has a rather *narrow* scope, in the sense that it only applies to a restricted number of languages (in this case, two). On the other hand, the property Sibilance has a *wide* scope, since it applies to all languages. Peripherality and PosteriorityCD have the same scope, and are both only relevant if the value f.Ident(Sib) has been chosen for the property Sibilance. This allows us to represent the Property Analysis in the form of a Property Treeoid (fig. LXVIII), which reflects the extent of the scope of single properties as vertical depth. The highest property

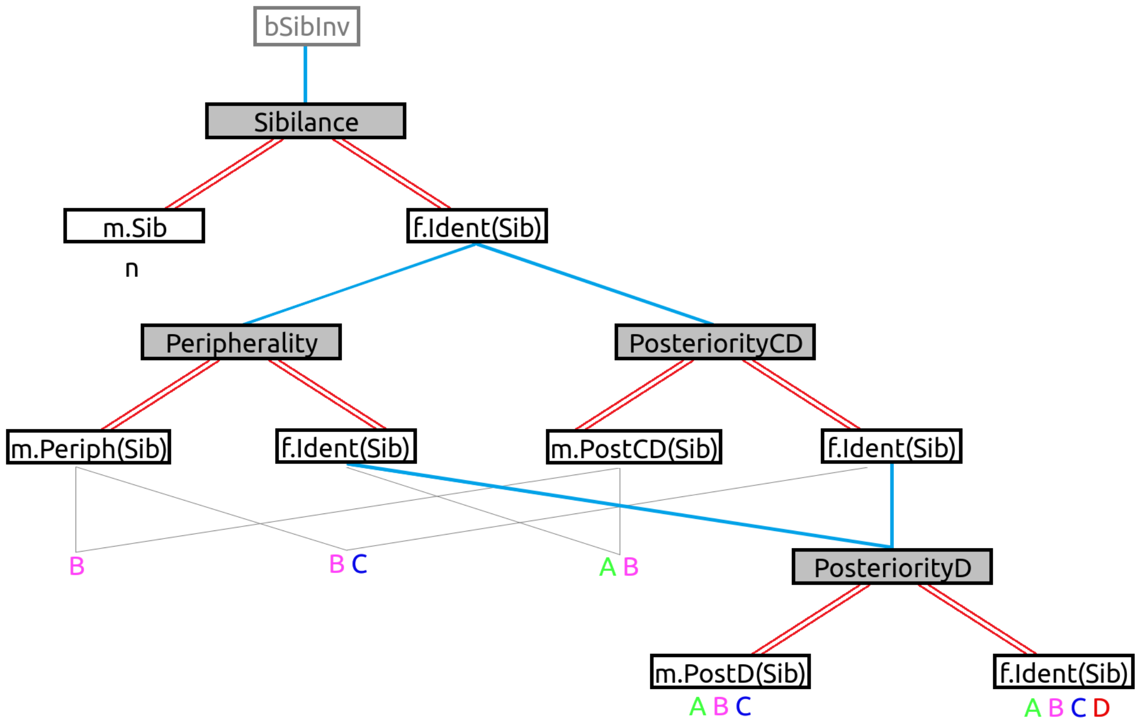


Figure LXVIII: *Property Treeoid of the basic sibilant typology. Note that the red double lines mean that a value for a property* must *be chosen.*

in this treeoid, Sibilance, is the one with the widest possible scope. The property with the narrowest scope, PosteriorityD, is expectably the lowest one. The treeoid shows straightforwardly that this property is only relevant if f.Ident(Sib) has been chosen for the three other properties. Indeed, the scopes reveal an important fact about the sibilant typology: when looking at two languages in the Factorial Typology, e.g. the sibilantless and the ABCD language, one is tempted to think superficially that they are simply distinguished by whether m.Sib dominates f.Ident(Sib). Instead, the much larger difference between them is that it suffices for m.Sib to dominate f.Ident(Sib) to obtain the sibilantless language, whereas *all* properties must choose f.Ident(Sib) to obtain ABCD.

5.2.2.4 Border points and typohedron

Each of the 6 languages of the basic SibInv typology consists of a certain number of possible constraint rankings that yield that language. For example, the perfectly faithful language ABCD requires that f.Ident(Sib) dominate all markedness constraints (cfr. LXV). However, the ranking of the markedness constraints among themselves does not change anything: whatever this ranking be, the resulting language will be ABCD. For example, a possible ranking yielding ABCD is:

(79) f.Ident(Sib) ≫ m.PostD(Sib) ≫ m.Periph(Sib) ≫ m.Sib ≫ m.PostCD(Sib)

Since there are 4 markedness constraints, there will be 24 possible constraint rankings yielding the ABCD language. Now, if one permutes two values involved in a property, one gets another language: permuting f.Ident(Sib) and m.PostD(Sib) in (79) will yield a language that does not allow D but maps all other sibilant faithfully, i.e. the ABC language. This means that ABCD and ABC are ‘neighbours’: they share at least one *border point* (Merchant and Prince 2017; Merchant and Krämer 2018), which is a permutation of two adjacent constraints that yields one or the other language, depending on the direction of the permutation. Not all languages share border points: for example, the B language requires that m.Periph(Sib) *and* m.PostCD(Sib) dominate f.Ident(Sib) (cfr. fig. LXVI), while the ABCD language requires that f.Ident(Sib) dominate all other constraints. It is thus impossible to ‘go’ from an ABCD language to a B language by permuting two adjacent constraints only *once*: one would need to permute f.Ident(Sib) with the higher ranked among m.Periph(Sib) and m.PostCD(Sib), and then to permute it again with the lower ranked among m.Periph(Sib) and m.PostCD(Sib). To ‘go’ from the ABCD language to the B language, one must thus either pass through a BC language (if f.Ident(Sib) is first permuted with m.Periph(Sib)) or through an AB language (if f.Ident(Sib) is first permuted with m.PostCD(Sib)). Border points play a key role in understanding language change, because they embody *minimal variation* from a language to another (see Alber 2001 et seq.).

When OTWorkplace (Prince et al. 2007-2021) evaluates the candidates under all possible constraint rankings, it notes which languages have at least one border point, i.e. which languages are ‘neighbours’. This information can then be visualised as a *typohedron*, which is the geometric representation of adjacent grammars. The typohedron, which was first introduced by Prince (2014) and Merchant and Prince (2017: 174-183), represents language-changing permutations between constraints on several dimensions. Figure LXIX illustrates the flattened (two-dimensional) typohedron of the basic sibilant typology, with the presence of border points as single lines (left) and with the number of border points as line thickness (right). In this 2D typohedron, the vertical axis coincides perfectly with the extent to which a language is faithful to the underlying forms (from

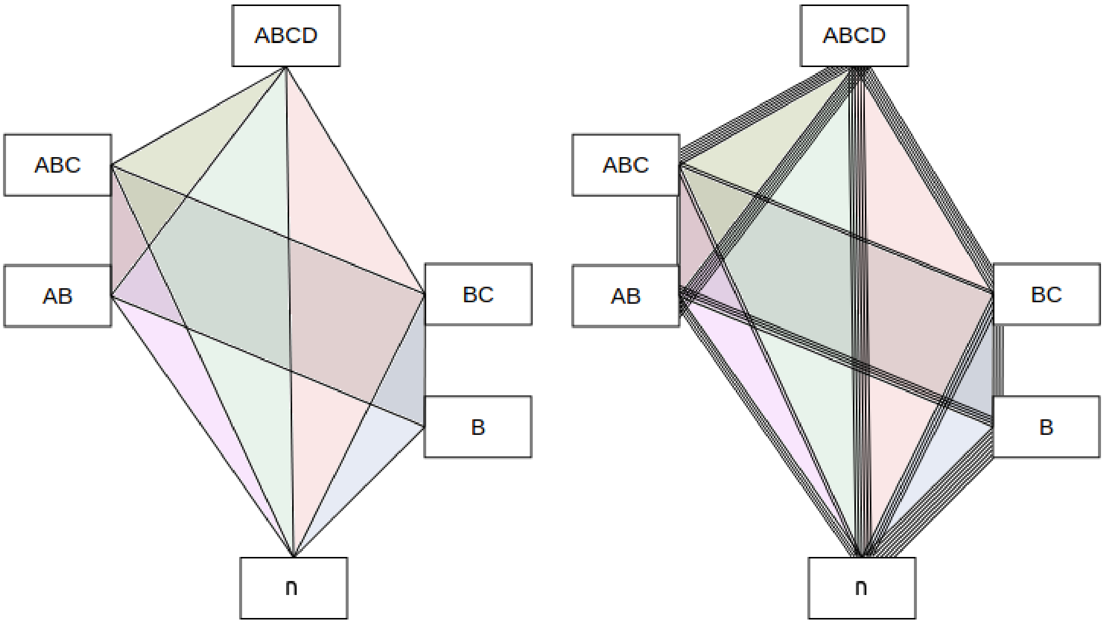


Figure LXIX: *Flattened 2D typohedron of the basic sibilant typology, without and with the number of border points as line thickness.*

perfectly faithful, above, to the totally unfaithful sibilantless language, below). The horizontal axis coincides with the non-central vs. central dichotomy (with central languages on the right), but imperfectly, because the sibilantless language is indifferent to centrality. It was placed in the middle to signal this indifference wrt. centrality, rather than being 'halfway' between non-central and central languages. Also, this horizontal axis possesses the imperfection that the ABCD language could not be put on the same axis as the other non-central languages, on the left, although it is also a non-central language. It was placed in the middle in order not to hide the line connecting it to the ABC language and the line connecting the ABC language to the AB language, i.e. to avoid opacity through superimposition of lines, but it should be located on the left as a non-central language.

A third relevant dimension is posterior markedness, expressed as the dichotomy 'with vs. without posterior sibilants'. Added to the faithfulness scale and the centrality dichotomy, the 6 languages of bSibInv can now be represented in 3D in fig. LXX. In this figure, the border points have been removed but the axes now correlate with the faithfulness, centrality and posterior markedness of the different languages. The coloured area linking languages to each other does not correspond with the border points, but correlates in color with faithfulness (dark = faithful, bright = unfaithful). The sibilantless language is placed halfway on the centrality dichotomy because it is unclassifiable in this respect, and is also placed halfway on the posterior markedness dichotomy for the same reason. Central languages are located towards the right/deep side, non-central languages towards the left/near side. Languages without posterior sibilants are situated above the others, and more faithful languages are situated more towards the left. On the faithfulness axis, the sibilantless language is located at 0, B at 1, AB and BC at 2, ABC at

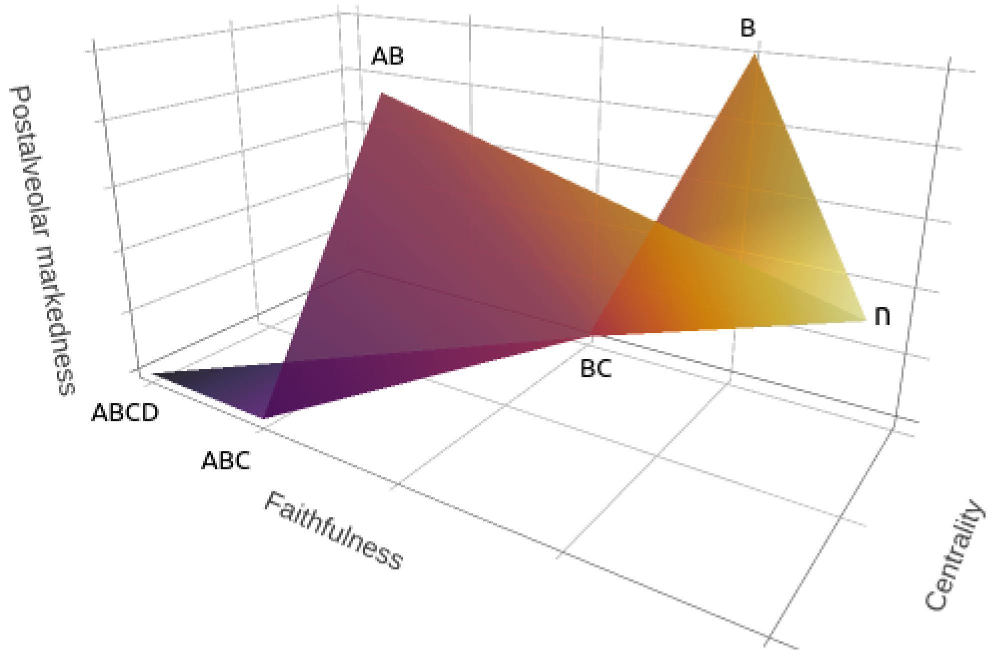


Figure LXX: *Three-dimensional representation of the basic sibilant typology.*

3 and ABCD at 4; this coincides with the number of contrasting PoAs for sibilants that they allow, and with the color coding.

The Property Analysis can be integrated into this 3D representation, as in fig. LXXI. The first property on the right, Sibilance, distinguishes the sibilantless language from

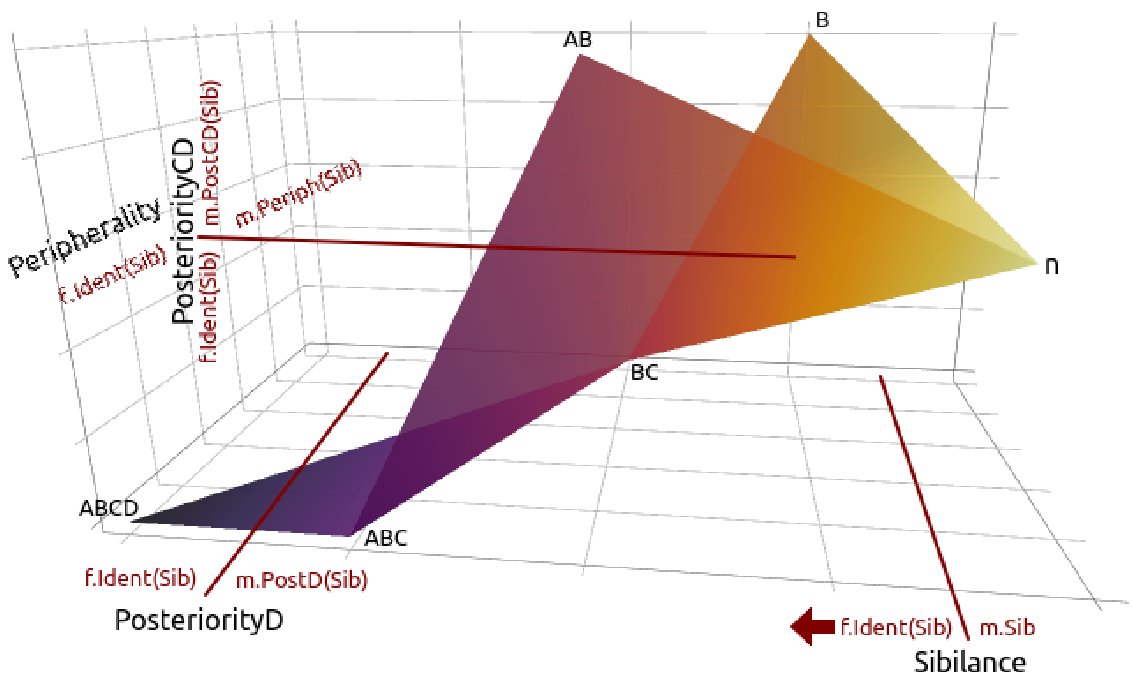


Figure LXXI: *Three-dimensional representation of bSibInv with the properties of the PA.*

the other languages. The two possible values, f.Ident(Sib) and m.Sib, are written on the relevant side of the line that delimits the zones in which one property value is chosen rather than the other. On the right side of this line, m.Sib dominates f.Ident(Sib) and the language is thus sibilantless. On the left side of this line, f.Ident(Sib) dominates m.Sib in *all* languages, a wide scope which is indicated with the red arrow. This arrow means that the languages in that direction all have f.Ident(Sib) above m.Sib, how far away they might be. On the contrary, the property PosteriorityD only distinguishes ABC and ABCD *locally*, and is moot for all other languages. This is indicated with the fact that there is no red arrow stretching from m.PostD(Sib) towards the right. In the middle of the figure, two properties distinguish central from non-central languages and languages with from languages without posterior sibilants. They are represented by one and the same line: higher than that line, the languages have m.PostCD(Sib) dominating f.Ident(Sib) and below that line, languages have posterior sibilants (property PosteriorityCD). Towards the left (i.e. closer to the viewer), languages are non-central while towards the right (i.e. further away from the viewer, deeper), languages have m.Periph(Sib) dominating f.Ident(Sib) (property Peripherality). This yields 4 possible values and is enough to distinguish the languages with sibilants unambiguously (except for the difference between ABC and ABCD): [central, no posterior sibilants] corresponds to the upper right/deep part of that zone, i.e. the B language; [central, posterior sibilants] to the lower right/deep part of that zone, i.e. the BC language; [non-central, no posterior sibilants] to the upper left/near part of that zone, i.e. the AB language; and [non-central, posterior sibilants] to the lower left/near part of that zone, occupied by both ABC and ABCD, which need to be further distinguished by PosteriorityD. Abstracting away from the sibilantless language, there is thus a square consisting of the two properties centrality and posterior markedness, with ABC and ABCD occupying together one corner of the square. The line for the Peripherality and PosteriorityCD properties does not go all the way to the right where the sibilantless language is, because it does not have any scope there; these properties are only relevant (i.e. non-moot) for all languages with sibilants.

The two preceding figures, fig. LXX and LXXI, are not properly speaking typohedra, because they do not include the border points (although they do have the axes faithfully representing property values). Figure LXXII illustrates the typohedron. The typohedron allows us to represent the basic OT typology in its full detail. Moreover, given the assumption that a natural language can only follow border points (i.e. swap crucially ranked constraints minimally) to change from a language to another in minimal steps (Alber 2001, 2014; Alber and Meneguzzo 2016), this typohedron makes explicit how the different sibilant inventories are linked to each other and can evolve. For example, an AB language could not become a BC language without becoming another language first.

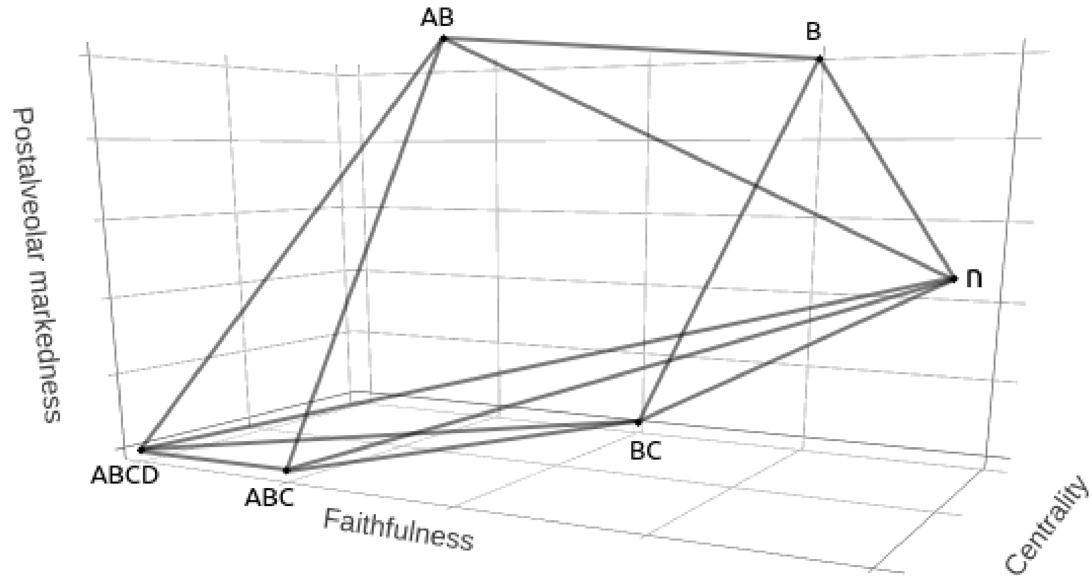


Figure LXXII: *Three-dimensional typohedron of the basic sibilant typology.*

5.2.2.5 The typology of eSibInv and its differences wrt. bSibInv

Recall that eSibInv.GEN and eSibInv.CON are the following:

(80) Input - output mappings

't' = affricate, 's' = fricative, 'u' = voiceless, 'v' = voiced

UI = {sAu, sAv, tAu, tAv,
sBu, sBv, tBu, tBv,
sCu, sCv, tCu, tCv,
sDu, sDv, tDu, tDv, n}

Input = any element of UI

Output = any element of UI

Candidate = x → y where x, y ∈ UI

(81) eSibInv.CON

M.SIB= *A, B, C, D

M.PERIPH(SIB)= *A, D

M.POSTCD(SIB)= *C, D

M.POSTD(SIB)= *D

M.AFFR(SIB)= *t

M.VOICE(SIB)= *v

F.IDENT(SIB): Custom function (see (76) above).

The extended typology adds sibilant affricates and voicing distinctions to the basic typology, together with the constraints M.AFFR(SIB) against sibilant affricates and M.VOICE(SIB) against voiced sibilants. These two constraints are categorical and thus issue at most 1 violation (either a sibilant is an affricate or not; either it is voiced, or not). After positing the possible accidental gaps for each inventory resulting from the extended typology (in 5.4.2), the predicted inventories can finally be compared to the attested inventories in the SibInv database (in 6.1).

The presence of voicing and fricative - affricate distinctions multiplies by two and two again the number of candidates: the 25 candidates (5 outputs for all 5 inputs, i.e. 4 sibilants + 1 non-sibilant) of the basic typology become 289 candidates (17 outputs for all 17 outputs, i.e. 4 x 2 x 2 sibilants + 1 non-sibilant) in the extended typology. For reasons of space, only the first 20 input - output pairs of eSibInv are represented in table 14 with their respective violations. One shall notice that F.IDENT(SIB) assigns the same number

		*A, B, C, D	*A, D	*C, D	*D	[custom]	*v	*t
input	output	M.SIB	M.PERIPH(SIB)	M.POSTCD(SIB)	M.POSTD(SIB)	F.IDENT(SIB)	M.VOICE(SIB)	M.AFFR(SIB)
sAu	sAu	1	1	0	0	0	0	0
	sBu	1	0	0	0	1	0	0
	sCu	1	0	1	0	2	0	0
	sDu	1	1	1	1	3	0	0
	sAv	1	1	0	0	1	1	0
	sBv	1	0	0	0	2	1	0
	sCv	1	0	1	0	3	1	0
	sDv	1	1	1	1	4	1	0
	tAu	1	1	0	0	1	0	1
	tBu	1	0	0	0	2	0	1
	tCu	1	0	1	0	3	0	1
	tDu	1	1	1	1	4	0	1
	tAv	1	1	0	0	2	1	1
	tBv	1	0	0	0	3	1	1
	tCv	1	0	1	0	4	1	1
	tDv	1	1	1	1	5	1	1
	n	0	0	0	0	6	0	0
sBu	sAu	1	1	0	0	1	0	0
	sBu	1	0	0	0	0	0	0
	sCu	1	0	1	0	1	0	0
etc.	etc.							

Table 14: *Violation Tableau of the extended sibilant typology. The percentage of grey is proportional to the number of violations. The other 269 of the 289 input - output mappings are absent for reasons of space, but are of course part of the extended typology.*

of faithfulness violations as in the basic typology, with the addition of 1 violation for a mapping from voiceless to voiced or vice-versa, and 1 violation for a mapping from fricative to affricate or vice-versa.

The resulting Factorial Typology contains 21 languages, which are shown in fig. LXXIII. To highlight what the resulting pattern is, perfectly faithful mappings have been indicated with a colouring corresponding to the place of articulation, showing us which contrasting sibilants each language has. The spacing in between sets of four rows in fig. LXXIII highlights a familiar pattern: in fact, the factorial typology of eSibInv is precisely like that of bSibInv, but with four variants for each language (except the sibilantless one). Each of the 5 languages with sibilants of the basic typology comes in 4 versions: one without voicing and without affricates, one without voicing but with affricates, one with voicing but without affricates, and one with voicing and affricates. Adding the

Inputs->	sAu	sBu	sCu	sDu	sAv	sBv	sCv	sDv	tAu	tBu	tCu	tDu	tAv	tBv	tCv	tDv
L.01	sAu	sBu	sBu	sBu	sAu	sBu	sBu	sBu	sAu	sBu	sBu	sBu	sAu	sBu	sBu	sBu
L.02	sAu	sBu	sBu	sBu	sAu	sBu	sBu	sBu	tAu	tBu	tBu	tBu	tAu	tBu	tBu	tBu
L.03	sAu	sBu	sBu	sBu	sAv	sBv	sBv	sBv	sAu	sBu	sBu	sBu	sAv	sBv	sBv	sBv
L.04	sAu	sBu	sBu	sBu	sAv	sBv	sBv	sBv	tAu	tBu	tBu	tBu	tAv	tBv	tBv	tBv
L.05	sAu	sBu	sCu	sCu	sAu	sBu	sCu	sCu	sAu	sBu	sCu	sCu	sAu	sBu	sCu	sCu
L.06	sAu	sBu	sCu	sCu	sAu	sBu	sCu	sCu	tAu	tBu	tCu	tCu	tAu	tBu	tCu	tCu
L.07	sAu	sBu	sCu	sCu	sAv	sBv	sCv	sCv	sAu	sBu	sCu	sCu	sAv	sBv	sCv	sCv
L.08	sAu	sBu	sCu	sCu	sAv	sBv	sCv	sCv	tAu	tBu	tCu	tCu	tAv	tBv	tCv	tCv
L.09	sAu	sBu	sCu	sDu	sAu	sBu	sCu	sDu	sAu	sBu	sCu	sDu	sAu	sBu	sCu	sDu
L.10	sAu	sBu	sCu	sDu	sAu	sBu	sCu	sDu	tAu	tBu	tCu	tDu	tAu	tBu	tCu	tDu
L.11	sAu	sBu	sCu	sDu	sAv	sBv	sCv	sDv	sAu	sBu	sCu	sDu	sAv	sBv	sCv	sDv
L.12	sAu	sBu	sCu	sDu	sAv	sBv	sCv	sDv	tAu	tBu	tCu	tDu	tAv	tBv	tCv	tDv
L.13	sBu	sBu	sBu	sBu	sBu	sBu	sBu	sBu	sBu	sBu	sBu	sBu	sBu	sBu	sBu	sBu
L.14	sBu	sBu	sBu	sBu	sBu	sBu	sBu	sBu	tBu	tBu	tBu	tBu	tBu	tBu	tBu	tBu
L.15	sBu	sBu	sBu	sBu	sBv	sBv	sBv	sBv	sBu	sBu	sBu	sBu	sBv	sBv	sBv	sBv
L.16	sBu	sBu	sBu	sBu	sBv	sBv	sBv	sBv	tBu	tBu	tBu	tBu	tBv	tBv	tBv	tBv
L.17	sBu	sBu	sCu	sCu	sBu	sBu	sCu	sCu	sBu	sBu	sCu	sCu	sBu	sBu	sCu	sCu
L.18	sBu	sBu	sCu	sCu	sBu	sBu	sCu	sCu	tBu	tBu	tCu	tCu	tBu	tBu	tCu	tCu
L.19	sBu	sBu	sCu	sCu	sBv	sBv	sCv	sCv	sBu	sBu	sCu	sCu	sBv	sBv	sCv	sCv
L.20	sBu	sBu	sCu	sCu	sBv	sBv	sCv	sCv	tBu	tBu	tCu	tCu	tBv	tBv	tCv	tCv
L.21	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n

Figure LXXIII: *Factorial Typology of the extended sibilant typology (generated in OTWorkplace; Prince et al. 2007-2021).*

voicing and fricative vs. affricate distinctions to the basic typology thus simply adds 3 new variants for every language with sibilants, without 'telling us much news'. The Property Analysis of the extended typology confirms that there is no structural change in the behaviour of the typology other than obtaining 4 variants of each language with sibilants, as can be seen in the Property Analysis of fig. LXXIV. The two last properties on the right of the PA table exhibit the recurrent pattern 'bb-ba-ab-aa' all the way down. On a Property Treeoid, the two new properties would have the same scope as PERIPHERALITY and POSTERIORITYCD, i.e. all languages that have sibilants, as can be seen in the PA of fig. LXXIV. They are also totally independent from the other properties, determining whether affricates or voiced sibilants are allowed to surface in any language with sibilants. On a typohedron, there would be four dots at every language with sibilants which would be connected by border points with two of the three other variants at that language (e.g. the AB with voicing and affricates is connected to the AB without voicing with affricates and to the AB with voicing without affricates, but not to the AB without voicing and without affricates, because this requires two permutations of crucially

CON
f.Ident(Sib)
m.Sib
m.Periph(Sib)
m.PostCD(Sib)
m.PostD(Sib)
m.Voice(Sib)
m.Affr(Sib)
{}.dom
{}.sub

Props	a	b
Sibilance	f.Ident(Sib)	m.Sib
Peripherality	f.Ident(Sib)	m.Periph(Sib)
PosteriorityCD	f.Ident(Sib)	m.PostCD(Sib)
PosteriorityD	f.Ident(Sib)	m.PostD(Sib)
VoicedSibilants	f.Ident(Sib)	m.Voice(Sib)
AffricateSibilants	f.Ident(Sib)	m.Affr(Sib)

PA table	Sibilance	Peripher.	Poster.CD	Poster.D	VoicedS.	Affr.S.	ok?
L.01	a	a	b	*moot*	b	b	::
L.02	a	a	b	*moot*	b	a	::
L.03	a	a	b	*moot*	a	b	::
L.04	a	a	b	*moot*	a	a	::
L.05	a	a	a	b	b	b	::
L.06	a	a	a	b	b	a	::
L.07	a	a	a	b	a	b	::
L.08	a	a	a	b	a	a	::
L.09	a	a	a	a	b	b	::
L.10	a	a	a	a	b	a	::
L.11	a	a	a	a	a	b	::
L.12	a	a	a	a	a	a	::
L.13	a	b	b	*moot*	b	b	::
L.14	a	b	b	*moot*	b	a	::
L.15	a	b	b	*moot*	a	b	::
L.16	a	b	b	*moot*	a	a	::
L.17	a	b	a	*moot*	b	b	::
L.18	a	b	a	*moot*	b	a	::
L.19	a	b	a	*moot*	a	b	::
L.20	a	b	a	*moot*	a	a	::
L.21	b	*moot*	*moot*	*moot*	*moot*	*moot*	::

Figure LXXIV: *Property Analysis of the extended sibilant typology (generated in OTWorkplace; Prince et al. 2007-2021).*

ranked constraints), and which would be connected to their respective correspondent at other languages in the same way as in the typohedron of the basic typology.

To these 21 inventories, the well-dispersed 2s variants and every possible accidental gap will be added in 5.4.2, so that the resulting sibilant inventories will be comparable with the inventories in the SibInv database.

5.2.2.6 The absent well-dispersed 2s and its implications for SibInv

Since auditory dispersion was not included as a constraint in the SibInv typologies, no well-dispersed language was predicted to exist. By *well-dispersed*, it is meant 'having precisely one empty PoA slot between two occupied PoAs', i.e. A - C or B - D with 4 PoAs. *Overdispersed* inventories have two or more empty PoA slots in between, which could only be A - D in the case of 4 PoAs.

As 5.4 will show, the AC language, despite not being predicted by the SibInv typologies, can emerge diachronically from an AB language (by retracting B to C), from a BC language (by fronting B to A) or from an ABC language (by letting B merge with A and/or C). Crucially, other well-dispersed or overdispersed inventories (AD, BD, CD, ABD, ACD, BCD) *do not* emerge in the NN simulations and, what is more, simulations in 5.4.1.6 show precisely that if they are artificially given as input to generations of learners, they spontaneously shift over time to inventories predicted by the SibInv typologies or to the AC language. This is because there always exists a sibilant inventory (either predicted by SibInv or AC) that is 'better' than these unstable inventories:

- In an AD language, auditory dispersion cannot operate because the categories are too far apart for their phonetic ranges to overlap. An overlap in phonetic ranges is crucial to the development of auditory dispersion (see 5.3.1). Articulatory constraints (both centralisation and posterior markedness) therefore shift D forward to C because it is less or equally marked with respect to all markedness constraints, and the language becomes an AC language.

- A BD language is equivalent to AC in terms of dispersion, but more marked than AC (5 vs. 4 violations for the two PoAs, cfr. table 8). It therefore shifts to AC.
- A CD language is inferior to AC in terms of dispersion and more marked than AC (6 vs. 4 violations for the two PoAs). It therefore shifts to AC.
- In the ABC language, distributedness distinctions enhance the auditory contrasts between the three sibilant PoAs. The ABD and ACD languages only add PoA dispersion to half the sibilants at the expense of having D, which is more marked than B and C (3 and 2 violations more, respectively). The BCD language is equivalent in terms of dispersion to ABC, but more marked than ABC (7 vs. 5 for the three PoAs). They therefore shift to ABC.

In OT terms, these unstable inventories are 'harmonically bounded' wrt. the predicted ones. The AC language also is wrt. AB and BC in bSibInv, but differs from the unstable inventories in being an enhancement of auditory dispersion occupying two PoA slots of an ABC language. One could object that ABD and ACD are also enhancements of auditory dispersion occupying three PoA slots of an ABCD language, but there is a difference wrt. the AC language: ABD and ACD *already allow* for two adjacent sibilant PoAs (A-B and C-D, respectively), and could therefore very well allow for two adjacent sibilant PoAs by being ABC and would thereby avoid the more marked segment D. This is not the case for AC, which does not allow for two adjacent sibilant PoAs.

The AC language is thus not predicted by the SibInv typologies because they do not include auditory dispersion constraints, but it nevertheless emerges diachronically and remains stable thanks to its auditory dispersion advantage. It remains to be explained what happens in OT terms when a child is exposed to native language input from an AC language, i.e. how this language is to be represented as an OT grammar. In what follows, I address two possible ways, one with BiPhon-OT (Bidirectional Phonology and Phonetics in OT) and one with an ungrounded constraint.

5.2.2.6.1 BiPhon-OT

In BiPhon-OT (Boersma 1998, 2011*b*), perception is learnt earlier than production, in the period during which children perceive but do not (or barely do) produce linguistic output (Boersma and Hamann 2006, 2008). This is called prelexical perception in Boersma and Hamann (2008). A child exposed to (almost exclusively) A and C as linguistic input ranks her/his cue constraints in such a way that sibilants are mapped to one of these two categories in perception (compare e.g. the language-particular cue constraints *[20.2 Erb]/s/ or *[39.6 Erb]/ʃ/ in Boersma and Hamann 2008: 233). Should a sibilant be heard that phonetically corresponds to B, it will fall in the domain of acoustic overlap of the categories A and C and be reanalysed in perception as belonging to one of these two surface forms (based on its acoustic proximity to the perceptual prototype of A or C, as embodied in the cue constraint ranking). Richness of the Base (i.e. the assumption that there are no language-particular limitations on possible inputs), which in BiPhon-OT does not apply to the underlying form but to the auditory form (Boersma and Hamann 2006: 11), will therefore not provoke reanalysis to a new category B for the few phonetically B tokens that might be perceived by the learning child. When the grammar learnt through perception starts to be used in production by the child, m.PostD(Sib) is ranked higher than f.Ident(Sib) because of the absence of D in the native input, f.Ident(Sib) higher than m.PostCD(Sib) because of the presence of C and f.Ident(Sib) higher than m.Periph(Sib) because of the presence of A in the native input. This in fact corresponds precisely to the domination hierarchy of the ABC language in bSibInv, as shown in the Hasse diagram of fig. LXXV. No lexically stored form

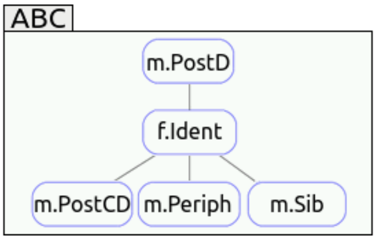


Figure LXXV: *Hasse diagram of the ABC language, generated in OTWorkplace (Prince et al. 2007-2021).*

in the grammar of the child will thus contain B as a sibilant, and the child will neither perceive nor produce B. For an adult speaker of an AC language to start perceiving and producing forms with B, it will first be required to change one's cue constraints to include constraints like *[28.2 Erb]/s̠/, i.e. cue constraints that acknowledge the existence of another category than /s/ and /ʃ/. As this will add a new |s̠| category in the lexicon if successful, the speaker will then have to change her/his sensorimotor constraints to be able to realise [[s̠]] as [s̠]. Therefore, from the perspective of BiPhon-OT, the 'AC problem' is unproblematic to the extent that /B/ cannot arise in the grammar of a child who has been exposed to [A] and [C] as input.

5.2.2.6.2 An ungrounded constraint

If Richness of the Base however is considered to apply to underlying forms, there *must* be lexical forms in an AC language that have an underlying B; if not in native vocabulary, at least words of foreign origin with an original B will have input - output candidates with an underlying B. Nothing then impedes B to surface in the AC language, quite the opposite: it surfaces as the optimal candidate regardless of the constraint ranking, as long as sibilants are allowed to surface (i.e. as long as f.Ident(Sib) dominates m.Sib). The unavoidable mapping of B to B in an AC language is illustrated in table(au) 15. This problem can be solved with an ungrounded constraint against the B sibilant PoA (Hayes

/B/	*D	[custom]	*A, D	*C, D	*A, B, C, D
output	m.PostD(Sib)	f.Ident(Sib)	m.Periph(Sib)	m.PostCD(Sib)	m.Sib
A	0	1!	1	0	1
☞ B	0	0	0	0	1
C	0	1!	0	1	1
D	1!	2	1	1	1

Table 15: *Undesired election of B as winning candidate in an AC production grammar.*

1999: 270f.). According to Hayes, during language acquisition, the child tries to induce from the surrounding linguistic input what are the existing constraints restricting the surfacing output forms, and what is their ranking. This is called *inductive grounding* by Hayes. A large majority of these induced constraints is universal, because the human speech organs do not significantly differ from one population to another, so that a constraint such as voicing markedness for sibilants (m.Voice(Sib)) will exist in every human being's phonology, because voiced sibilants are harder and more effortful to produce for all human beings. These constraints are *grounded*, because they are based on actual articulatory experience. On the other hand, ungrounded constraints are also induced from the surrounding input but are not based on articulatory motivations; in fact, they are not motivated at all. The only reason they exist is that there is a gap in the input distribution, so that the child induces the existence of the constraint by the fact that what it bans (almost) never occurs in the linguistic input. In Hayes (1999: 270)' terms:

> [T]he source [JK: of ungrounded constraints] is induction, in this case not over the learner's phonetic experience but over the input data: eventually, the child figures out such constraints from negative evidence; that is, from systematic, consistent, long-term absence of a particular structure in the input data. (…) [I]nductively grounded constraints [JK: are] learned from internal experience[,] and learned [i.e. ungrounded] constraints [are] learned from gaps in input data.

A similar point is made by Vaux and Samuels (2006: 21):

> There is ample evidence that humans form generalizations over the data to which they are exposed; so for example a child exposed to outputs of a grammar containing no e's may well postulate a systematic gap via rule or constraint in their synchronic grammar that produces this distribution (though they also may not, resulting in an accidental gap).

This difference between systemic and accidental gaps will be of importance in 5.4.2, where it is shown how accidental gaps appear. Since gaps in the distribution are language-specific, ungrounded constraints are language-specific. The internal experience of voicing markedness is equivalent for all human beings since all have similar speech organs, but a gap is by definition a lacking segment at a place where it would have been expected and is possible; there will thus be languages that fill that gap and thus that do not have this ungrounded constraint.

A well-dispersed AC language can be represented as an ABC language as in bSibInv, with the addition of an ungrounded constraint *B(UNGROUNDED). Such a constraint cannot be present in bSibInv because this would imply that it is universal. Instead, when auditory dispersion transforms an AB or BC language into an AC language, the child who is exposed to the new input distribution ranks the relevant constraints with *B(UNGROUNDED) and M.POSTD(SIB) dominating F.IDENT(SIB), which itself dominates the other sibilant-related constraints of the typology. This results in the desired AC pattern, as shown in table(au) 16 and the Hasse diagrams in fig. LXXVI.

/B/	*B	*D	[custom]	*A, D	*C, D	*A, B, C, D
output	*B(UNGR.)	M.POSTD(SIB)	F.IDENT(SIB)	M.PERIPH(SIB)	M.POSTCD(SIB)	M.SIB
A	0	0	1	1	0	1
B	1!	0	0	0	0	1
☞ C	0	0	1	0	1	1
D	0	1!	2	1	1	1

Table 16: *Desired election of C as a winning candidate in an AC production grammar.*

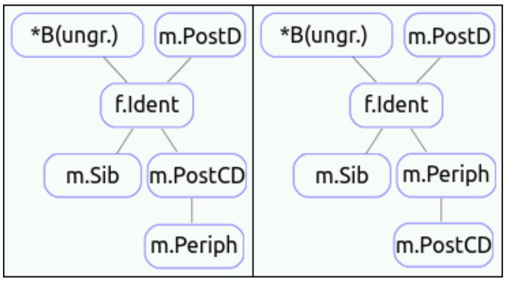


Figure LXXVI: *Hasse diagram of the AC language, generated in OTWorkplace (Prince et al. 2007-2021).*

Note that the ranking of centralisation (M.PERIPH(SIB)) with respect to posterior markedness (M.POSTCD(SIB)), which was indifferent in an ABC language, now becomes crucial: depending on whether one or the other dominates the other one, a B in loanwords is reanalysed as A or C. This means that if one swaps M.PERIPH(SIB) and M.POSTCD(SIB) in table 16, a foreign B will be reanalysed to A rather than to C. The constraint *B(UNGROUNDED) cannot be a universal constraint, because if it is included in bSibInv, it yields four additional languages out of which three are unattested. These languages are: ACD, AC, A and C. Each of these languages has the same grammar as the ABCD, ABC, AB and BC languages respectively, but with *B dominating F.IDENT(SIB) each time. All languages with sibilants thus get a correspondent without B, except the B language (because its result is indistinct from the sibilantless language).

5.2.3 Production and perception

According to the BiPhon model, "a language user uses the same constraints when she speaks as when she listens, with the same rankings" (Boersma 2011*b*: 34). This means that the constraint ranking of e.g. a BC language will not be different in production or perception. The production grammar is the same as the perception grammar (abstracting away from the articulatory constraints used only when speaking). Importantly, this does not mean that speakers explicitly and voluntarily adapt their speech for listeners and the other way around: such effects arise non-teleologically and outside of grammar (Boersma 2011*b*). An example illustrating the perception grammar of a BC language is given in table(au) 17. Having created and ranked these cue constraints (among many

[20.9 ERB]	*[20.9 ERB]/B/	*[20.8 ERB]/B/	*[20.8 ERB]/C/	*[20.9 ERB]/C/
/A/				
/B/	1!			
☜ /C/				1
/D/				

Table 17: *Violation Tableau illustrating the perception of a [[D]] of 20.9 ERB in a BC language. Note that in the case of perception, the winning candidate is indicated with a leftward-pointing hand (Boersma 2011*b*: 36).*

other cue constraints) during childhood with the help of the Gradual Learning Algorithm (Boersma and Hamann 2008: 235-238), the perception mechanism has come to identify this [[D]] as the phonological surface form /C/, because *[20.9 ERB]/B/ is ranked above *[20.9 ERB]/C/. Table(au) 17 also includes the two *[20.8 ERB] constraints to illustrate their ranking, even if they play no role in evaluation, as much as the candidates /A/ and /D/ that have no associated cue constraints. Since sibilants with lower ERB values at the lower end of the continuum (i.e. with more extreme ERB values) are less likely to have occurred in the native language input, the constraints that penalise them have been reranked more rarely, leading to a cue constraint ranking in which the *[20.9 ERB] constraints have risen or descended more than them (Boersma and Hamann 2008: 238).

The way in which cue constraints filter the perceived input towards the acquired SF categories and these SF categories are realised as auditory and articulatory forms in production is rather straightforward. One peculiar typologically attested case, however, deserves to be addressed here: the case of 0s languages with sibilant allophones (detailed in 3.2.2.1). As the example of Kiribati illustrates, underlying /ti/ surfaces as a sibilant by means of stop palatalisation despite the absence of phonemic sibilants; the grammar is posited to look as shown in table(aux) 18 and 19. In this first stage, the palatalisation constraint Pal([ti]), which disfavours the tongue transition gestures from the low [t] to the high [i], dominates Uniformity (which prohibits the coalescence of two segments into one) and *Pal([c]) (representing the articulatory markedness of the palatal segment [c]). Additionally, Dep, Max and Ident(Vowel/Consonant) ban alternative repair strategies, namely epenthesis, deletion or changing the nature of the vow-

/ti/	*ti	[custom]	[custom]	[custom]	[custom]	*[c]
output	Pal([ti])	Max	Dep	Ident(V/C)	Uniformity	*Pal([c])
ti	1!	0	0	0	0	0
t	0	1!	0	0	0	0
i	0	1!	0	0	0	0
təi	0	0	1!	0	0	0
te	0	0	0	1!	0	0
pi	0	0	0	1!	0	0
☞ c	0	0	0	0	1	1

Table 18: *Production Violation Tableau illustrating the palatalisation of /ti/ in Kiribati, a 0s with allophonic sibilants. The intermediary step with [ci] has been skipped.*

el/consonant, respectively. If all these constraints dominate Uniformity and *Pal([c]), a speaker will produce /ti/ as [c]. However, spontaneous affrication occurs when producing a palatal stop (Jacobs and Berns 2013; Hall et al. 2006; Recasens 2013: 17) so that it will frequently be realised as a sibilant ([ɕ], i.e. B). As soon as the frequency of affricated versions of palatal stops in the surrounding input of children is large enough, they will categorise them as underlying sibilants and m.Sib (which is ranked highest from the start in acquisition, see D.) will get ranked lower than f.Ident(Sib) as a result of reranking in the Gradual Learning Algorithm (see 5.4.1). A language with palatal stops and without

[c]	*[c]	*A, B, C, D	[custom]	*C, D	*A, D	*D
output	*Pal([c])	m.Sib	f.Ident(Sib)	m.PostCD(Sib)	m.Periph(Sib)	m.PostD(Sib)
A	0	1	1	0	1!	0
☜ B	0	1	0	0	0	0
C	0	1	1	1!	0	0
D	0	1	2	1!	1	1
c	1!	0	4	0	0	0

Table 19: *Perception Violation Tableau illustrating the perceptual reanalysis of palatalised /ti/ [c] in Kiribati, a 0s with allophonic sibilants.*

sibilant phonemes, like the majority of Australian languages, will thus have m.Sib dominating f.Ident(Sib) because markedness constraints are ranked higher than faithfulness in child speech and there are not enough affricated palatal stops in the native input to reverse this ranking. When this happens, the language maps them to B because it is the least marked sibilant.

One might object that it is unrealistic that the grammar uses structural constraints that are originally purely articulatorily motivated (like e.g. m.Periph(Sib)) in perception, since articulatory markedness is not directly involved in perception. Now, even if one takes this objection into account, the resulting pattern is not any different, as I will show here. Removing all articulatorily motivated structural/markedness constraints, only faithfulness remains; this means that a D in an AB language would be perceived as a D rather than as a B. This is far from impossible: while it can be claimed that a native speaker of an AB language really *believes* that the underlying category of an incoming C *is* a B, since their acoustic vicinity make them easily confusable for an untrained listener, it is harder to claim that something as radically posterior as a D is really heard as a 'pure' B by a native speaker of an AB language: B is the phonetically closest category, but still distant enough so that the hearer might recognise that the acoustic token [[D]] is far from a canonical B-pronunciation and could constitute a category on its own. In this case, a D would thus be heard as a D, as much as any sibilant or non-sibilant would be mapped faithfully. Nevertheless, when the native speaker of an AB language tries to repeat the word heard, the result automatically becomes B, as shown in table(au) 20.

	|D|	*D	*C, D	[custom]	*A, D	*A, B, C, D
	output	M.POSTD(SIB)	M.POSTCD(SIB)	F.IDENT(SIB)	M.PERIPH(SIB)	M.SIB
	/A/	0	0	3!	1	1
☞	/B/	0	0	2	0	1
	/C/	0	1!	1	0	1
☜	/D/	1!	1	0	1	1
	/n/	0	0	4!	0	0

Table 20: *VT of the production and perception of D in an AB language. The leftward-pointing arrow indicates that the speaker has previously heard [[D]] and categorised it as |D|, under the assumption made above that only F.IDENT(SIB) is active in perceiving an SF.*

Table(au) 20 reveals that even if the native speaker of an AB language has heard [D] and categorised it as /D/, as soon as (s)he pronounces the word just heard (s)he will produce a B, because the articulatorily-based constraints are then active again. For example, a native speaker of modern French might hear a Venetian /s̺/ as a distinct category /s̺/, i.e. realise that it is neither native French /s/ nor native French /ʃ/, but as soon as (s)he will say the word heard to another native French speaker, (s)he will produce a [s] or [ʃ], having the constraint *B(UNGROUNDED) ranked above F.IDENT(SIB), and that second native speaker will hear /s/ or /ʃ/ accordingly. Rejecting the active role of articulatorily motivated markedness constraints in perception is thus not even problematic for the sibilant typology proposed here: the nativisation effect will arise in any case. The observation made by Bybee (2001: 54) that there is a "strong tendency for speakers to reuse a single set of highly entrenched neuromotor patterns and to substitute members of this set for novel or less common configurations" is thus reflected in this typology, without needing any structural, inventory-related constraint such as a hypothetical *DIFFERENTGESTURE, similar to the hypothetical *ASYMMETRY in 4.1.1.2. Speakers reuse their own articulatory configurations for those of non-native phonemes because either they perceive the non-native phoneme as a native phoneme (the bidirectional hypothesis), or they hear the difference between the non-native and native category but are not able to reproduce the non-native category faithfully because articulatorily-based markedness constraints against the non-native phoneme are undominated by faithfulness.

5.3 A sibilant typology in Neural Networks

As illustrated in 2.3.1, the conflict between the discrete, categorical behaviour of phonemic inventories and the gradient, multiply divisible phonetic continuum poses a challenge to any linguistic theory. Neural Networks aim to resolve this conflict by reconciling gradient data ranges and discrete categories in one single framework (Iskarous 2017; Boersma et al. 2020; Grimaldi 2018). They allow us to model emergent category creation based on variable input, i.e. modelling how children learn the sibilant inventory of their native language from the variable sibilant realisations of their environment, without knowing beforehand which are the phoneme categories. The BiPhon assumption of bidirectionality is applied to Neural Networks in the sense that in production, the connection weights of the network are reused symmetrically: the connection from node A to node B is the same as the connection from node B to node A (Seinhorst 2012: 18), and activity will spread evenly in both directions.

A neural network is organised as a series of nodes arranged in layers, as illustrated in fig. LXXVII. The lower layer in this neural net corresponds to the Auditory Form

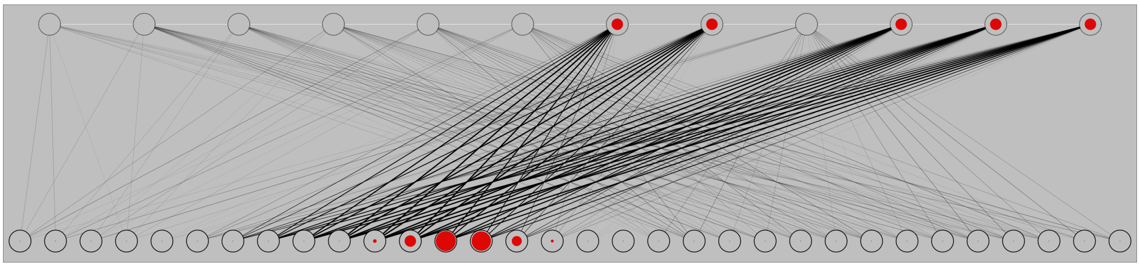


Figure LXXVII: *Example of a neural network (in Praat; Boersma and Weenink 2019). This network has been trained with input data from a B language.*

layer (AudF), which is an abstraction of the phonetic values that can be perceived on the Centre of Gravity continuum. In conformity with the IPA convention of representing articulations that are further front starting from the left (e.g. /p, t, k/ instead of /k, t, p/), the frontmost sibilant [s̪] corresponds to the leftmost node of the AudF layer. This means that from left to right, the CoG decreases from approximatively 7500 Hz, [s̪], to 2000 Hz, [s̠] (or similarly on the perceptual ERB scale). On the continuum of 32 AudF nodes represented above, the first 25% of nodes (i.e. nodes 1 to 8) correspond to the phonemic sibilant category A, nodes 9 to 16 to B, nodes 17 to 24 to C and nodes 25 to 32 to D.

The upper layer corresponds to the Surface Form layer (SF), which does not constitute a continuum but rather is an abstraction of the link between auditory forms and higher-level representations (Underlying Form, lexicon, semantics etc.). Being more to the left or to the right on the SF layer does thus not 'mean anything', and the same acoustic region of AudF in the same language does not always create strong connections to the same SF nodes (quite the contrary: the pattern varies randomly from one simulation to another). The two layers illustrated here, AudF and SF, are initially interconnected with every possible connection between two nodes of different layers, with very small random connection weights (between 0, inclusive, and 0.1, exclusive, i.e. $0 \leq randomweight < 0.1$). The connection weight, i.e. how 'strong' a connection is, is positively correlated with line thickness on the figure so that the strongest connections are the darkest and thickest. As can be seen in fig. LXXVII, the connection weights are strongest around nodes 7 to 15 on the AudF layer, among which most correspond to the B category (9 to 16). The white connections between all SF nodes are negative weights of -0.1, created "to make sure that the SF nodes inhibit each other's activities (...) This inhibition militates against different SF nodes being on at the same time. As a result, assuming small random differences in activities between SF nodes (caused by the different random initial weights), different SF nodes will come to specialize in different regions of AudF, so that they can be on at different times" (Boersma et al. 2020: 139, 146).

The red circle inside the nodes represent activation (i.e. neural excitation), with redder nodes being more strongly activated. In fig. LXXVII, a sibilant noise has been heard with a CoG located in between nodes 13 and 14, though slightly closer to node 13 (with an activation of 0.942), corresponding to a sibilant realisation situated in between [s̺] and [ʃ], although closer to the former. This incoming token does not only activate node 13, but also every single node with an activation strength correlating with distance (in a Gaussian shape, i.e. drastically less for nodes further away). This faithfully reflects the behaviour of the human ear (Chládková 2014: 89; Boersma et al. 2020: 141): a perceived CoG activates a small region of the basilar membrane (with a standard deviation of 4 %), but with most of the activation at the epicentre. A crucial aspect is that when activity spreads and the weights are updated, the connections between strongly activated nodes are reinforced; at the same time, stronger connections allow for activity to spread more. It thus becomes clear that in fig. LXXVII, incoming tokens in the auditory region of B have been heard much more frequently than in other regions and therefore, the strength of these connections has been reinforced. The algorithm used here to update the weights, the *inoutstar* learning rule (Boersma et al. 2020: 133ff.), does nothing when two nodes have no activation, weakens a connection if only one of the two nodes is activated, and strengthens a connection if both of its nodes are activated.

It would be an undesirable property of the network if activity spread from a strongly activated AudF node to a SF node, which would then send back much of this activity to other AudF nodes that are less strongly activated, which would then send back some activity to SF nodes, and so forth. To prevent this 'contagious' behaviour, all AudF nodes are *clamped* in perception before spreading activities, which allows them to send activity but not to receive any activity. A clamped node is represented with a full circle, while an unclamped node is drawn with dashed lines, as is visible in fig. LXXVIII.

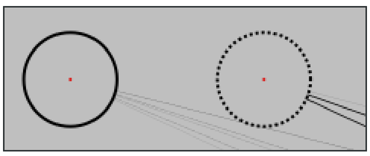


Figure LXXVIII: *Detail of a clamped and an unclamped node, respectively (in Praat; Boersma and Weenink 2019).*

To summarise, our network as depicted in fig. LXXVII is organised in two layers, AudF and SF, interconnected with all possible connections between nodes of different layers, which have initially small random weights, and with negative weights of -0.1 between all SF nodes to inhibit simultaneous activations. When an incoming token falls on e.g. 48% of the continuum (i.e. halfway between [s] and [ʃ]), the closest AudF nodes are activated strongest, with the strength of activation following the pattern of a Gaussian curve. The clamped AudF nodes then spread activation to SF nodes, with more activation being transmitted through stronger connections. The connection weights are then updated, proportionally to the combined activation of the nodes that they connect.

Readers interested in a more detailed account of Neural Networks simulations with sibilant inventories are referred to Boersma et al. (2020).

5.3.1 Boersma et al. (2018)'s model

Boersma et al. (2020) model the acquisition of several sibilant inventory types by neural networks, illustrating the emergence of probabilistic categorising behaviour in children from a variable acoustic input. Since this acoustic input is perceived without anything telling the learning network which category it was, this method of learning is called *unsupervised* learning, which has been shown to be preferable to supervised learning in 5.1.1.3 (*Discarding predefined categories in learning perception*). In total, Seinhorst (2012) and Boersma et al. (2020) simulate the successful acquisition of a 1s inventory with an initial sibilant realisation /s̪̠/ that automatically becomes /s̠/, of a well-dispersed 2s with /s - ʃ/, an overly-dispersed 2s that becomes well-dispersed, a less-dispersed 2s that becomes well-dispersed, a 3s with /s̪ - ɕ - ṣ/ and a well-dispersed 2s that becomes such a 3s (by means of the shift /s^j/ > /ɕ/). Seinhorst (2012: 36-40) also simulates the acquisition of a bimodally distributed inventory, in which one of the two sibilant categories is distributed both left and right from the other category. One shall notice that their simulations do not leave room for less-dispersed languages like the AB or BC languages, which forcibly become AC languages after a few generations. Also, their 1s and 3s languages end up being centred around the centre of the continuum rather than slightly towards the left/front, because they do not include posterior markedness in their model.

In all cases, auditory dispersion emerges as follows in their simulations: an incoming token falling precisely in between e.g. the two categories of a less-dispersed 2s gives a part of its activation to the SF nodes of one phoneme category and another part of its activation to the SF nodes of the other phoneme category (i.e. it activates both categories ambiguously); a token falling on the left of the front category or on the right of the back category, i.e. in a peripheral zone, gives all of its activation to the closest category[1] and therefore activates that category more strongly than if it had fallen on the other side, in between the two categories, as illustrated in fig. LXXIX. When the network will 'speak', i.e. activate and clamp the SF nodes of a category and spread activities towards AudF, peripheral values on AudF will be activated more strongly, since the connections towards

[1] Recall that from the perspective of the language user and the NN model, auditory distance does not exist as a concept and is not required for anything to function.

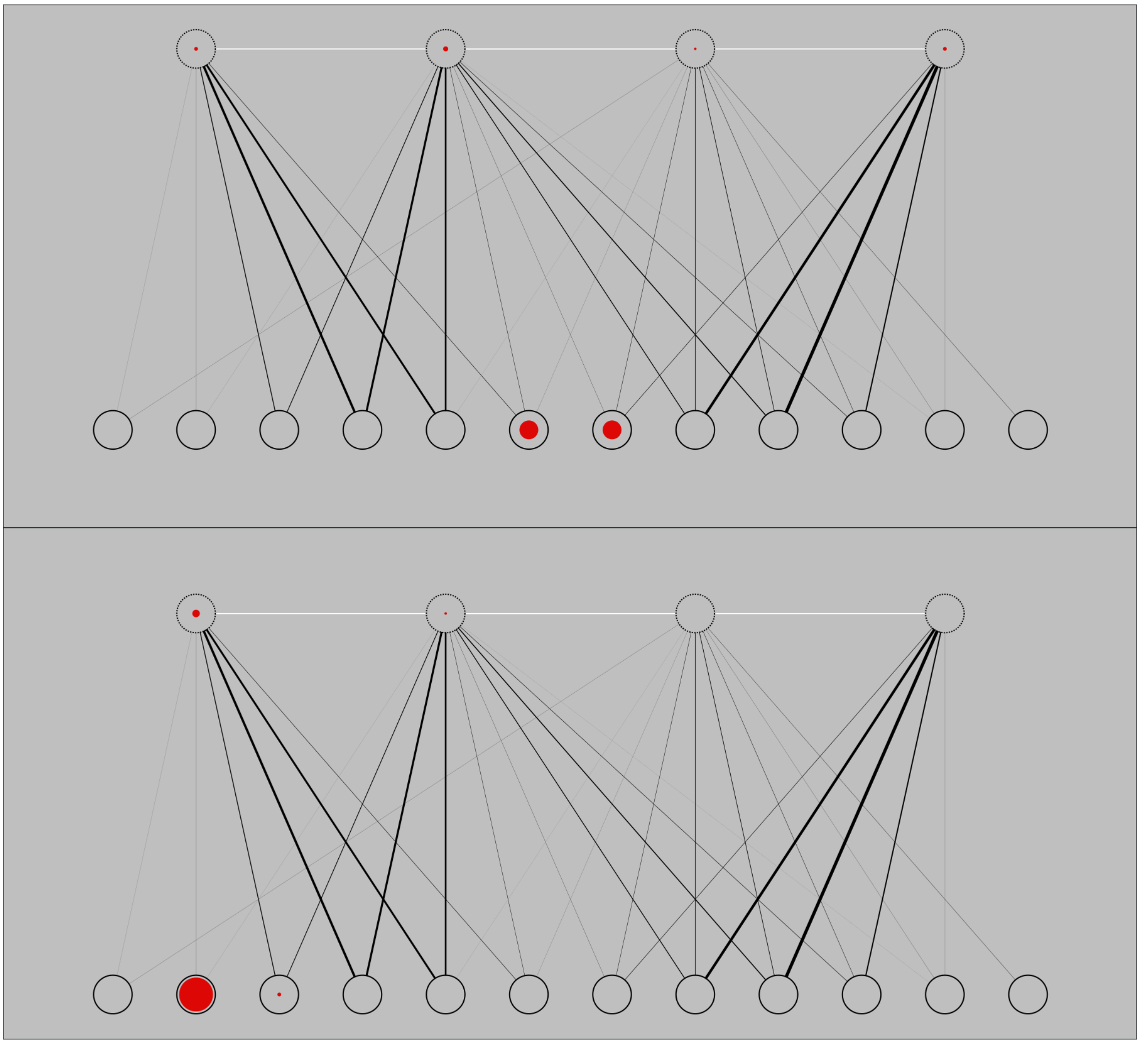


Figure LXXIX: *Illustration of how auditory dispersion arises in Neural Networks (in Praat; Boersma and Weenink 2019).*

and from peripheral zones have been reinforced more strongly. More activation in peripheral zones then means a larger probability of being pronounced, and the following generations will thus perceive an input distribution with more tokens around peripheral zones (if not counteracted by centralising or fronting articulatory markedness).

Fig. LXXIX depicts two activation patterns in one identical neural network that has been trained with tokens randomly and stochastically drawn with equal probabilities from two phoneme categories, situated at 4.3/12 AudF nodes (30%) and 8.7/12 AudF nodes (70%), respectively. It can be seen by the line thickness that the first two SF nodes have come to correspond to the first category, while the last two SF nodes (and especially the last node) have come to correspond to the activation of the second category. The incoming token in the lower example is situated at a distance of 2.2 left from the peak of the first category, i.e. at 2.1. The incoming token in the upper example is situated at the same distance of 2.2 from the peak of the first category, but on the other side, which is perfectly in the middle of the continuum and perfectly in between the two categories, at

AudF node 6.5. The total activation is 1.0 at the two nearest AudF nodes and the activity is spread 400 times in both cases. Fig. LXXX represents the activations of the 4 SF nodes as

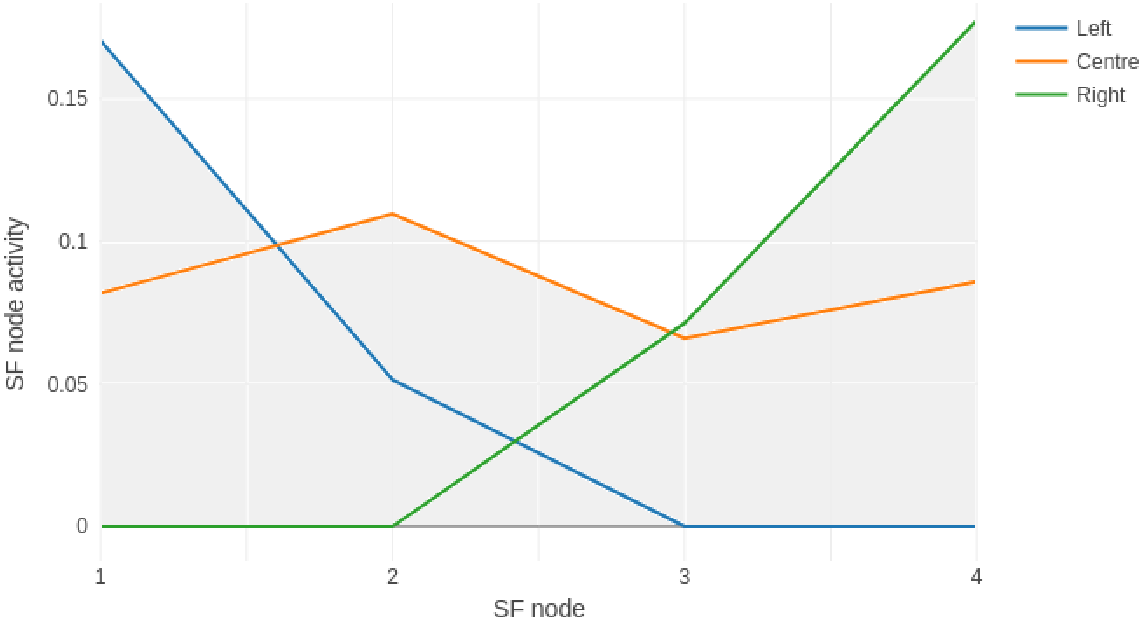


Figure LXXX: *Comparison of the activities of the 4 SF nodes in the examples of fig. LXXIX, together with the activities of the 4 SF nodes when AudF node 10.9 is activated (i.e. 2.2 on the right from the second phoneme category) and activities spread 400 times.*

a percentage of total activation. When the incoming token is left from the first category, the activation of the two first SF nodes is 115.92% that of the two first SF nodes when the incoming token is in the middle, which means that a peripheral token activates the nodes of the nearest category more strongly than a token at precisely the same distance that is also near another category. On the other side of the second category, an incoming token situated 2.2 on the right of the second category activates SF nodes 3 and 4 with 163.69% of the activation transmitted by the central token to these two nodes. If the weights were updated, the peripheral tokens would logically have reinforced the connections towards their respective phoneme category more strongly than the central token.

The NN model by Boersma et al. can thus model at the same time unsupervised, gradual category creation as well as auditory dispersion, without using any teleological means and with only the assumption of bidirectionality. This model is (Boersma et al. 2020: 104):

> [A] model that accounts for at least four types of valid behavioural data, namely 1) the generalizations that phonologists have found within and across languages, 2) the phenomena that psycholinguists and speech researchers have found by observing speakers, listeners, and language-acquiring children, 3) the mergers, splits, chain shifts and other sound change phenomena found by historical phonologists and dialectologists, and 4) the phenomena that have been observed when languages come in contact, such as loanword adaptations. Besides having to account for all these

types of behavioural data, the model will have to be compatible with what is known about the biology of the human brain, because that is where language is produced and comprehended.

It is thus not only biologically more plausible in the sense that it parallels neural activation patterns, but also because categories are never explicitly transmitted to the learning child but induced from the varying input. It does not only account for linguistic patterns attested synchronically, but also for diachronically observed sound shifts. The assumption of bidirectionality and the assumption that the network changes much less in adults accurately reflect the observation that loanwords with a non-native phoneme are treated in perception as having a less prototypical exemplar of a native phoneme, but not as an independent phoneme if it is too rare in the input and/or if the learning flexibility of the network is slowed down by age.

In the next subsection, I describe what features I add to the existing NN model in this dissertation (more precisely, in 5.4.1), and how they contribute to an even better understanding and modelling of the typology of sibilant inventories.

5.3.2 Additions to Boersma et al. (2018)'s model

A first theoretical problem is that depending on how one (fine-)tunes the learning ability and flexibility of the neural network, it can or cannot learn languages with 4, 5, 6 etc. distinct places of articulation. There is little doubt that human beings can most probably not acquire a sibilant inventory with e.g. 10 PoAs for sibilants, but this is not certain for the computer: in fact, it all depends on how the neural network is built. For example, the script illustrated in LXXXI and LXXXII (provided by S. Hamann and P. Boersma) does not or does let the network learn inventories with 10 PoAs, depending on the values specified for the parameters *peak_sharpness* and *auditory_sharpness*. It is

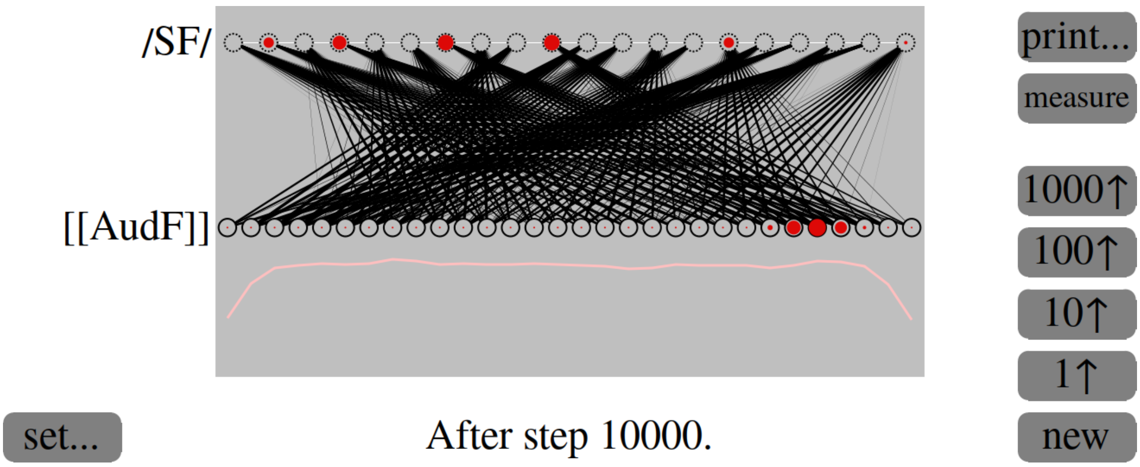


Figure LXXXI: *Simulation of a network not learning an inventory with 10 PoAs, with a peak sharpness of 3 and an auditory sharpness of 25 (in Praat; Boersma and Weenink 2019).*

of course a radical example, but illustrates the problem caused by the precision of computers, which can have a much larger perceptual acuity than humans if asked to do so.

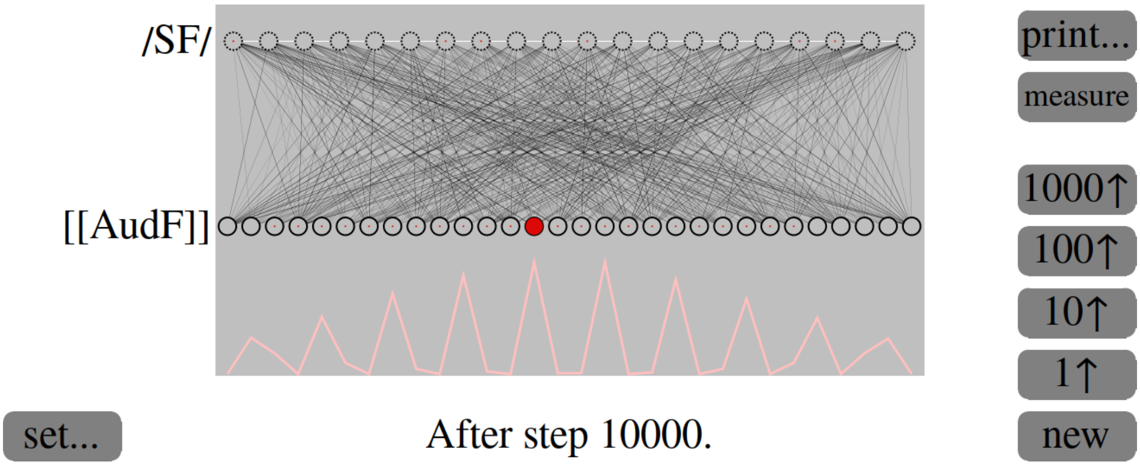


Figure LXXXII: *Simulation of a network learning an inventory with 10 PoAs, with a peak sharpness of 100 and an auditory sharpness of 100 (in Praat; Boersma and Weenink 2019).*

To tackle this problem, the perception experiment in 2.3 was designed to quantify the average learnability of sibilant inventories for human ears, and thus allows us to tune the neural network according to this average perceptual sensitivity. Boersma et al. (2020: 155f.) experiment with the number of nodes and show that the larger the number of SF nodes in the network, the more categories can be learnt; a network with 10 SF nodes succeeds in learning at most 4 categories. The results of the perception experiment in 2.3, nevertheless, allow us to tune the learning ability of the network in an even finer way so that the low percentage of humans able to distinguish 5 categories might match the percentage of simulations in which 5 categories are effectively learnt.

Secondly, to fill the gap of the attested sibilant inventories not modelled by Boersma et al. (2020) and Seinhorst (2012), I propose to illustrate their acquisition. This includes adding the AB and BC 2s inventories as well as the 4s inventory to the model, but also to include posterior markedness and therefore to show how 1s and 3s languages end up being centred around B, i.e. at approximatively 37.5% of the acoustic continuum, rather than 50%. Since the retroflex and non-retroflex types of 3s and 4s are not distinguished primarily by their PoA/CoG values but rather by their distributedness (Nowak 2006; Hu 2008), retroflex and non-retroflex 3s variants are not distinguished in the simulations (which only have one continuum representing PoA).

Thirdly, the NN simulations I perform in 5.4.1 differ from the preexisting ones by linking the neural network, which models the interface between Surface Form and Auditory Form, to the OT grammar, which determines the mapping from an Underlying Form to a Surface Form. With this innovative way of formalising the interaction between perception and production, combining Neural Networks and Optimality Theory, I aim to show that they are compatible frameworks that complement and do not necessarily contradict each other. The network first learns and deduces which are the categories based on the input distribution, and the OT grammar then models the production and reranking process that leads to the correct production for the inventory learnt. It must be emphasised here that simulations of the acquisition and transmission of sibilant inventories in NN are *also* successful without the OT grammar. Neural networks alone can account for the typology of sibilant inventories, as much as OT alone can (with the particular exception of auditory dispersion, see 5.2.2.6); I combine both here to highlight their successful interaction and compatibility.

Furthermore, I propose to model language contact and functional load within Neural Networks: instead of being exposed to the input distribution of one type of sibilant inventory, the neural network is fed with the input distribution of all possible pairs of sibilant inventories (except the 4s inventory), with different contact intensities. For example, I run the language acquisition simulation with a network only exposed to a 1s, AB 2s, BC 2s and 3s, but also with e.g. a network exposed to input from a 1s 20% of the time and exposed to input from an AB 2s 80% of the time, the same with 40 - 60% instead, with 60 - 40% instead, and so forth. Besides this, I include functional load in the form of different frequencies between the phoneme categories in the input, since it is widely observed that the frequencies of sibilant phonemes vary widely in natural languages (see Peust 2008 and fig. LXXXIII). With four possible functional load percentages for each category (20, 40, 60 or 80% frequency) and five different percentages of language contact intensity (0, 20, 40, 60 and 80%), it becomes much more likely that *all* possible sibilant inventories will have been predicted and modelled in this dissertation.

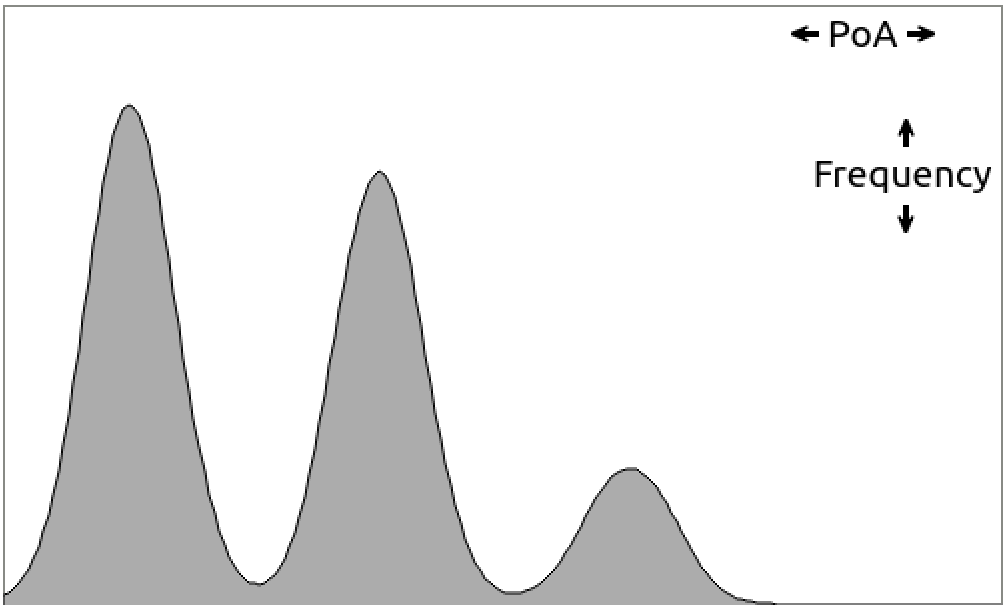


Figure LXXXIII: *Example of language contact and functional load: the input distribution of a child growing up in an environment with 60% French (78.6% /s, z/ vs. 21.4% /ʃ, ʒ/) and 40% Central Greek (100% /s̺, z̺, t͡s̺, d͡z̺/). A is centred at 12.5%, B at 37.5% and C at 62.5% of the auditory continuum. Phoneme frequencies are based on Peust (2008).*

5.4 Proposed model: Optimality Theory and Neural Networks

In the next subsection, I intend to model the acquisition and evolution of all possible sibilant inventories in a combination of OT and NN, using the software Praat (Boersma and Weenink 2019). Praat can load, modify and save OT grammars as an object named *OT-Grammar* (with an eponymous file extension) as well as create, modify and save neural networks as an object named *Network* (with an eponymous file extension). Importantly, I will not only show in 5.4.1.5 that the inventories predicted by the OT grammar and the well-dispersed 2s are acquired stably over the generations, but also in 5.4.1.6 that one can start with the initial distribution of an inventory that is not attested (e.g. a language with only C), and that over the course of time, it becomes an attested inventory (in this example, a language with only B). When the distribution of a well-dispersed AC language is fed to the network, although it was not predicted by the OT typology as explained in 5.2.2.6, it is learnt stably.

In subsection 5.4.2, it will be argued that the inventories predicted by the OT and NN model are still abstractions, in the sense that they occur in natural languages as a vast range of different variants with or without accidental gaps, i.e. as 'surface inventories'. For each abstract inventory, there will thus be several possible variants (one with a gap at X, one with a gap at Y etc.), of which only a small proportion is attested and present in the SibInv database. These surface inventories will be matched with the inventories of the database in 6.1.

5.4.1 OT and NN: the interaction of production and perception

Several publications in Optimality Theory reveal the success with which the supervised or unsupervised learning of OT grammars has been modelled computationally (cfr. e.g. Jarosz 2006). The algorithms proposed in the literature aim to account for how children learn the lexicon *and* the segment inventory as well as the phonological constraint ranking of their native language at the same time, and several among these algorithms manage to do so in an unsupervised manner, i.e. by only having access to the unstructured phonetic data. Recall from 5.1.1.3 that a model is preferable if it can model the acquisition of language without 'telling' the learning mechanism which are the categories, as much as a child acquires a language in real life without being told most of the time e.g. 'This is a /ʃ/'.

The main idea of the proposed model is to start with one stochastic OTGrammar object for each language predicted by the basic OT sibilant typology (except for the sibilantless language, which is not relevant here because the network would have nothing to learn); then, to let a hypothetical perfect 'adult' speaker produce the input distribution for each of these OTGrammars, which will be heard by a 'child' through a neural network; then to let this 'child' learn and rerank the relevant constraints to end up producing tokens from the acquired sibilant inventory as an 'adult' speaker, then transmitting these produced tokens to further generations of 'children'. In these simulations, the ‘deeper’ domains of language are in the hands of OT, which evaluates and maps an Underlying Form (UF) into a Surface Form (SF) and inversely, while the actual process of production and perception towards or from surface acoustic input/output is modelled with Neural Networks. The former is thus responsible for the phonological, discrete behaviour and the latter for the interface between discrete categories and continuous phonetic values.

The entire simulation is controlled by a simple Python script, provided here as script 1 as well as in 8.5.1 in the appendix.

Script 1: *Python script that runs the OT-NN scripts.*

```
#!/usr/bin/env python3

import os

os.system ('/usr/bin/praat --run "AdultOTandNN.praat" "30000"')
# The number specified as argument is the number of tokens per
    generation.

for x in range(1, 21):
   # range(1, 11) for 10 generations, (1, 37) for 36 generations etc.
   gen = "{}"
   print("The current generation is generation " + gen.format(x))
   os.system ('/usr/bin/praat --run "ChildOTandNN.praat" "' + gen.
       format(x) + '"')
   # The number specified as argument refers to the generation number.
```

This script executes the Praat scripts one after the other. The two Praat scripts open existing files, perform their calculations and create/overwrite files based on these calculations. How the entire simulation of production and perception works is illustrated in fig. LXXXIV.

A. Production of the 'perfect adult' input distribution

The *AdultOTandNN.praat* script (detailed in 8.5 in the appendix) takes each of the 5 languages with sibilants predicted by the factorial typology of the basic production grammar (in 5.2.2), and assumes that an adult speaker has acquired in a typical way the ranking of the OT grammar of her/his native sibilant inventory. These 5 grammars are created as OTGrammar objects in Praat and are located in the folder *InitialObjects*, whose contents do not change during the simulations. It is abstracted away from voicing and affricates, which are not necessary distinctions for the needs of the simulations. Non-sibilants are not included in the 5 OTGrammars of the adult speaker, since we only want to model the acquisition of the sibilant inventory by the children, not their acquisition of stops, and because sibilants are almost never accidentally stopped in adult speech (cfr. Baroni 2014*b*: 34). However, the M.SIB constraint and non-sibilant outputs are included in the OTGrammars of the child, because very young children first realise sibilants as stops (cfr. Fletcher 1989: 746, cited in 4.6.1.3), something which shall be modelled as a part of early sibilant production in *ChildOTandNN.praat* (D.).

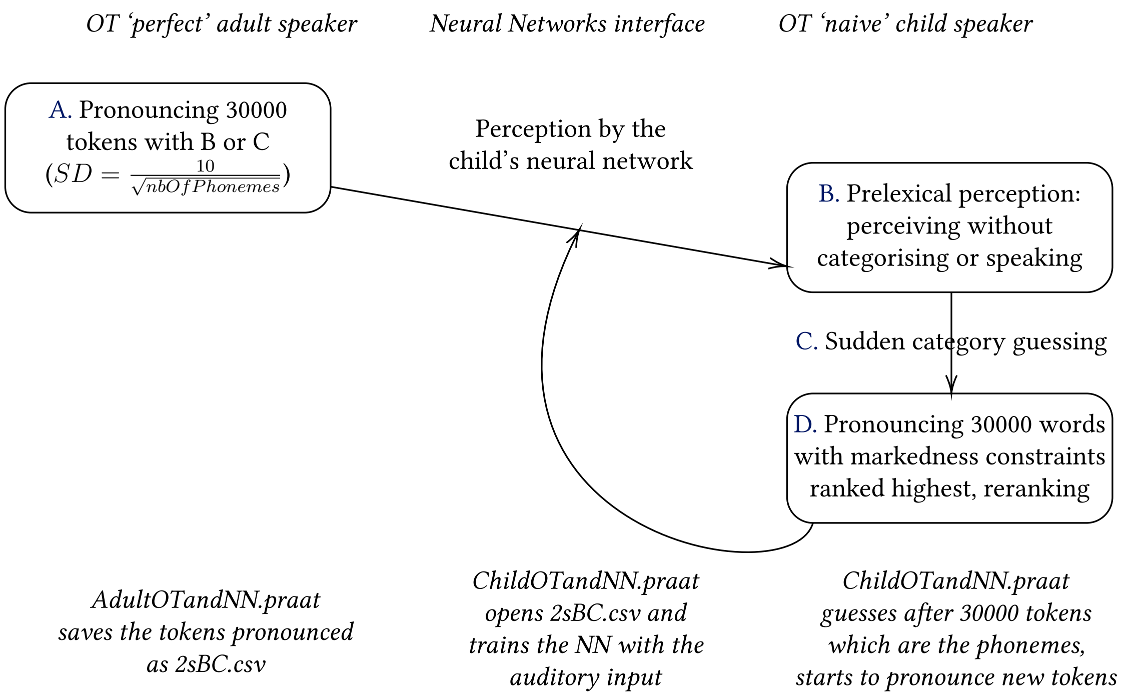


Figure LXXXIV: *Schematic representation of how OT and NN interact in the learning process of the sibilant inventory BC.*

In each OTGrammar, the fact that one constraint crucially dominates another is expressed as the fact that its ranking value is 10 points higher than the constraints that it must dominate (this ranking difference of 10 is fairly standard in Praat OTGrammars, cfr. Boersma 2007). For example, in all 5 OTGrammars, m.Sib is ranked at 100 and f.Ident(Sib) at least at 110. Constraints that are not crucially ranked are assigned random ranking values between -10 wrt. the lowest crucially ranked constraint and +10 wrt. the highest crucially ranked constraint (for example, m.PostD(Sib) is randomly ranked between 90 and 130 in the BC language). For each token produced, the non-crucially ranked constraints are reattributed a new random ranking. The rankings in the 5 OTGrammars are summarised in table 21.

The 'adult speaker' first randomly selects one of the categories used in the native language, according to the functional load of that category, and can select one only among the existing categories (for example, in a 1s, only the B category is ever selected as the underlying input). An alternative to this selection mode would be selecting any possible input, which conforms more to the assumption of Richness of the Base (Prince and Smolensky 1993/2004) but has the disadvantage that it creates disparities in the functional loads: an ABC language will have C twice more frequently than A and B compared to the predefined functional loads, because not only C but also D is mapped to C. Nevertheless, this does not mean that the only outputs to ever be selected will be faithful to their input: the evaluation noise of 4 allows for unfaithful mappings to happen though rather infrequently (0.04% of the time with a constraint ranking distance of 10, according to Boersma 2007). This accurately mirrors the fact that in the lexicon, only and precisely the existing phonemes of a language are stored (e.g. /C/), but in performance, a tongue

	Ranking value:	1s	2s (AB)	2s (BC)	3s	4s
Crucial	120	M.PERIPH(SIB), M.POSTCD(SIB)	M.POSTCD(SIB)	M.PERIPH(SIB)	M.POSTD(SIB)	
	110	F.IDENT(SIB)	F.IDENT(SIB)	F.IDENT(SIB)	F.IDENT(SIB)	F.IDENT(SIB)
	100	M.SIB	M.SIB, M.PERIPH(SIB)	M.SIB, M.POSTCD(SIB)	M.SIB, M.PERIPH(SIB), M.POSTCD(SIB)	M.SIB, M.PERIPH(SIB), M.POSTCD(SIB), M.POSTD(SIB)
Random	$R \in \begin{cases} max+10 \\ min-10 \end{cases}$	M.POSTD(SIB)	M.POSTD(SIB)	M.POSTD(SIB)		

Table 21: *Constraint ranking values for the OTGrammar objects of the 5 initial languages. Note:* crucial *means that without that particular constraint ranking hierarchy, the resulting language becomes another language.*

slip or accidental coarticulation might lead from time to time to other phonetic realisations (e.g. [B]). For example, in one sample of the 4s language, unfaithful mappings happen 4.226% of the time in the speech of the adult speaker (2.823% A, C or D towards B, and 1.403% D towards C).

After having selected the pronounced surface form, this SF category is converted to a number on a scale from 0 (dental sibilant, i.e. the most anterior one) to 100 (subapical retroflex sibilant, i.e. the most posterior one). This happens by selecting a Gaussian random value around a peak at 12.5% of the continuum (A), 37.5% (B), 62.5% (C) or 87.5% (D) respectively, with a standard deviation of $\frac{10}{\sqrt{nbOfPhonemes}}$, as illustrated in fig. LXXXV and LXXXVI. The standard deviation is inversely proportional to the number of phone-

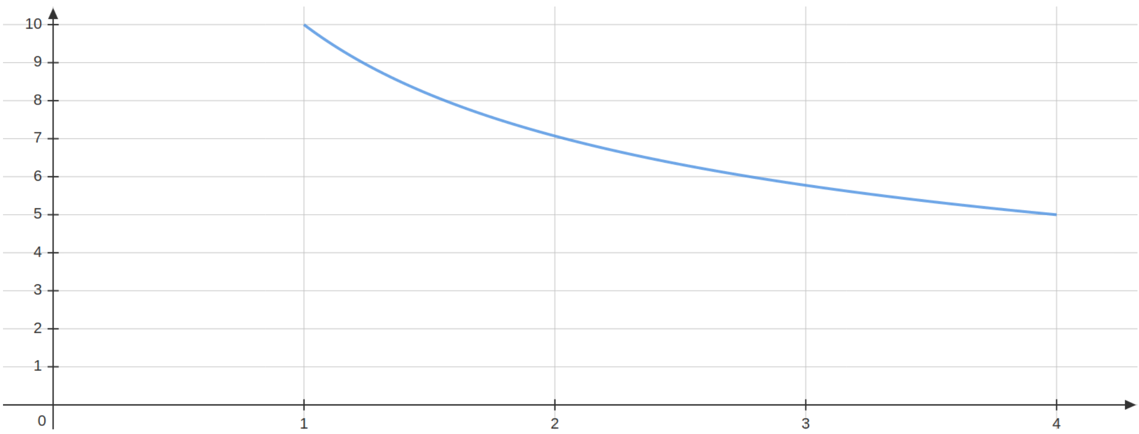


Figure LXXXV: *The percentual standard deviation value for the Gaussian random selection of categories (vertical axis) as a function of the total number of phoneme categories in the inventory (horizontal axis), from 1 to 4.*

mic categories, as observed in native languages (cfr. e.g. Boersma and Hamann 2008; Jaggers and Baese-Berk 2019; Sieber 2020).

A last precaution is taken to ensure that the randomly selected phonetic value is not

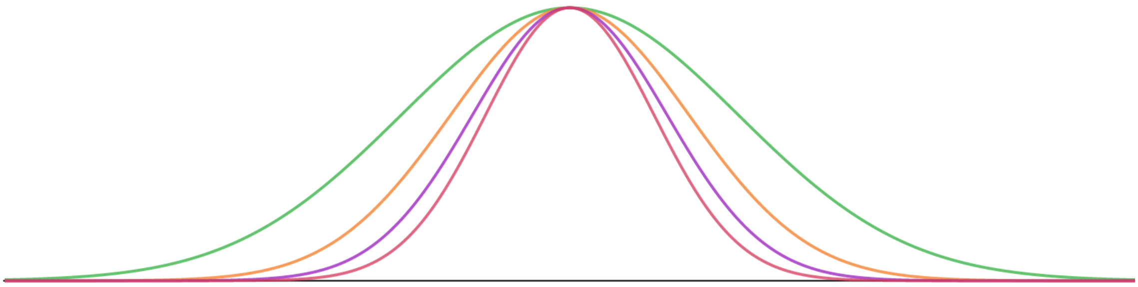


Figure LXXXVI: *Visual representation of the relative widths of the input distributions for 1 (green) to 4 (red) phoneme categories.*

situated below 0 or above 100, as illustrated in script 2.

Script 2: *Selection of a random phonetic value for the category A (centred around 12.5% of the auditory continuum) in Praat (Boersma and Weenink 2019).*

```
standev = 10/sqrt(nbOfPhonemes)
if pronouncedSFtoken$ = "A"
   repeat
      pronouncedvalue = randomGauss(12.5, standev)
   until pronouncedvalue >= 0 && pronouncedvalue <= 100
endif
```

The pronounced tokens (together with the intended UF and SF category for each token, which will not be accessible to the child) are then saved as an input distribution and placed in the folder *ChildOTandNN/Distributions*, which the script *ChildOTandNN.praat* will access.

B. Prelexical perception by the 'child' neural network

The script *ChildOTandNN.praat* loads the input distribution in the *Distributions* folder, and creates an empty neural network. This symmetric network has a spreading rate of 0.01, linear activity clipping, an activity leak of 1, shunting set to 0, activities between 0 and 1, weights between -1 and 1, a learning rate of 0.01 and a weight leak of 0. Instar and outstar are both set to 0.5, which guarantees bidirectionality. The SF nodes are connected to all other SF nodes with weights of -0.1, and all the connections between AudF and SF are assigned random weights between 0 (inclusive) and 0.1 (exclusive). All of these parameters are set to the same values as those specified in Boersma et al. (2020). The network I use has 32 AudF nodes and 12 SF nodes, as already shown in fig. LXXVII.[2] The AudF nodes, from 1 to 32, correspond to a perceptual continuum in ERB from 33 to 20 ERB, which corresponds to a perceptual continuum from 9610.92 Hertz to 1788.95 Hertz. The ERB scale is used here, because it accurately reflects the human perceptual acuity with equal distances in different auditory regions, unlike the Hertz scale. A representation of the acoustic values in ERB and Hertz per AudF node

[2] The reason for the two additional nodes in each layer compared to Boersma et al. (2020) is the divisibility of 32 and 12 by 4. In the latter case, the 12 SF nodes can furthermore be subdivided evenly into 1, 2, 3 or 4 categories.

is provided in fig. LXXXVII. The values used here are only a little different from those

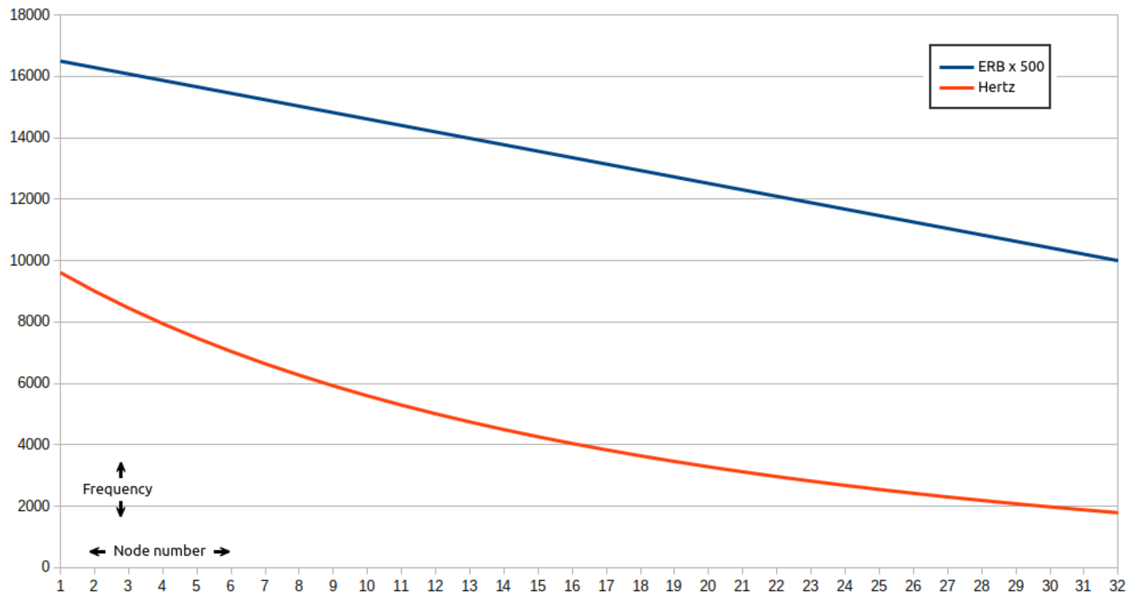


Figure LXXXVII: *The values in ERB and Hertz associated with AudF nodes 1 to 32. ERB values have been multiplied by 500 only to be able to represent them on the same figure as Hertz.*

in Boersma and Hamann (2008: 233), who "assume somewhat arbitrarily that relevant spectral mean values for sibilants universally lie between 20.0 Erb (1789 Hz) and 36.0 Erb (16177 Hz)". With my subdivision of the AudF continuum in 4 categories (A, B, C and D), using 36 ERB as an upper threshold would situate the average phonetic values for the Polish sibilants that they report (pg. 230f.) too much towards the lower end of the continuum, close to a BCD rather than ABC language; with a continuum from 33 to 20 ERB, the values reported are located near the centre of each category (A, B, C).

The network perceives the 30000 tokens in the input distribution as follows: for each token, the activities of the entire network are first set to 0; then, the activities of the AudF nodes are determined according to a Gaussian activation shape with a standard deviation of 0.05 (5%) of the continuum.[3] The formula used for activity spreading at AudF is the following:

$$a_n = e^{\frac{-0.5\cdot(n-gaussianpeak)^2}{(0.05\cdot(audf.NbOfNodes-1))^2}}$$

[3] Boersma et al. (2020: 141) use 0.04 instead, which according to them reflects the extent of activation spread on the human basilar membrane. Note that with a smaller value, the network is more likely to learn one large category as several adjacent categories instead, and can thus not learn a category with a large standard deviation (such as e.g. 10% in a 1s) as one category. The simulations with 0.04 did not allow to learn categories as wide as the two sibilants in a well-dispersed 2s (width which is necessary to allow auditory dispersion to operate), while the simulations with 0.06 merged categories as close as A-B or B-C. Setting the parameter to 0.05 reconciled both, allowing to model auditory dispersion with distant categories (AC) without merging adjacent categories (AB and BC).

where a is the activity of an AudF node n, n the number of node n (from 1 to 32), $gaussianpeak$ the centre of the Gaussian peak i.e. the precise phonetic value that has been perceived on the auditory continuum, and $audf.NbOfNodes$ the number of nodes (which must be lowered by 1, because the extent of the continuum is from node 1 to 32 rather than from 0 to 32). $(n - gaussianpeak)^2$ amounts to the auditory distance between the node whose activity is set and the location of the token perceived on the auditory continuum. An example of activity spreading in a network that has perceived 750 tokens from an ABC distribution is provided in fig. LXXXVIII. Connections whose

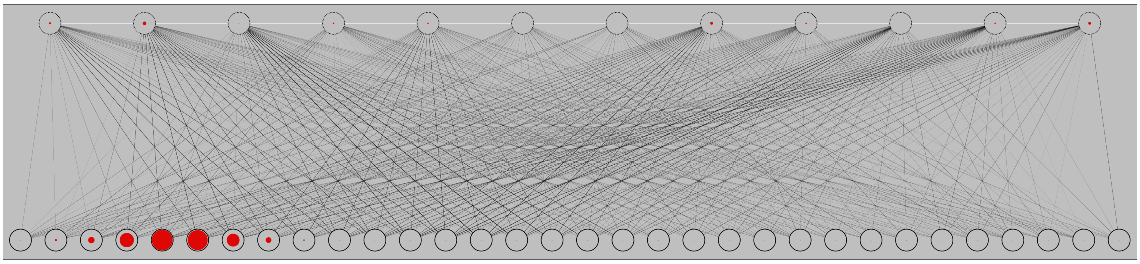


Figure LXXXVIII: *Activation pattern at AudF (lower layer) and SF (upper layer) when perceiving category A in a 3s, after 750 learning steps. It can already be seen that some SF nodes are activated more strongly than others.*

nodes are both activated will be reinforced proportionally, those with only one node activated will be weakened proportionally, and those whose nodes are both inactive will not change. This leads to the fact that connections with auditory zones that are almost never activated keep their initial random weights and low SF activations, while connections from auditory regions that correspond to a category will be strongest around the centre of the category and activate a subset of SF nodes very strongly. Fig. LXXXIX shows a network after 30000 tokens from a 3s distribution.

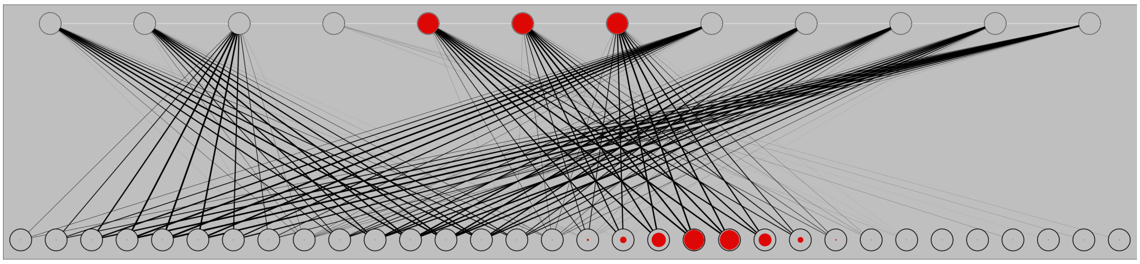


Figure LXXXIX: *Activation pattern at AudF (lower layer) and SF (upper layer) when perceiving category C in a 3s, after 30000 learning steps.*

The fact that auditory zones that are almost never used in the language keep their initial weak random weights constitutes a difference with respect to the OT typology. Indeed, all OT grammars force the mapping of non-native sibilants to some category (in the proposed typology, the phonetically closest category). Here, non-native sibilants are not mapped to any category if they are distant from existing categories, meaning that a speaker of an AB language does not identify a [D] as /B/, but as no category in particular, an 'unfamiliar sound' that is not part of the inventory (as much as clicks are not assimilated by native speakers of languages without clicks to native categories such as stops, but as 'unfamiliar/non-speech sounds'; Best et al. 1988; Fuhrmeister 2020). Nevertheless, an acoustically close non-native sibilant (e.g. a C not too distant from the B category in an AB language) will activate the SF nodes for B, though more weakly than a prototypical B token; this reflects the fact that a non-native phoneme is perceived as the nearest category but as a *bad exemplar* of this category proportionally to the acoustic distance. This is thus a major point of divergence wrt. the OT typology and Boersma et al. (2020)'s model: a nearby non-native sibilant (e.g. [ʃ] in a 1s) is perceived as a non-prototypical exemplar of a native phoneme category, while a very distant non-native sibilant (e.g. [s̺] in a 1s) is not perceived as an exemplar of any category.

C. Postlexical perception: category guessing

Now, categorical perception has emerged in the network of the child: the number of stable activation patterns has come to reflect the number of categories in the input distribution, though the child was never told what categories there are. This distributional category learning behaviour was observed in language acquisition (Chládková 2014: 91). In fig. XC, the activation patterns in a network fed with a BC input distribution are twofold: SF nodes 2, 9 and 11 correspond to category B, while SF nodes 1, 6, 10 and 12 correspond to category C. A peripheral auditory input activates category B more strongly than one in between the two categories. Although the network now exhibits two categories after having heard 30000 tokens, the categories have not yet been 'named' by the script; therefore, the 'child' now posits the existence and names (A, B, C, D) of the categories. To this purpose, one 'paces' through AudF in the sense that activity is spread from AudF to SF as illustrated above 32 times, one for each AudF node. The script for a Praat button that allows to pace through AudF is provided in 8.5.4 in the appendix. SF nodes that have an activation higher than 0.5 at the same time are stored as being one category, and redundant categories are filtered (e.g. a token in between two categories can activate both categories above 0.5 at a time, which is then logically considered by the script to be a distinct pattern of its own, although it is of course not). Then, the clamped AudF nodes are unclamped and the unclamped SF nodes are clamped, all activities in the network are set to 0, and each category is then successively activated by setting its SF nodes at an activity of 1. Looping through AudF, it is calculated for each category whether most of its activation at AudF is located at the first (A), second (B), third (C) or fourth (D) fourth of the continuum. The categories are then assigned a corresponding name (e.g. *Category1B* for B in a BC language, *Category2B* for B in an ABC language). It is of course not realistic for children to 'suddenly realise' which are the categories rather than gradually realising which they are during the learning period, but for reasons of computational speed, the faster and simpler procedure of category guessing used here was preferred to using a stepwise incremented counter of the probability for a certain category to exist.

After the categories have been named, their respective relative frequencies (i.e. their functional loads) are calculated by the script. To this purpose, 10000 tokens are randomly drawn from the original input distribution, and are categorised as one of the categories.

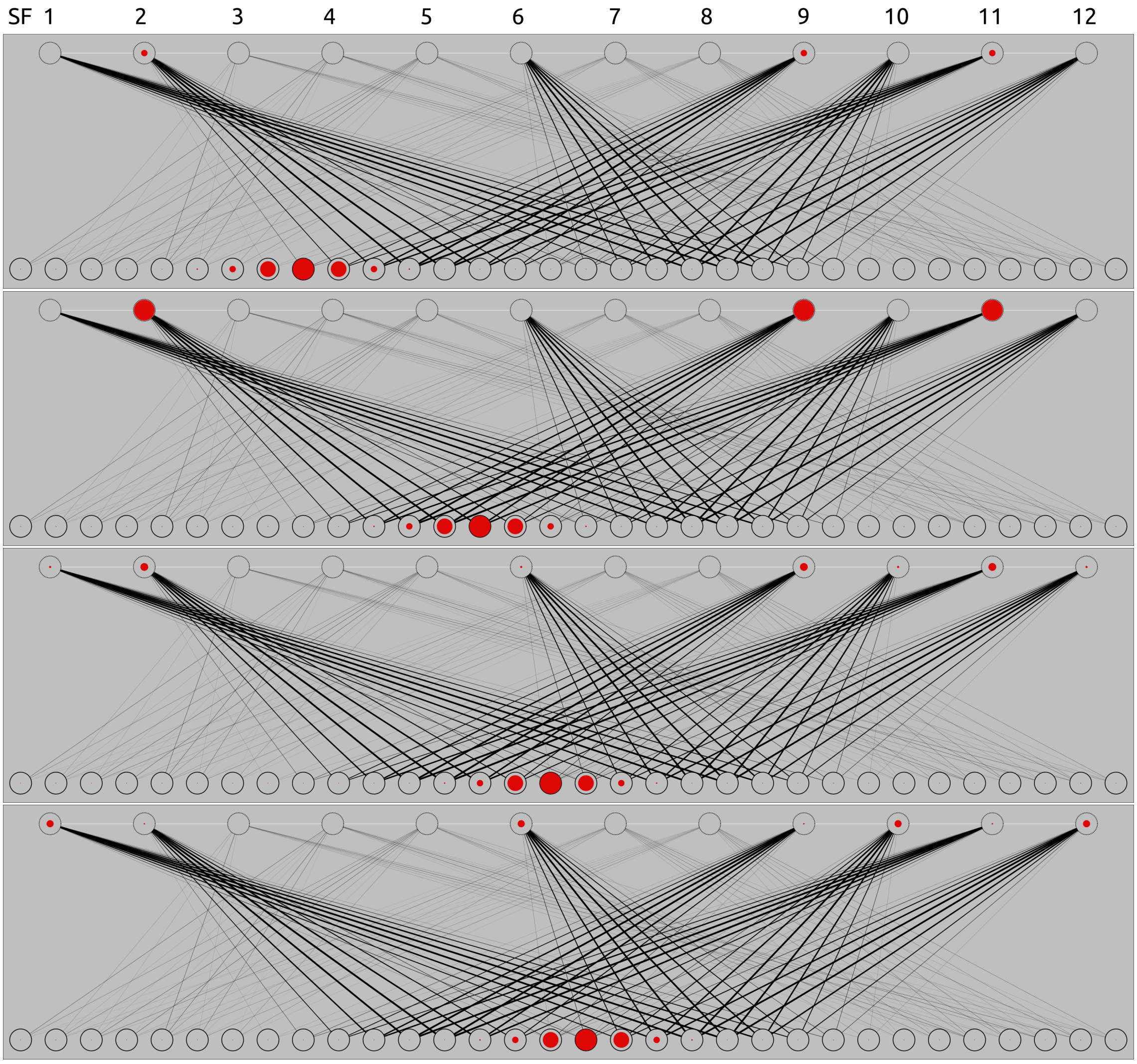


Figure XC: *Activation patterns in a BC language after learning.*

This allows us to obtain the functional load of each category, which will be slightly different at every generation (since 10000 tokens are selected randomly out of 30000), but will remain stable since there is no reason for it to increase or decrease (e.g. 53.27%, then 51.68%, then 51.97% etc.). To test whether the intended SF category of the adult distribution corresponds to the category perceived by the network, i.e. if my algorithm for categorisation works as it should, I compared both and obtained an overall error of 0.0005, i.e. 99.95% accuracy (for the BC language). The weights are also updated at the same time as categorisation occurs, so that the network will have been trained with 40000 tokens in total.

D. Production and constraint reranking

The child now starts to speak and attempts to pronounce the sibilants of her/his mother tongue 30000 times, selecting each time a native category randomly with a probability equal to its functional load. In order to speak, the child makes use of two devices: an

OTGrammar object and two articulatory nodes (ArtF). The token randomly selected to be pronounced will first be evaluated as a UF by the OTGrammar and, if its SF output is a native sibilant, the network pronounces it by switching on the nodes of the SF category and selecting a pronounced value on the continuum randomly, based on the activities of the AudF nodes as probabilities. The entire procedure is illustrated in fig. XCI.

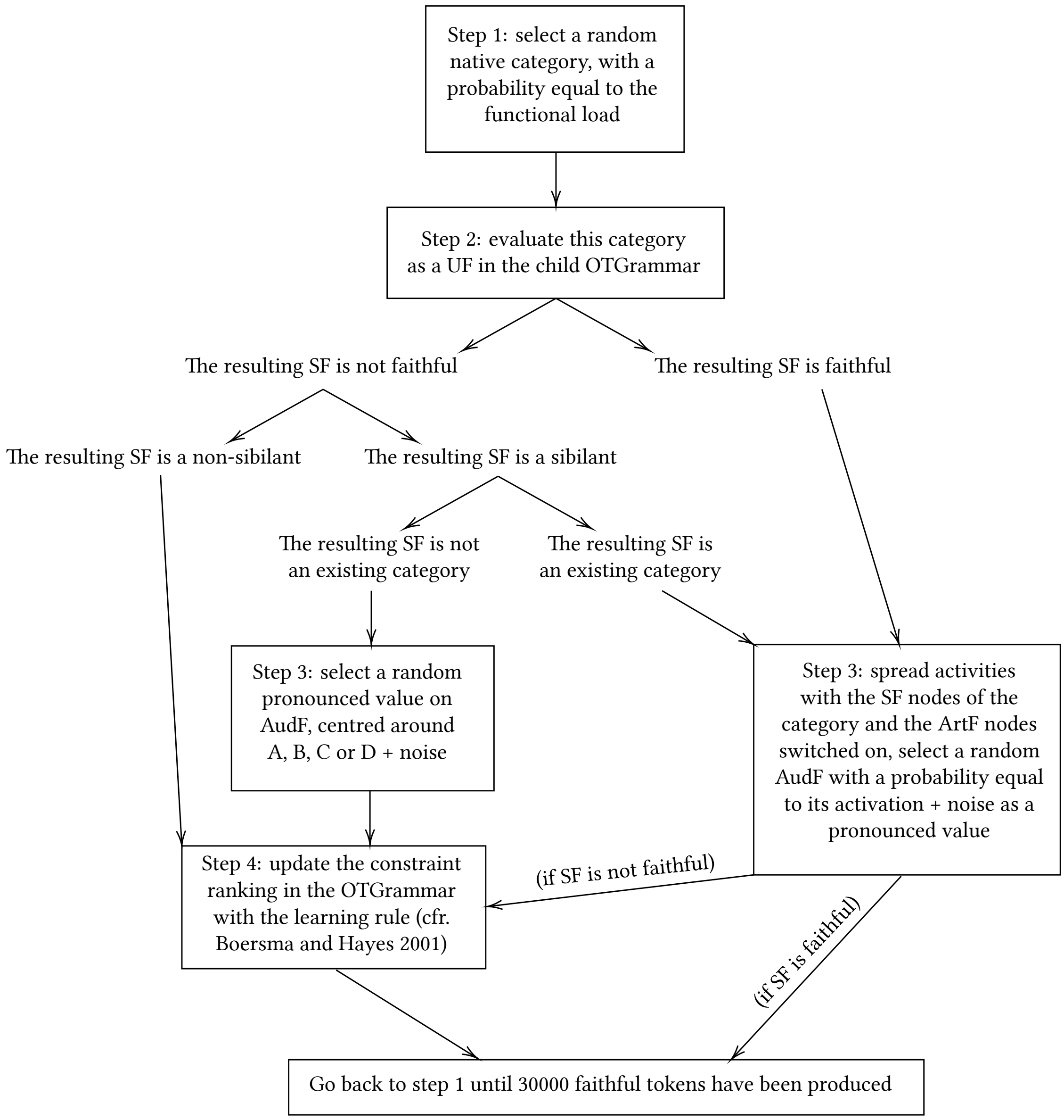


Figure XCI: *Procedure for the production of 30000 sibilant tokens by the child.*

The OTGrammar that the child uses has all sibilant markedness constraints ranked higher (at 120) than the faithfulness constraint F.IDENT(SIB) (100), something which is thought in the literature to reflect the initial state of child grammars (see e.g. Smolensky 1996: 723; van Oostendorp 2005*b*: 33). Each time a category is selected for production, this category is evaluated as a UF in the OTGrammar; at first, only non-sibilants are produced because of the high ranking of the markedness constraint M.SIB above F.IDENT(SIB). However, when the UF and SF are different (i.e. when the mapping is non-faithful), the child corrects this error by lowering the ranking of the constraints that penalise the faithful ('correct') candidate by 0.1, and by raising the ranking of the constraints that penalise the wrongly elected candidate by 0.1, after violation marks present for both candidates have been cancelled out (according to the learning rule presented in Boersma and Hayes 2001: 52f.). The entire production process, from probabilistic category selection to constraint reranking (in case UF and SF are different), is repeated until 30000 faithful tokens are produced by the child, which are the tokens that will be perceived by the next generation.[4] Fig. XCII shows the evolution of the ranking values of the constraints in the 3s language as a function of the number of tokens produced.

As a result of this gradual learning procedure (cfr. the Gradual Learning Algorithms in Boersma 1997*a*; Boersma and Hayes 2001; Boersma and Hamann 2008: 236), the grammars of the languages come to evaluate and select optimal candidates precisely as their counterpart in the OT typology presented here, with matching constraint rankings.[5] One observes that as the ranking values of the constraints change, the nature of the tokens pronounced by the child changes and matches what is cross-linguistically attested in terms of order of acquisition. Fig. XCIII and XCIV represent the overall number of tokens produced chronologically by the child in a 3s. One observes mainly 1) that the first tokens produced are non-sibilants, a short-lived stopping phase; 2) that mismatches become increasingly rare with time; 3) that the least marked sibilant, B, is acquired earliest, and 4) that the order of acquisition is sensitive not only to markedness but also to statistical frequency (C is by chance slightly more frequent (2.29%) than A in the output distribution of the previous generation, and thus acquired slightly earlier). The resulting production pattern of the child in this simulation accurately parallels the attested acquisition pattern: young children produce stops instead of sibilants at first, something which is known as the stopping phase (Barlow and Gierut 1999: 1487f.); then, any category is realised as B for some period, until the non-faithful mapping A/C > B reduces in frequency and the number of mismatches (i.e. of non-faithful mappings) starts to stagnate. As expected, with the relevant constraints being reranked to correct the mismatches, the accuracy of the child improves at each error. Both the effects of markedness (B acquired first) and of statistical frequency (C acquired earlier than A because of the

[4] By doing so, mispronunciations are removed from the input distribution that the next generation will learn. In reality, children also hear the errors made by other children, but these errors are diluted in the extremely large number of tokens heard in a human life. If I allow the next generation to hear the mispronunciations also, a statistical bias towards B arises: B being by far the most frequent mispronunciation, its relative frequency increases in every generation, until all other categories eventually disappear. This is a problem with 30000 tokens, but not with the zillions of tokens perceived in a life. Also, real children do not aim to reproduce precisely the relative frequencies of the phonemes, unlike this script.

[5] Crucially *and* non-crucially ranked markedness constraints remain at 120, if they do not penalise the categories of the language (see LXV - LXVII). This means that in the B language, the ranking value stays at 120 for all constraints except F.IDENT(SIB) at +/- 110 and M.SIB at +/- 105; in the AB language, M.POSTCD(SIB) and M.POSTD(SIB) stay at 120, with F.IDENT(SIB) at +/- 115 and M.PERIPH(SIB) slightly above M.SIB at +/- 110. The BC language has the same ranking as AB but with M.PERIPH(SIB) and M.POSTCD(SIB) swapped. The ranking in the AC language is the same as that of ABC: M.POSTD(SIB) at 120 (this time crucially), F.IDENT(SIB) at +/- 118, M.PERIPH(SIB) and M.POSTCD(SIB) at +/- 112 and M.SIB at +/- 109. Due to the absence of B, M.PERIPH(SIB) and M.POSTCD(SIB) end up only slightly lower in AC compared to ABC. The ABCD language has F.IDENT(SIB) at +/- 118 above the other constraints, which are in between +/- 108 and +/- 112.

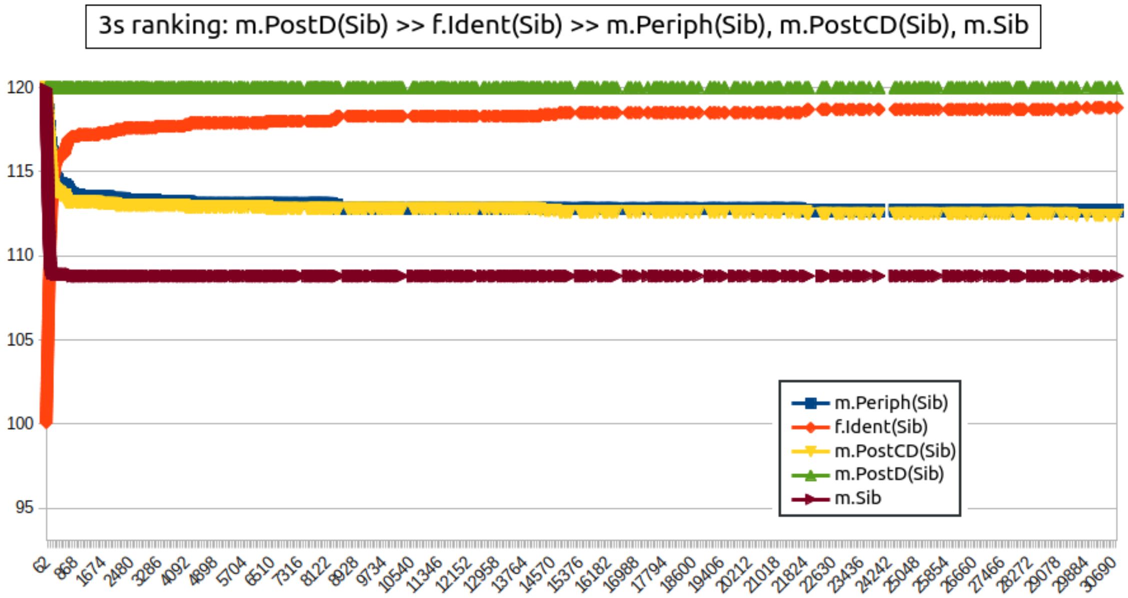


Figure XCII: *Evolution in a 3s inventory of the ranking values of the 5 constraints of the basic sibilant typology, as a function of the number of tokens produced (chronologically).*

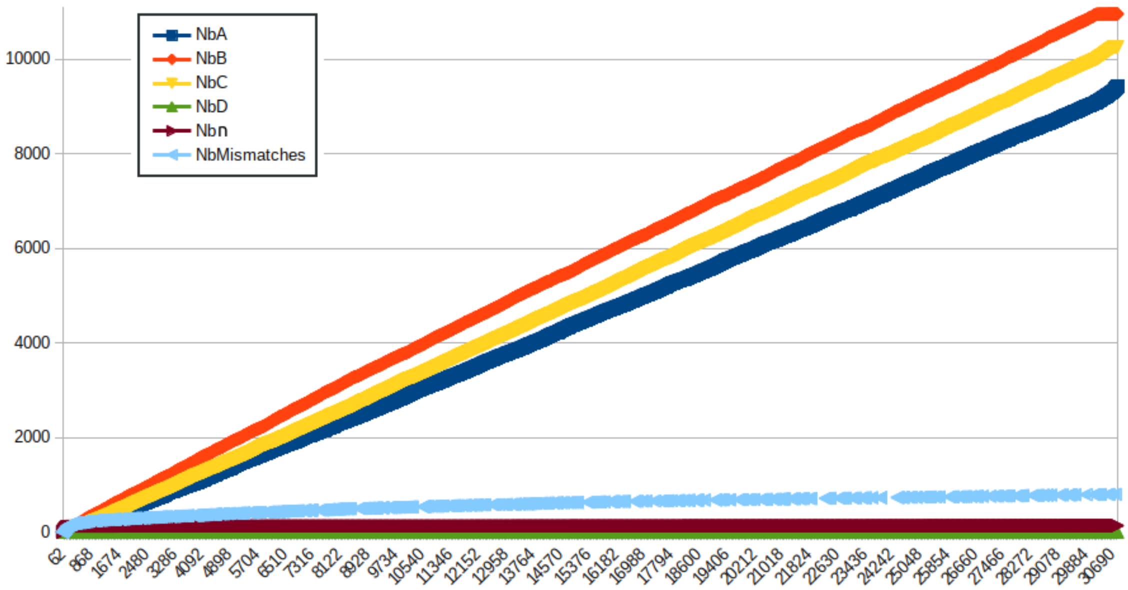


Figure XCIII: *Evolution of the number of tokens produced (non-sibilant, A, B, C or D) in a 3s language, as a function of the total number of tokens produced (chronologically).*

2.29%) on order of acquisition witnessed in this simulation correspond to attested effects in real-life acquisition (Li and Munson 2016). Indeed, Li and Munson (2016) showed that both markedness and language-specific frequency exert an influence on the acquisition of standard Chinese /ɕ, ṣ s̠/, with B (/ɕ/) acquired earlier than the statistically frequent C (/ṣ/), itself earlier than the rarer A (/s̠/). Noteworthily also, the grammar of a child acquiring a well-dispersed AC language is that of the ABC language, with the only difference being that it does not have any underlying /B/ in the lexicon; this corroborates

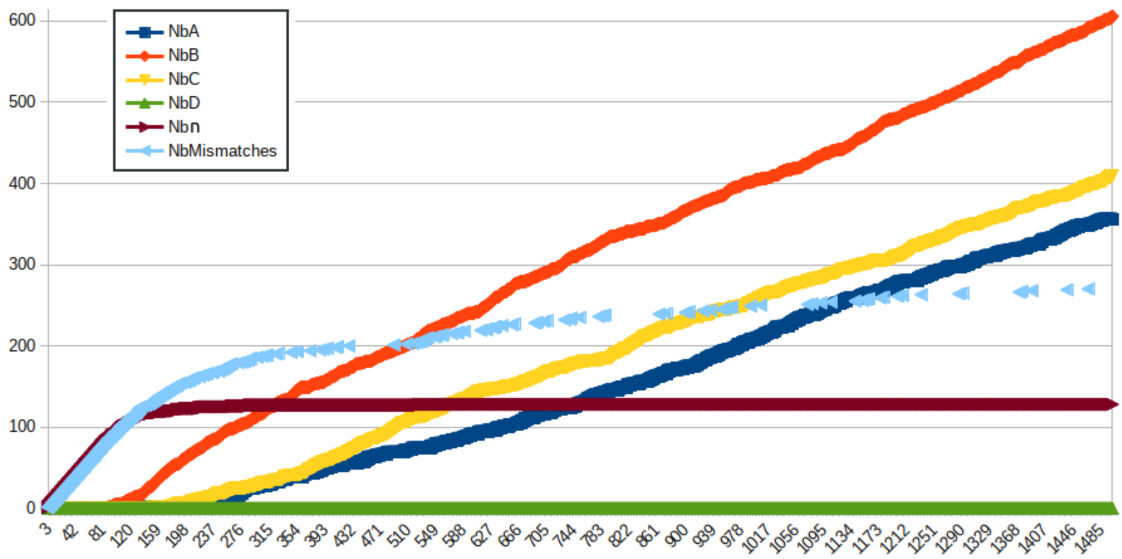


Figure XCIV: *Fig. XCIII, zoomed in on the first 1500 tokens.*

what was observed *a priori* in 5.2.2.6.

Whenever the underlying sibilant selected to be pronounced is a native sibilant, be it the result of a mispronunciation or not, the neural network is used to select a pronounced value. To this purpose, an articulatory (ArtF) node for centralisation and one for posterior markedness have been created, with negative (i.e. inhibitory) connections towards each AudF node, similarly to those in Boersma et al. (2020). To capture the markedness of peripheral (formulas (1)-(2)) and posterior (formulas (3)-(4)) sibilants, the formula of the strength of the ArtF connections is as follows:

$$d_n = -\left|\frac{n - \frac{audf.NbOfNodes-1}{2} - 1}{audf.NbOfNodes - 1}\right| \tag{1}$$

$$c_n = 0.25 \cdot d_n^2 - 0.8 \tag{2}$$

$$d_n = -\frac{n - \frac{audf.NbOfNodes-1}{4} - 1}{audf.NbOfNodes - 1} \tag{3}$$

$$c_n = 0.25 \cdot d_n^2 \tag{4}$$

where c_n is the strength of the connection towards AudF node number n, $audf.NbOfNodes - 1$ the extent of the AudF continuum, and 0.25 a coefficient to make the differences in inhibition between different AudF nodes smaller.[6] Squaring d_n serves to assign increasingly inhibited weights in a parabolic shape, similar to the ArtF node in Seinhorst (2012) and Boersma et al. (2020: 150). The -1 outside the fraction serves to centre the formula around 16.5 rather than 15.5 (for centralisation) or 8.75 rather than 9.75 (for posterior markedness), since the continuum ranges from 1 to 32 rather than from 0 to 31. The division by 2 and 4, respectively, refers to the least inhibited point of the continuum, which is the middle point of the continuum for centralisation and one fourth of the continuum for posterior markedness.[7] The formula of the posterior markedness node differs from that of centralisation to the extent that the least marked point is located further left (1/4 rather than 1/2), that the left side of the inhibition (at A) is cut out, and that having the absolute value ($|...|$) is no longer required because of this (there are no positive values if the function is cut at $y = 0$). To the total ArtF inhibition on the entire continuum, -0.8 is furthermore added by adding these to the formula of the centralisation node (since it covers the entire continuum). This means that at the centre and one fourth of the continuum, the total inhibition will be -0.815625, at the centre of B -0.8078125, at the centre of A -0.83515625 and C -0.8390625, at the centre of D -0.9328125. The two inhibition functions, together with their sum, are shown in fig. XCV. The inhibition within the domain of each category (A, B, C and D) can be visualised

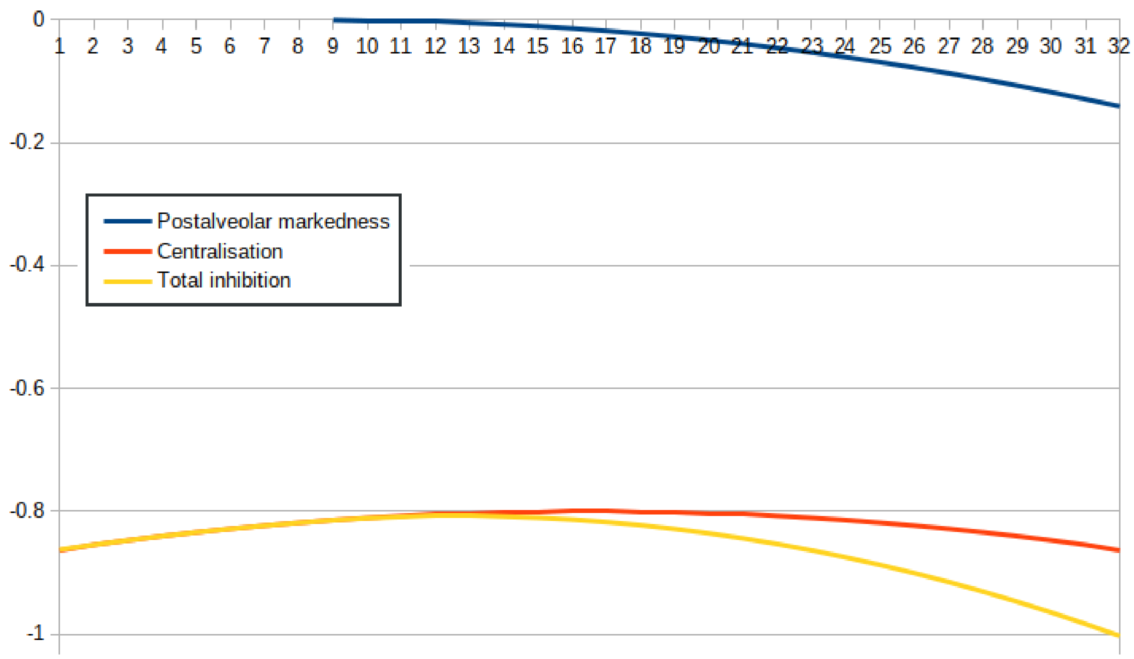


Figure XCV: *Strength of the inhibitory connections between ArtF and AudF, as a function of AudF node number.*

[6] If this value is too high, auditory dispersion cannot prevent all categories to eventually merge to B.

[7] Posterior markedness starts at B rather than at the middle of the continuum because since it is parabolic, its values around B are extremely small and do not make B marked for posterior markedness (much the opposite: the inhibition with this formula is weakest precisely at B and has precisely the same value at 1/4 and 1/2 of the continuum). With this formula starting at 1/4, the centre of C is slightly more marked than that of A. If posterior markedness starts at 1/2 of the continuum instead, A ends up being more marked than C, precisely the opposite of the desired posterior markedness effect.

with more precision as in fig. XCVI, which depicts the total inhibition for the 8 nodes of each category from least to most inhibited.

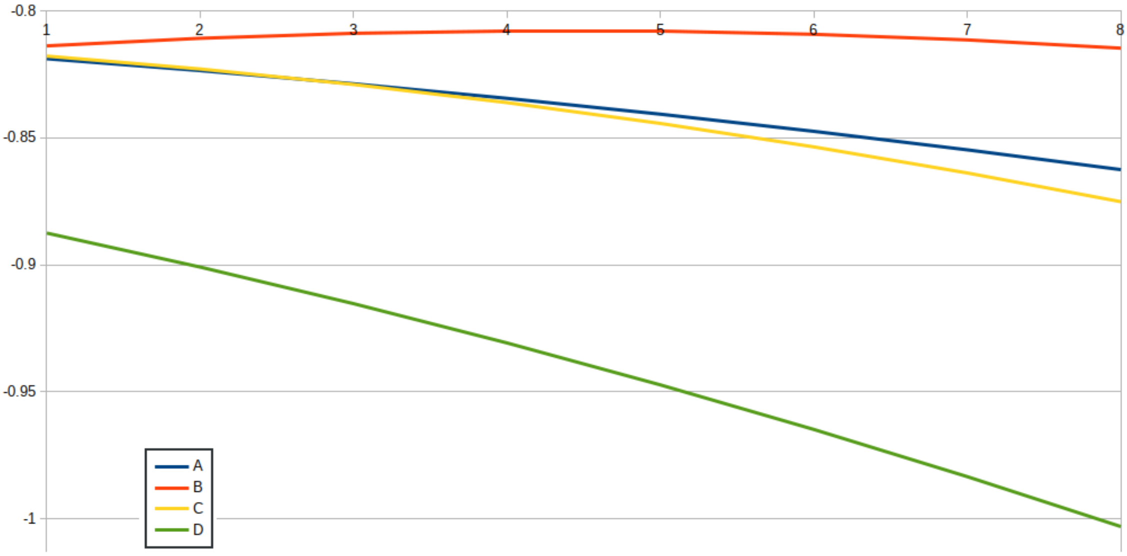


Figure XCVI: *Strength of the inhibitory connections between ArtF and AudF, per category (A, B, C and D). The least inhibited part of the curve is situated on the left for each category, i.e. A and B have been reverted. The right on this figure therefore corresponds to the right (i.e. more posterior) on the AudF continuum for C and D, but the right on this figure corresponds to the left (i.e. more anterior) on the AudF continuum for A and B.*

The network, which is ready to speak, is depicted in fig. XCVII. The SF nodes for the

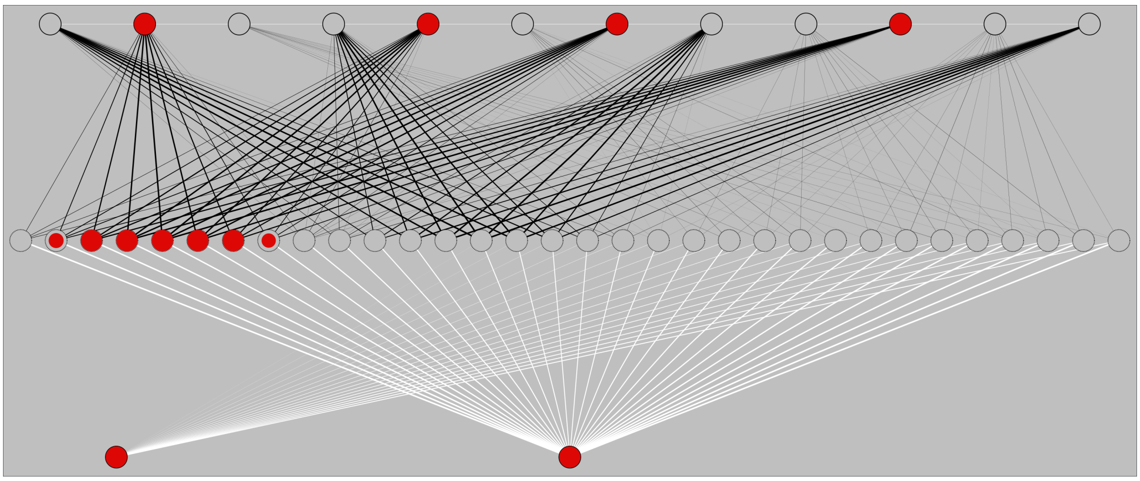


Figure XCVII: *Neural network of a language with two sibilants (A and B), producing the category A.*

native category to be produced (in fig. XCVII: A) are switched on, and the ArtF nodes are set at an activity of 1. In case the existing categories have a different number of SF nodes per category (which is not the case in this example: 4 A and 4 B), the activation of the SF nodes of the category with the lowest number of SF nodes (e.g. 3, if A had 3 and B had

4) are 1 for each SF node and those of the SF nodes of other categories are the number of the nodes of the category with the lowest number of SF nodes (e.g. 3) divided by its number of nodes (e.g. 4), so e.g. 3/4 = 0.75 per SF node. This ensures that each category in a network is activated with the same strength (3 x 1 = 4 x 0.75). The SF and ArtF nodes are clamped while the AudF nodes are unclamped, to allow activity to spread from SF and ArtF to AudF. After spreading activities 500 times, the network comes to look as depicted in fig. XCVII. It is important to observe that the combination of excitation (from SF) and inhibition (from ArtF) prevents the nodes with weaker excitatory connections to be activated, and simultaneously allows the nodes with stronger excitatory connections to be activated, despite the inhibition. For example, the first AudF node has no activation despite its connection to the SF nodes of category A: the inhibitory connection towards this extremely peripheral node is simply stronger than its excitatory connection to the SF nodes. In order to choose what pronounced value on the AudF continuum will be produced, a random AudF node and a random uniform value between 0 and 1 are chosen (e.g. node 2 and 0.61186982). If the activation of the random node is larger than the random uniform value (which, in this case, is true), that node will be that of the pronounced value. The actual pronounced value is then selected randomly out of a Gaussian probability distribution with that node as a centre/mean and a standard deviation of $\frac{0.2}{(audf.NbOfNodes-1)} \cdot 100$, which equals 0.64% with an AudF continuum of 31 nodes. If the activation of the random node is not larger than the random uniform number, both are selected randomly again.

If the category does not exist (e.g. the child has pronounced a B in a well-dispersed AC language), a pronounced value is selected randomly instead with the centre of a Gaussian probability distribution aligned at the centre of the category (12.5% for A, 37.5% for B, 62.5% for C and 87.5% for D), with a standard deviation of $\frac{0.2}{(audf.NbOfNodes-1)} \cdot 100$, i.e. the same as the one described just above.

After having pronounced 30000 faithful tokens, the script saves the new distributions and a copy of the distribution as well as the network, so that the next generation might learn from the input distribution just saved.[8]

5.4.1.5 The stable acquisition of predicted inventories

Now, the five inventories with sibilants predicted by the basic sibilant typology (B, AB, BC, ABC and ABCD) will be learnt by 150 generations of learners, i.e. *ChildOTandNN.praat* will be run 150 times for each language. This allows us to provide an estimate of the learnability and diachronic stability of each inventory.

The 1s language with a single sibilant category at **B** was learnt without any change, as

[8] If there is an endless loop in the script, which happens e.g. when a category has excitatory connections too weak to propagate any activity on AudF, so that the script will look forever for an AudF whose activity is larger than the random uniform value, the loop is forcefully exited and the distribution *not* saved. The next generation will then try again to learn the same distribution. Such cases occur relatively rarely (ranging from e.g. 7.3% of the time for 150 generations of the B language, or 4% of the time for 150 generations of BC, to 0.6% for 150 generations of ABC).

fig. XCVIII shows. The algorithm described in C. that ‘guesses’ which are the categories

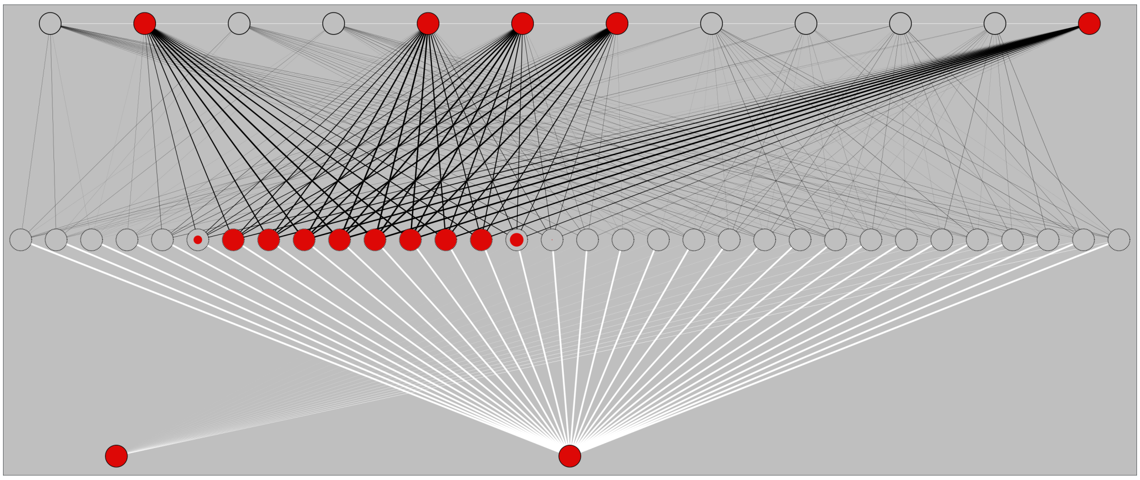


Figure XCVIII: *Activation in the neural network of the B language at generation 150. It can be seen that despite the large number of generations, practically nothing has changed in the network, showing the stability of the B pattern.*

and assigns them a name, during the 150 generations of acquisition of the 1s, did not ever recognise any other category than B. The 1s can thus be considered a very stable pattern.

Interestingly, some generations of the 1s have been exposed to such a wide input distribution, with a variation ranging from the end of A to the beginning of C, that the 1s language is reanalysed as one having two barely distinct B categories, one slightly more anterior and one slightly more posterior. This would correspond in IPA terms with a /s̠ - ɕ/ inventory. The next generation, however, reinterprets these categories as one wide B category, merging them again to /s̠/. This phenomenon, in fact, reflects what is cross-linguistically attested in some languages with /s̠ - (t͡)ɕ/ (e.g. Ancient Greek and Late Proto-Finnic, see 3.3.1.2), through the depalatalisation of alveopalatals (which combines the reduction of articulatory effort by not requiring to raise the tongue dorsum, and in this case the perceptual confusion between alveolar retracted and alveopalatal, cfr. 3.1.5.1). Such a language is depicted in XCIX, with the activation patterns for the two categories superimposed with an opacity of 50% each. The reason for which the network learns a very wide B category as two categories instead of one is the fact that the parameter of the standard deviation of perceived tokens on AudF, set to 0.05 here and 0.04 in Boersma et al. (2020) (cfr. B.), allows only to learn a limited range on AudF as one single category. Above the range threshold, all of the SF nodes within one and the same group of SF nodes cannot be activated by the activation spread when a token activates some of them, resulting in distinct activation patterns.

The **AB** language was acquired stably during the first attempt until generation 52, when it became a well-dispersed AC language (although continuing to oscillate back and forth between AB and AC until being AC 6 times in a row, from generation 65 to 70). The same change occurred in the second simulation attempt, although earlier (at generation 22 already). It seems puzzling why the AB language is not acquired stably and becomes

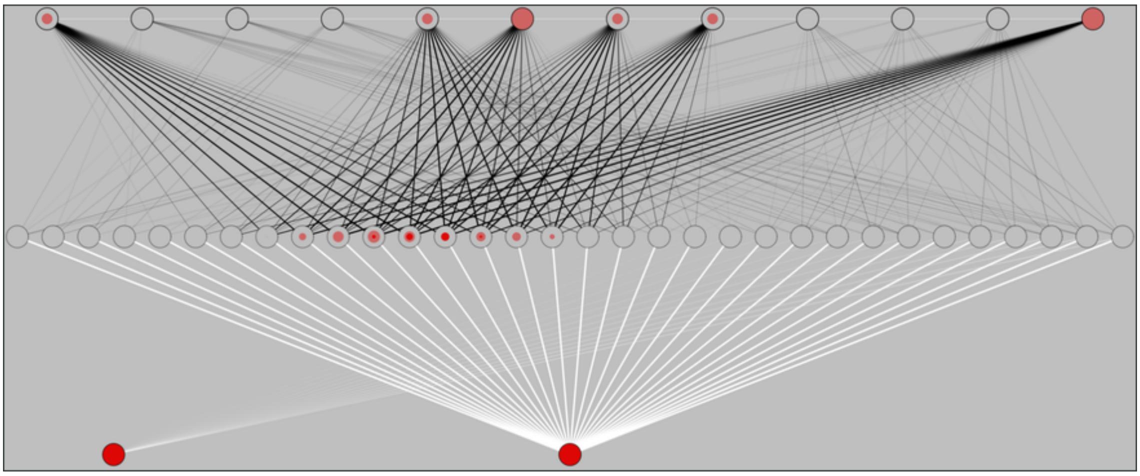


Figure XCIX: *Activation in the neural network of the B language at generation 9, with the two categories superimposed (with an opacity of 50%). The proximity of the two categories will make them merge at the next generation.*

an AC language over time, while the BC language is acquired stably (as shown below). The difference between the diachronic behaviour of AB and BC could have been due to an extremely small difference in inhibition (e.g. of 0.00102 between the last A node and the first C node), as can be seen in fig. XCVI. To test whether the different behaviours of AB and BC are due to this small difference in inhibition, I reverted the inhibitions from A to C, with the centre of B (around nodes 12 and 13) as 'mirror' point. This lead to the reverted inhibition pattern displayed in fig. C. Running the simulations for AB and BC

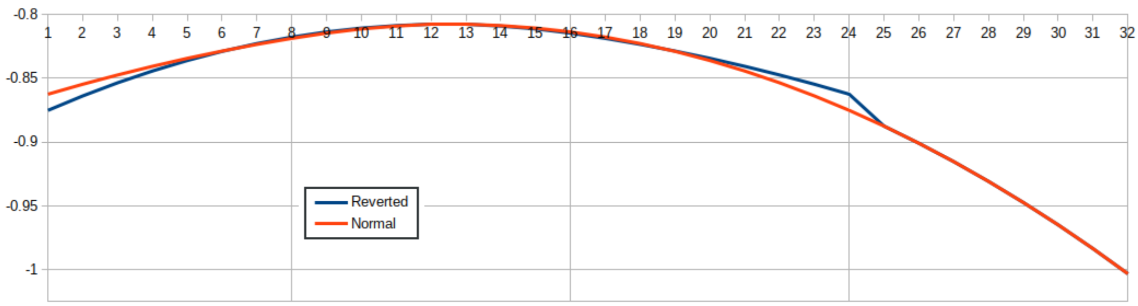


Figure C: *Reverted inhibition pattern from A to C. It can be observed that D is unchanged and that the inversion at B is barely visible.*

again with this inhibition pattern, the same as previously resulted: AB becomes AC (at generation 21) while BC remains BC. The explanation for their different behaviour might thus rather be another difference between A and C, namely that C borders rightwards with D while there is nothing leftwards from A. When auditory dispersion drives B and C apart, B can push C to overlap a bit towards D, and C can push B to overlap a bit towards A. However, in the AB language, B cannot repel A towards the left while A pushes B towards the right, i.e. towards C. Auditory dispersion would thus have a stronger effect in the AB language because A is 'constrained in a corner' of auditory space, and thus necessarily provokes an initial movement towards the back that might push B to C.

It is therefore predicted that attested AB languages like the dialect of Évolène (see 3.2.2.3) are diachronically unstable, something which is potentially witnessed by the rarity of this pattern and the fact that it changes promptly in contact with other varieties (e.g. with Swiss standard French). It nevertheless remains a possible inventory, even if prompt to transition into another.

The **BC** language was acquired stably and does not look any different between e.g. generation 3 and generation 150, as illustrated in CI. This detail of the activation pattern for category C reveals to what extent practically nothing has changed over time, showing the stability of the BC language. When learning the BC language for 150 generations, no

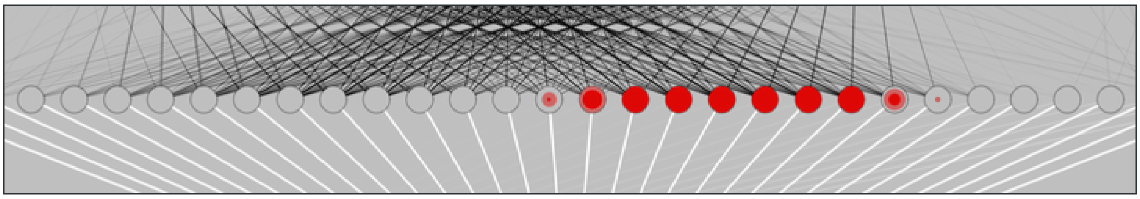


Figure CI: *Superimposed images of the activation for category C in the neural network of the BC language at generation 3 and 150 (with 50% opacity).*

reanalysis of the inventory with any A or D category occurred, i.e. it remained precisely BC for 150 entire generations.

The **3s** language was learnt stably from generation 1 to 150, as witnessed by fig. CII. This is corroborated by the category identification algorithm, which recognised exclusively A, B and C across all generations. The **4s** language was learnt somewhat stably,

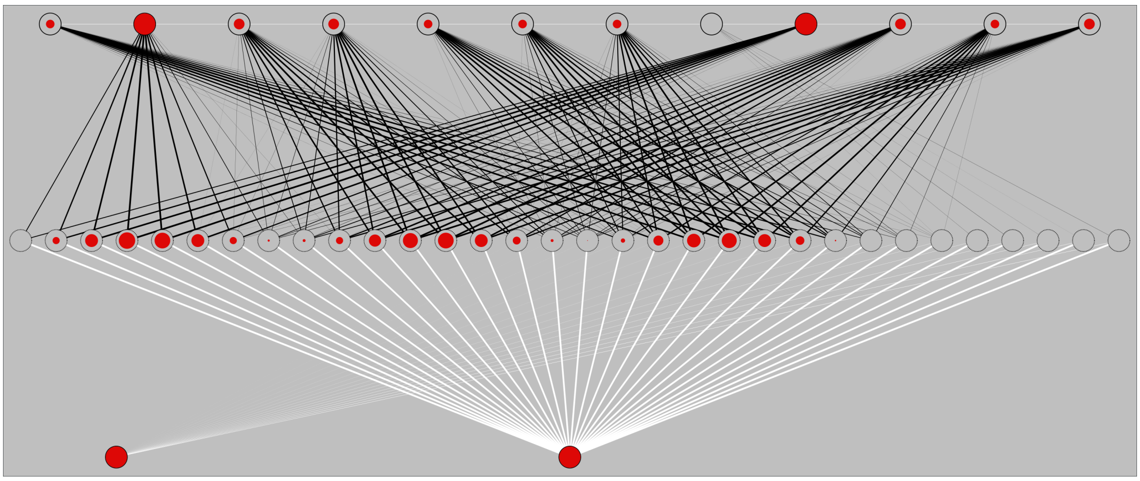


Figure CII: *Activation for all three categories in the neural network of the ABC language, at generation 150.*

although becoming a 3s before the 150th generation. In one simulation, B merged to C already at generation 36, followed by the expected gradual shift from D to C, i.e. becoming the attested 3s. In another simulation however, the 4s remained stable until generation 124, with the merger of D to C. It can thus remain stable for a certain time, but not indefinitely.

To summarise, changes (dispersion or mergers) do happen within the predicted inventories, but only in the AB and ABCD language. In both cases, the edges of the auditory continuum seem to play a crucial role, pushing neighbouring categories inwards without being pushed back. In simulations which do not abruptly and unrealistically cut possible frequencies lower or higher than 1789 and 16177 Hertz respectively, this 'edge effect' could not be active anymore.[9] Also, natural languages might disperse more efficiently than the languages simulated here by using distributedness distinctions as additional acoustic information, something which was not simulated here.

In unattested inventories (e.g. BCD or D), predictable changes happen unavoidably in any simulation and they become attested/predicted inventories, as shown in what follows.

5.4.1.6 The automatic conversion of unattested inventories

What I call here 'unattested' inventories are inventories that were not predicted by the basic OT sibilant typology and are not attested in natural languages, but would still be possible in principle. There are 16 logically possible inventories, considering the maximum of 4 different categories (nihil, A, B, C, D, AB, BC, CD, AC, BD, AD, ABC, ABD, ACD, BCD, ABCD). Removing the 6 patterns already predicted by the OT typology from the 16 logically possible patterns, one obtains the following inventories: A, C, D, CD, AC (which however *is* stable although unpredicted, as will be demonstrated in the next subsubsection), BD, AD, ABD, ACD, BCD. With the exception of the AC pattern, these 10 inventories are all set excessively towards the left of the ideal pattern (A), excessively towards the right (C, D, CD, BD, ABD, ACD, BCD) or overdispersed, i.e. set excessively towards both left and right (AD).

The input distribution for each of these languages was generated by means of a Praat script (Boersma and Weenink 2019) that generates random values on the 0%-100% continuum with Gaussian-shaped probabilities (with the function *randomGauss*). The probabilities were equally distributed among all categories. Out of the unattested 1s languages (A, C and D), the C language was the first to become B in the simulation (at generation 61). The 1s with D became one with C already at generation 22, but continued much more slowly towards B, which it only became at generation 217; A became B at generation 181. In a 1s, the rapidity of change is expectably proportional to the strength of the inhibition differences: the nearer a category is to B, the slower it moves because the smaller the inhibition differences become (compare XCV). The CD, BD and AD languages also all became an attested inventory after a certain number of generations. The CD language became a BD language at generation 19 already, moving further to become BC at generation 65. The BD pattern became BC at generation 36, which

[9] In reality, the auditory continuum is not abruptly cut at some point, but the auditory values that can be perceived and pronounced are both gradually fading towards the extremities and speaker-specific, in the sense that individuals reach different maxima and minima in perception and production. Simulations in which A and D can 'overflow' to some restricted and gradual extent towards further extremities of AudF would thus perhaps allow AB and ABCD to be stable patterns and certainly be more biologically plausible, unlike when values below 0% and above 100% of the continuum are clipped away.

approximatively corresponds to the 46 generations that CD required to move from BD to BC. Very interestingly, AD experienced a slightly different path of change: the two overdispersed categories initially moved towards B and C respectively, with D reaching C first at generation 56 (as expected because of the stronger inhibition differences) and A then reaching B at generation 80, but the BC pattern soon became an AC pattern again, stabilising further as an AC language. It then remained AC for 27 generations in a row, and I stopped the simulation at the 27th generation since it seemed not to be changing anymore. To illustrate the path of change from AD to AC, fig. CIII details which categories were identified by the algorithm for each generation. My explanation

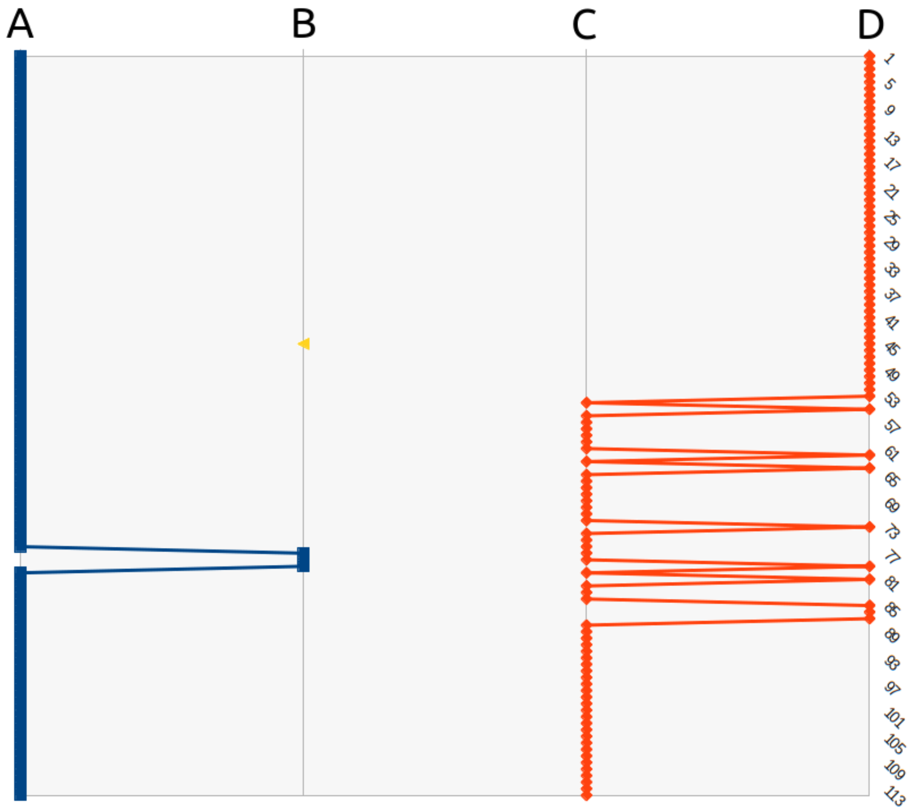


Figure CIII: *Evolution of the category identification in a 2s with A and D, as a function of generation number (from top to bottom). The yellow triangle is a short-lived third category that has been perceived by the network at generation 47 and then disappeared again.*

for this behaviour would be that although A and D started at equal distances from the centre of the continuum (precisely 37.5%), the stronger inhibition differences from D to C made it reach C faster (which is corroborated by the fact that e.g. at generation 55, the distance between the edge and D was 2 against 1 for A, and by the fact that the 1s with D became C much earlier than the 1s with A became B). When the original A reaches the border with B, the original D has already become C and pushes A back forward. It thus seems that although the BC pattern is stable across generations, as shown in the previous subsubsection, it becomes an AC language as soon as the location of B and C are not necessarily at the cardinal 37.5% and 62.5%, respectively, but with B slightly more leftwards.

Turning now to the unattested 3s, we also observe that they all become the ABC language. The ABD pattern becomes ABC at generation 58, and then does not budge an inch, staying at ABC for the next 62 generations (until the end of the simulation). The ACD pattern becomes ABD at generation 59, then ABC at generation 90, remaining as such for 25 consecutive generations. The BCD language moves slightly more slowly: it becomes ACD at generation 95, ABD at generation 99, with A and B oscillating back and forth until D becomes C at generation 140, thereby reaching the ABC pattern.

The unattested 1s, 2s and 3s have thus all become an attested inventory with the same number of contrasts. Furthermore, the behaviour of the AD language has shed light on how the BC language also can become AC. The next subsubsection goes more deeply into the motivations for the emergence of the well-dispersed AC language and its stability.

5.4.1.7 The well-dispersed 2s and its implications for the typology II

As has been shown above, the AC language can emerge spontaneously in a less-dispersed 2s (systematically in AB, although with a *caveat* wrt. the edge as a possible unrealistic trigger for this; and occasionally in BC when the categories are slightly leftwards rather than the 'canonical' 37.5% and 62.5%). I would like to make two additional observations here: 1) the AC language also emerges when the B and C categories are sufficiently wide or left-skewed or left-situated, and 2) the AC language is in itself a stable pattern.

When simulating a BC language centred as usually around 37.5% and 62.5%, but with the standard deviation of a 1s (i.e. 10%) for both categories, B and C first repel each other, but with C predictably moving less quickly rightwards, since the inhibition is stronger towards the right. The language then oscillates between AB and BC, finally becoming AC and staying as it is for 39 consecutive generations. Fig. CIV illustrates the

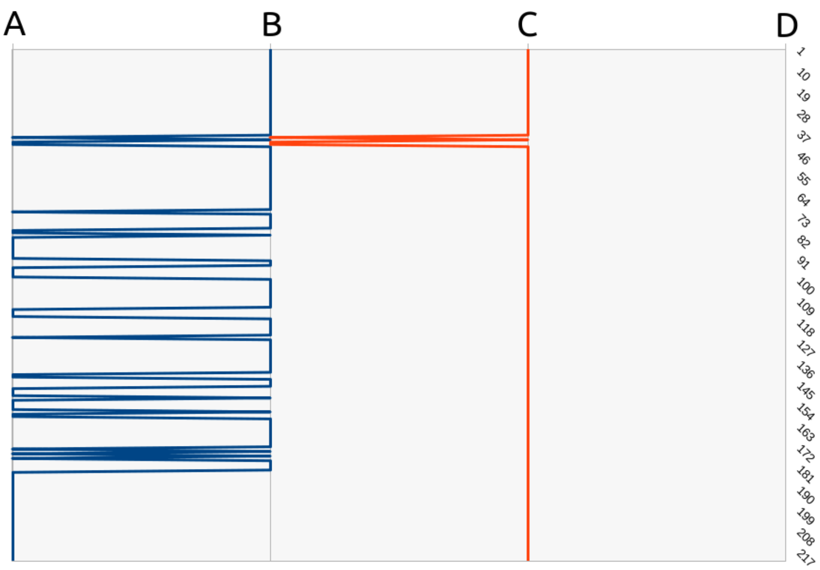


Figure CIV: *Evolution of the category identification in a 2s with two wide categories B and C (standard deviation: 10%), as a function of generation number (from top to bottom).*

evolution of the inventory as perceived by the category identification algorithm. An AC language can thus also emerge from a BC distribution, as soon as phonetic variation in the input distribution is large to a sufficient extent for the change to happen. I tried a simulation with a narrower standard deviation for B and C (namely the usual 7.07 for a less-dispersed 2s i.e. $\frac{10}{\sqrt{2}}$), but with an additional small category at 25% of the auditory continuum with a probability of $\frac{1}{6}$, which provokes the fact that the left tail of the B category will be wider than the right tail (see fig. CV). It results that even in such a case,

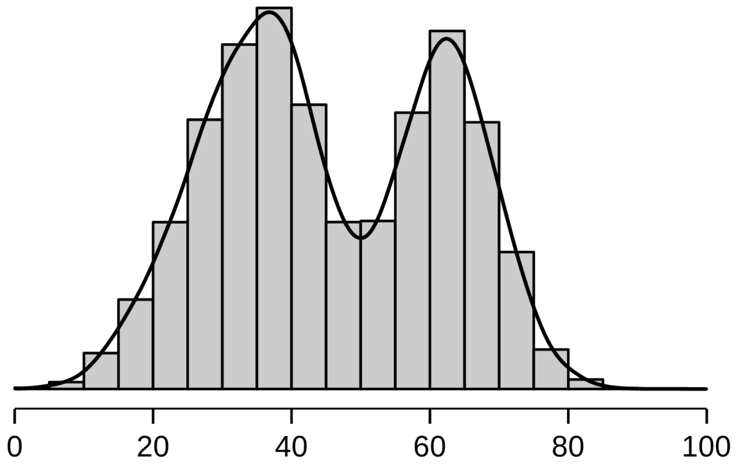


Figure CV: *Leftwards-modified input distribution with B and C (JASP Team 2021).*

some reanalysis to AC occurs, although less frequently and as a diachronically unstable phenomenon. Such a language has been identified as BC at 130 generations overall, and as AC at 9 generations in total (i.e. 6.47%). This confirms that change from BC to AC occurs when the BC distribution is left-skewed or set slightly left from the 'canonical' 37.5% and 62.5%, but also that the wider the standard deviation, the more likely it is that AC will emerge.

Secondly, if one centres the language to A (12.5%) and C (62.5%) with a standard deviation of 10% from the start, the language is learnt and transmitted stably. Fig. CVI illustrates this stability by showing the variation in activation patterns across 150 generations. Both categories are labelled A and C by the algorithm throughout *all* generations. As the simulations with language contact reveal (see the next subsubsection), AC is also very stable and comes out as a 'winner' of most simulations (i.e. A and C are still categories after 40 generations). For example, when crossed to various extents with the AB language in the native input, the neural network after 40 generations tends to exhibit ABC patterns if AC was less frequent than AB in the original input, but exhibits purely AC patterns if it was more frequent (see table 35 in the appendix). In contact with ABC, it tends to reduce the inventory to AC as long as AC was more frequent in the original input. Furthermore, it spontaneously emerges in some simulations that do not imply an AC distribution from the start (e.g. AB with ABC or AB with BC).

The reasons for which the AC language is acquired stably are twofold: firstly, the centre of B is the least marked point of the continuum, constituting the ideal middle way between centralisation and posterior markedness, as witnessed by the behaviour of

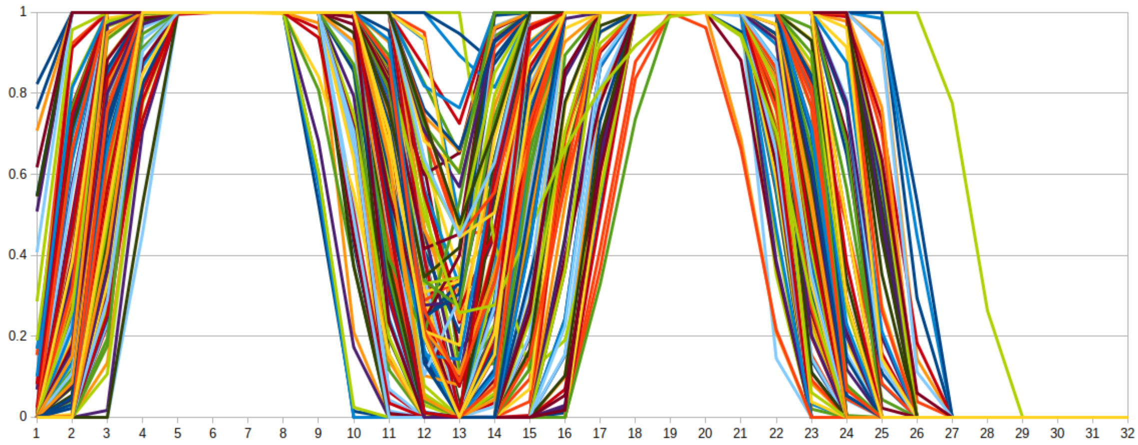


Figure CVI: *Evolution of the activation pattern in a well-dispersed 2s with A and C.*

the 1s; A and C are centred around this middle point and thus constantly strive towards that point. Secondly, the standard deviation of each category is large enough to allow them to 'meet' with each one side of their respective distribution at this centre point, which allows auditory dispersion to weaken the connections towards ambiguous and strengthen the connections towards peripheral tokens, thereby counteracting markedness. The first and the second reason counterbalancing each other, it results that nothing changes and the pattern remains stable. The AC pattern is thus not only a diachronically stable language, but it also emerges as soon as variation in the acoustic input (e.g. wide categories) is sufficient to cause a change.

This explains a crucial characteristic of the AC language: it is not predicted by the production grammar to exist, being suboptimal in markedness wrt. less-dispersed 2s; but these less-dispersed 2s are likely to become an AC language when they suffer a change because of variation (mergers, allophonic variation, language contact etc.). Once they have 'slided' into an AC language, this pattern is then stable and robust enough to remain the same across generations.

5.4.1.8 Language contact and functional load

Up to this point, homogeneous input distributions with equal functional loads have been learnt by the network. I have however also made simulations for different proportions of language contact between two sibilant inventories and different proportions of functional load within a sibilant inventory. Results are reported here for all possible combinations of 20, 40, 60 and 80% language contact intensities and functional loads (by default; the two exceptions are: in a 1s, the functional load is of course always 100%, and in a 3s, it is 20-20-60, 20-60-20, 60-20-20, 40-40-20, 40-20-40, 20-40-40 for the first, second and third sibilant respectively). Combinations with the 4s were not tested, because the number of simulations is already very high and because simulations with the 4s would expectedly lead to mergers in the 4s that are not of particularly interest. Combinations including one language twice were excluded (e.g. an AB language with a functional load of 40-60 in contact with an AB language with a functional load of 60-40), so that the two languages combined each time had to have different PoAs. This means that in total, 552 combinations were tested: 72 combinations involving the 1s, reported in table 22 here, and 480 combinations involving the three types of 2s (AB, AC and BC) as well as the 3s, incorporated into a table for combinations involving the AB language in the appendix because of its large dimensions (table 35 in 8.4). At the first generation of each combination, the child heard 30000 sibilants drawn from the two language distributions involved with probabilities equalling the contact intensity. The resulting distribution was passed on to the 39 following generations (the number of generations was limited to 40 because this required almost a month of continuous simulations on two computers). As can be

	AB		AC		BC		ABC	
	CI	FL	CI	FL	CI	FL	CI	FL
B	20-80	20-80 [AB]	20-80	20-80 [AC]	20-80	20-80 [BC]*	20-80	20-20-60 [ABC]
		40-60 [AB]		40-60 [AC]		40-60 [BC]*		20-60-20 [ABC]
		60-40 [AB]		60-40 [ABC]		60-40 [BC]*		60-20-20 [AB]*
		80-20 [AC]*		80-20 [ABC]		80-20 [BC]		40-40-20 [ABC]
	/	/	/	/	/	/		40-20-40 [BC]*
	/	/	/	/	/	/		20-40-40 [ABC]
	40-60	20-80 [AB]	40-60	20-80 [ABC]	40-60	20-80 [BC]*	40-60	20-20-60 [BC]
		40-60 [AB]		40-60 [BC]		40-60 [BC]		20-60-20 [ABC]
		60-40 [AC]*		60-40 [AB]*		60-40 [BC]		60-20-20 [ABC]
		80-20 [AC]*		80-20 [ABC]		80-20 [BC]		40-40-20 [ABC]
	/	/	/	/	/	/		40-20-40 [ABC]
	/	/	/	/	/	/		20-40-40 [ABC]
	60-40	20-80 [AB]	60-40	20-80 [ABC]	60-40	20-80 [BC]	60-40	20-20-60 [BC]
		40-60 [AB]		40-60 [ABC]		40-60 [BC]		20-60-20 [ABC]
		60-40 [AB]*		60-40 [ABC]		60-40 [BC]		60-20-20 [ABC]
		80-20 [AB]		80-20 [AB]*		80-20 [B]		40-40-20 [AB]
	/	/	/	/	/	/		40-20-40 [ABC]
	/	/	/	/	/	/		20-40-40 [BC]
	80-20	20-80 [B]	80-20	20-80 [BC]	80-20	20-80 [BC]	80-20	20-20-60 [BC]
		40-60 [B]		40-60 [BC]		40-60 [BC]		20-60-20 [B]
		60-40 [AB]		60-40 [B]		60-40 [B]		60-20-20 [AB]
		80-20 [AB]		80-20 [AB]*		80-20 [B]		40-40-20 [B]
	/	/	/	/	/	/		40-20-40 [AB]
	/	/	/	/	/	/		20-40-40 [BC]

Table 22: *Possible combinations of contact intensity and functional load with a 1s. [AB]* or [BC]* stands for an AB or BC language that is very close to being an AC language, [AC]* for an AC language that is very close to being an AB language.*

seen in table 22, it can occur that the language resulting from contact has adopted the sibilant inventory of one of the two contact languages (e.g. B + AB = AB), but it can also occur that a third sibilant inventory emerges from the contact (e.g. B + AC = ABC). One observes that attested patterns of language contact are matched in the results of these simulations: for example, the reduction of the Basque 3s to a 1s under the pressure of Castilian-dominant bilingual speakers mentioned in this dissertation is echoed in table 22, where the 80-20 contact (i.e. with clearly dominant Castilian Spanish) between a 1s and a 3s with a functional load of 40-40-20 (matching most closely the actual Basque functional loads, see Peust 2008: 108f.) yields a 1s. Another example is the combination of AB with AC that yields ABC in a majority of cases (see table 35 in the appendix), precisely what is observed by Alber and Kokkelmans (in prep.) in the contact between the Trentino and Tyrolean dialects in the border village of Salurn/Salorno (South Tyrol, Italy). Importantly, it turned out that all resulting languages were predicted to exist (i.e. B, AB, AC, BC, ABC and no other type).

5.4.2 From grammar to surface inventory

Adding the well-dispersed AC inventory modelled with Neural Networks to the basic sibilant typology, there are now 7 predicted sibilant inventories in total (sibilantless, B, AB, AC, BC, ABC, ABCD). These 7 predicted inventories are still an extremely small subset of the 84 different attested inventories in the SibInv database, something which is due to three facts: 1) each inventory with sibilants (6 out of 7) has four variants based on voicing as well as manner (fricative - affricate) distinctions, as predicted by the extended OT typology; 2) one finds many variants of inventories because of voicing and/or manner gaps (as witnessed by the attested subtypes in 3.2.2); and 3) auditory dispersion in the form of distributedness distinctions has not yet been taken into consideration. These three aspects will be handled successively in the three following subsubsections.

To account for how 7 predicted inventories become a very large number of predicted inventories, and how these many predicted inventories again boil down to 7 underlying patterns, I will describe here the mapping of what I name *grammar inventories* into *surface inventories* and vice-versa, something which will allow us to understand better how the abstract inventories in grammar and the observed phonetic realisations interact. In fact, what is called an *inventory* can correspond to one out of several concepts, ranging from the inventory predicted by OT to the inventory observed by the researcher, as detailed in fig. CVII. Everything in the model represented in CVII is emergent, except the OT grammar and the articulatory inhibition in NN (which correspond to the markedness constraints in OT). The learning child is exposed only to the surface inventory of a language, and based on this data, induces the structure of the underlying inventory. The other way around, the underlying grammar allows certain sibilants and prohibits other sibilants to exist as a phoneme, but accidental historical reasons might then also prevent some of the predicted sibilants to be a phoneme, although the underlying grammar allows them. Such a perspective is represented in the literature by e.g. Vaux and Samuels (2015: 590) and proponents of an emergent rather than grammar-based explanation for inventory structure (even if these proponents tend to see more emergence and accidentality in inventories than in this dissertation; recall 4.1.1.2). The precise phonetic way in which auditory dispersion is realised (by means of average frequency or distributedness, lip rounding as acoustic enhancement etc.) is also emergent (e.g. a consequence

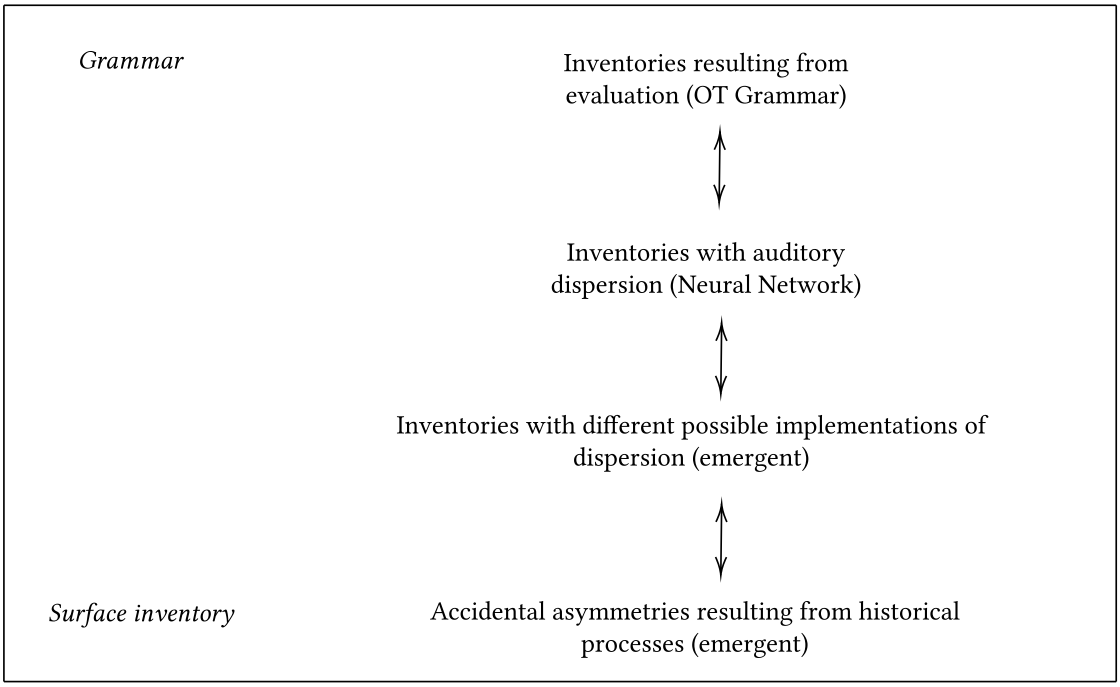


Figure CVII: *Schematic representation of the different layers involved between the OT inventory grammar and the inventory observed on the surface. Note that it is bidirectional, i.e. that children are only exposed to the surface inventory and induce the underlying grammar based on perceptual and articulatory learning.*

of the origin of the segments): for example, a well-dispersed 2s in which a palatal [c] is assibilated will have a /ɕ/ at the centre of the resulting 3s (e.g. standard Serbo-Croatian), but a 1s that phonemicises /s̪/ and a /ʃ/ from [ɕ] will have a /s̪/ at the centre of the 3s instead (e.g. Middle French). How a language can fit into this schematic representation is illustrated in CVIII, using the example of the standard German sibilant inventory. The OT grammar for standard German is that of a 3s (also historically, since Middle High German was a 3s), with the historical split and contextual redistribution of /s̪/ to /s/ and /ʃ/, which lead to a well-dispersed 2s. The absent /s̪/ in the lexicon and almost absent /s̪/ in the acoustic input provokes the induction of an ungrounded constraint *B (see 5.2.2.6) by the learning child (only under the assumption of the ungrounded constraint rather than the BiPhon-OT solution). The historical split, which occurred at the level of the neural network, lead to that change from 3s to 3s with ungrounded *B. In the other direction, the acoustic enhancement of the ABC contrast with distributedness distinctions was lost as this auditory dispersion phenomenon no longer emerged because of the proximity of the categories. On the surface, the gaps emerged as consequences of historical motivations: since there was no /d^h/ in Old High German, the High German sound shift could not create /d͡z/ parallelly to /t͡s/, etc.

To give another example, standard Polish (/s̪, z̪, t͡s̪, d͡z̪, ɕ, ʑ, t͡ɕ, d͡ʑ, ṣ, ẓ, t͡ṣ, d͡ẓ/) would have the grammar inventory ABC with voicing and affricates, the neural network inventory ABC, the implementation of dispersion /s̪ - ɕ - ṣ/, and no gap in the surface inventory.

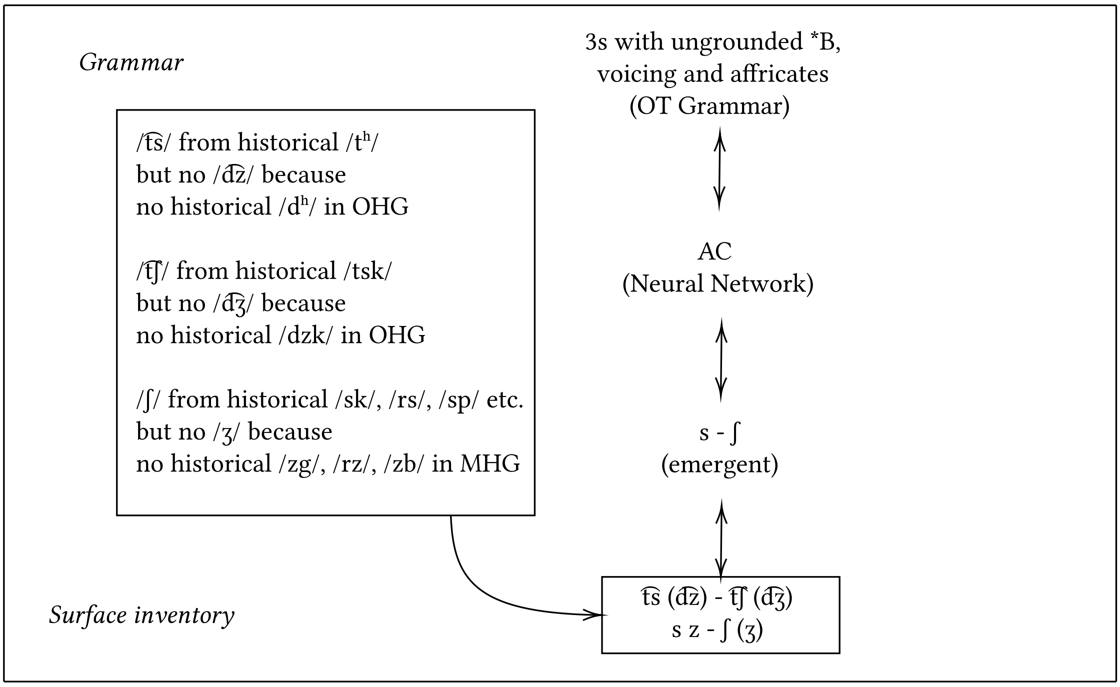


Figure CVIII: *Schematic representation of the different layers involved between the OT inventory grammar and the surface inventory, illustrated with Northern German.*

The important claim is that the gaps in the surface inventory are typically explainable from a historical perspective, but are emergent and can occur without any possible synchronic explanation. As shown in 3.2.3, they are possible and synchronically unpredictable at any position in the inventory, except for the unattested gap at /s/ (i.e. every language with sibilants must either have a (denti-)alveolar or retracted alveolar voiceless sibilant fricative). One will thus need to anticipate every possible gap, together with all possible combinations of voicing, manner and distributedness, to predict all possible surface inventories: this will be done in the next subsubsection.

5.4.2.1 Voicing and manner distinctions

The addition of the predicted voicing and manner (fricative - affricate) distinctions of the extended OT typology (see 5.2.2.5) to the 7 inventories resulting from the OT + NN simulations transforms them into 25 inventories, as detailed in table 23. In the extended OT typology, there are precisely 4 versions of each inventory that has sibilants and 1 version of the sibilantless language, i.e. (4 x 6) + 1 = 25. The next subsubsection will detail the gaps that can occur, taking into account PoA, manner and voicing.

5.4.2.2 Voicing and manner gaps

To determine what is a possible gap, we will adopt two conditions: 1) there must be at least one anterior voiceless sibilant fricative in every inventory (which is not a 0s), as evidenced in 3.2.3.1, and 2) a voiced sibilant must have a voiceless counterpart, as shown in 3.2.3.3. These two conditions are empirically motivated (more precisely, based

Main type	.		B		AB		AC		BC		ABC		ABCD	
Precise type	Vo	Af	Vo	Af	Vo	Af	Vo	Af	Vo	Af	Vo	Af	Vo	Af
	X	X	X	X	X	X	X	X	X	X	X	X	X	X
			V	X	V	X	V	X	V	X	V	X	V	X
			X	V	X	V	X	V	X	V	X	V	X	V
			V	V	V	V	V	V	V	V	V	V	V	V

Table 23: *The 25 possible inventories predicted by the extended OT sibilant typology (see 5.2.2.5) and auditory dispersion in Neural Networks.* Vo *stands for voicing distinctions, and* Af *for fricative - affricate distinctions.*

on typological observations in chapter 3) and can be found in the extended OT typology: the voicing universal directly as voicing markedness, and the anterior fricative universal as a combination of the unmarkedness of B and affricate markedness.[10]

The possible gaps will be detailed here per PoA, from 1 to 4. In a 1s, the possibilities are as reported in table 24; in a 2s, they are as in table 25. The logically possible

1s	0 gap	1 gap	2 gaps	3 gaps	4 gaps
XX	s̠	*	/	/	/
VX	s̠ z̠	s̠ *z̠	*	/	/
XV	s̠ t͡s̠	s̠ *t͡s̠	*	/	/
VV	s̠ z̠ t͡s̠ d͡z̠	s̠ t͡s̠ d͡z̠ s̠ z̠ t͡s̠ *z̠ t͡s̠ d͡z̠ *s̠ z̠ d͡z̠	s̠ t͡s̠ s̠ z̠ *z̠ d͡z̠ *z̠ t͡s̠ *t͡s̠ d͡z̠ *s̠ d͡z̠	s̠ *z̠ *t͡s̠ *d͡z̠	*

Table 24: *Possible voicing and manner gaps in a 1s. Note that* XX, VX, XV *and* VV *correspond to the values for voicing and manner, respectively, reported in table 23. An asterisk corresponds to a logically possible but unpredicted inventory, a slash to a logically impossible inventory (because one cannot have more gaps than segments).*

inventories that were ruled out are indicated with an asterisk (*) in the table. They are all inventories that violate at least one of the two conditions formulated above.

It becomes apparent that certain inventories in the table are 'homonymous': how can one distinguish a language without voicing and affricates, which therefore only has

[10]Opponents to these two conditions would then assume that the number of possible inventories equals the number of logically possible inventories given the 7 basic patterns. With voicing and affricates as possible distinctions, this would mean 71199 inventories, i.e. as much as 83.91% of the 84849 logically possible inventories if any basic pattern (AD, BCD etc.) is allowed, and include unrealistic phenomena such as e.g. a 4s with gaps at all sibilants, distinct from a 3s with gaps at all sibilants, distinct from a 0s.

	0 gap	1 gap	2 gaps	3 gaps	4 gaps	5 gaps	6 gaps	7 gaps	8 gaps
XX	s ʃ	s *ʃ	*	/	/	/	/	/	/
VX	s z ʃ ʒ	s ʃ ʒ s z ʃ *z ʃ ʒ *s z ʒ	s ʃ s z *ʃ ʒ *z ʒ *z ʃ *s ʒ	s *z *ʃ *ʒ	*	/	/	/	/
XV	s t͡s ʃ t͡ʃ	s ʃ t͡ʃ s t͡s ʃ *t͡s ʃ t͡ʃ *s t͡s t͡ʃ	s ʃ s t͡s *ʃ t͡ʃ *t͡s t͡ʃ *t͡s ʃ *s t͡ʃ	s *t͡s *ʃ *t͡ʃ	*	/	/	/	/
VV	s z t͡s d͡z ʃ ʒ t͡ʃ d͡ʒ	s t͡s d͡z ʃ ʒ t͡ʃ d͡ʒ s z t͡s ʃ ʒ t͡ʃ d͡ʒ s z t͡s d͡z ʃ t͡ʃ d͡ʒ s z t͡s d͡z ʃ ʒ t͡ʃ *z t͡s d͡z ʃ ʒ t͡ʃ d͡ʒ *s z d͡z ʃ ʒ t͡ʃ d͡ʒ *s z t͡s d͡z ʒ t͡ʃ d͡ʒ *s z t͡s d͡z ʃ ʒ d͡ʒ	[28 inv., out of which 19 have a *]	[56 inv., out of which 43 have a *]	[70 inv., out of which 57 have a *]	[56 inv., out of which 47 have a *]	[28 inv., out of which 24 have a *]	s *z *t͡s *d͡z *ʃ *ʒ *t͡ʃ *d͡ʒ	*

Table 25: *Possible voicing and manner gaps in a 2s.* XX, VX, XV *and* VV *correspond to the values for voicing and manner (see table 23). An asterisk corresponds to a logically possible but unpredicted inventory, a slash to a logically impossible inventory.*

/s̠/, from a language that allows e.g. voicing and has a voicing gap so that on the surface, one also finds only /s̠/? The answer is that in the former case, voiced sibilants are prohibited in the inventory grammar (m.Voice(Sib) ≫ f.Ident(Sib)) *and* allophonically (m.Voice(Sib) ≫ VoicingAssimilation), so that e.g. Icelandic /s̠/ does not surface as voiced in voiced environments (Hansson 2003; Flego and Berkson 2020); in the latter case, voiced sibilants are not phonemic but are found allophonically, e.g. in Castilian Spanish *desde* 'since' with [z̠d] (VoicingAssimilation ≫ m.Voice(Sib) ≫ f.Ident(Sib); see B.10). There are thus languages that have voiced sibilant phonemes (V), languages without voiced sibilant phonemes that have voiced sibilant allophones (V with gaps), and languages that have no voiced sibilant at all (X). This difference is accordingly reflected in the predicted possible gaps of table 24. However, the 1s with a gap at each sibilant phoneme would violate condition 1 about the presence of a voiceless anterior sibilant fricative, and are thus eliminated. Alternatively, one could consider that they do not violate it since they are a 0s, which would then represent the 'hidden 1s' type, i.e. a 0s that has sibilants allophonically, as described in 3.2.2.1.

The detailed calculations yielding the number of logically possible sibilant inventories as well as the number of predicted ones can be found in the appendix (section 8.3). It is explained there, with the help of mathematical formulas, how many sibilant inventories satisfy each of the two conditions individually (/s/-gap and voicing gap) and both conditions simultaneously. The number and proportion of predicted wrt. logically possible inventories for each number of sibilant PoAs, voicing and manner possibility as well as number of gaps are represented in fig. CIX. This figure has been split into

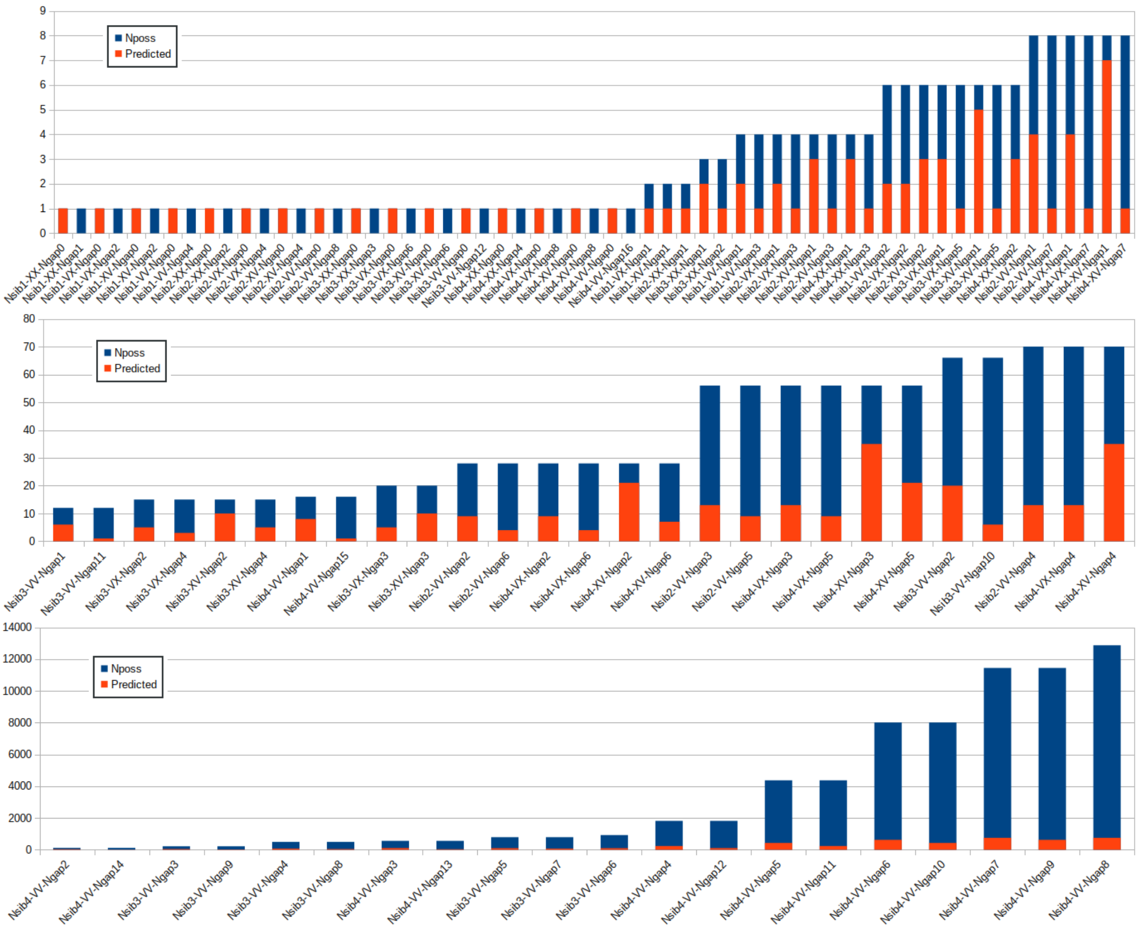


Figure CIX: *Percentual proportion of the number of predicted inventories with respect to the number of logically possible inventories, per number of PoAs, of voicing-manner distinctions and of gaps.*

three: with a number of logically possible inventories ranging from 0 to 8, from 12 to 70 and from 120 to 12870. The entire bar irrespectively of colour indicates the number of logically possible inventories for a given number of sibilant PoAs, voicing/affricate distinctions and gaps, while the red part of the bar amounts to the number of predicted inventories within these possibilities. The proportion of predicted inventories decreases proportionally as the number of logically possible sibilant inventories increases. This has to do with the very large number of possible unpredicted patterns when the number of possible inventories is itself very large. One observes also that in rows in which only the /s/-gap condition is active (by general lack of voiced sibilants), the proportion of predicted wrt. logically possible inventories decreases linearly. However, when also the voicing gap condition is active (i.e. in inventories that have voicing distinctions), this proportion decreases much more with fewer gaps, and less with more gaps. This is illustrated in fig. CX.

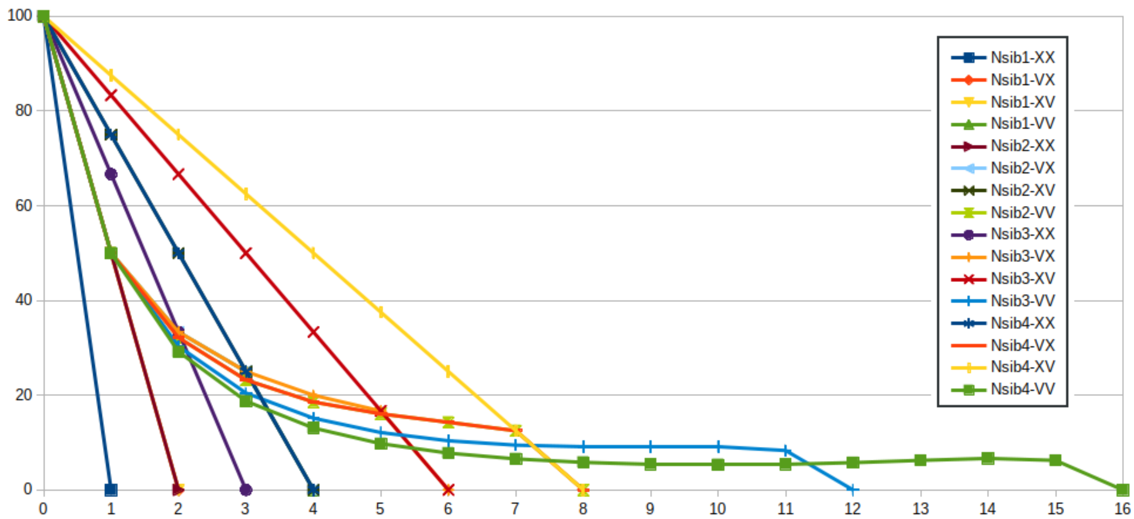


Figure CX: *Percentual proportion of the number of predicted inventories with respect to the number of logically possible inventories, per number of PoAs, of voicing-manner distinctions and of gaps.*

Summing up all the predicted inventories, one obtains the number of 11/26 for the 1s, 70/292 for each of the three 2s, 540/4232 for the 3s and 4564/66064 for the 4s. This amounts to 5325 predicted inventories out of 71198. Adding the 0s, one thus obtains 5326 predicted inventories out of 71199, which represents only 7.48% of logically possible inventories. The same number of predicted inventories constitutes 6.28% of the 84849 logically possible inventories if any pattern (e.g. AD) is allowed.

5.4.2.3 Distributedness distinctions

Besides voicing and manner, there is a third distinction that is made in sibilant inventories: distributedness. The NN model presented above did not include distributedness distinctions, which however also emerge as a way of enhancing contrasts, i.e. as auditory dispersion. One observes (as witnessed in 3.2.2.3, 3.2.2.4 and 3.2.2.5) that less-dispersed inventories (that is, inventories that have adjacent A, B, C and/or D) make use of distributedness contrasts, unlike the well-dispersed AC inventory. Not only do inventories with more phonemes enhance contrasts with e.g. apical-laminal contrasts where less crowded inventories allow free variation instead (Dart 1991, 1993), but distributedness distinctions are also more likely to occur if the categories are close to each other in terms of CoG.

Adding distributedness distinctions is thus only relevant for languages in which there are more than one sibilant PoA, and in which the segments are adjacent (i.e. the languages AB, BC, ABC and ABCD). The distinctions are expectedly also always with different values (either apical or laminal) for neighbouring segments, so that one gets e.g. apical A and laminal B or the other way around, but not twice the same value (something which would make the ‘contrast’ pointless by making it no contrast at all; cfr. Hall 1997*b*: 42). The possibilities are reported in table 26. This means that instead of having

# of PoAs	Inventory without distributedness	Inventory with distributedness
2	AB	A̻B̺
		A̺B̻
	BC	B̻C̺
		B̺C̻
3	ABC	A̻B̺C̻
		A̺B̻C̺
4	ABCD	A̻B̺C̻D̺
		A̺B̻C̺D̻

Table 26: *Possible distributedness distinctions in inventories with 2 or more sibilants.*

three possible 2s (AB, AC and BC), there will be five (A̻B̺, A̺B̻, AC, B̻C̺ and B̺C̻); instead of one possible 3s, there will be two (A̻B̺C̻ and A̺B̻C̺); and instead of one possible 4s, there will be two (A̻B̺C̻D̺ and A̺B̻C̺D̻). Reusing the numbers of the previous subsubsection, one can thus calculate that the number of predicted inventories will be: 1x1 (0s) + 1x11 (1s) + 5x70 (2s) + 2x540 (3s) + 2x4564 (4s), which amounts to 10570 predicted inventories, almost twice the number of predicted inventories without distributedness distinctions. The number of logically possible inventories, however, increases drastically when adding distributedness to the model: assuming that any segment can be absent, apical, laminal or present but unspecified for distributedness, this changes the initial formula for the number of possible segments to $4^{N_{seg}}$ rather than $2^{N_{seg}}$ (with N_{seg} being the number of segments in a given inventory if there is no gap). Nothing changes in a 0s but in a 1s, the rows get a total number of possibilities of 4, 16, 16 and 256 rather than 2, 4, 4 and 16 as in table 24, amounting to 292 logically possible inventories. A 2s then has 66064 possibilities, a 3s has 16785472 and a 4s has 4295098624. Considering that there are 1 logically possible pattern of 0s, 4 of 1s, 6 of 2s, 4 of 3s and 1 of 4s, the total number of logically possible inventories, taking into account any possible distributedness value, equals 4362638065. Out of these 4 billions, the 10570 predicted inventories represent only 0.00024%, which is around one four-thousandth of the theoretical possibilities ($\frac{1}{4127}$).

5.5 A sibilant typology in Distinctive Feature Theory

5.5.1 Pure DFT

In 5.1, it has been mentioned how Distinctive Feature Theory (Hall 2001; Clements 2009) predicts possible sibilant inventories based on a few distinctive features. The relevant common features of sibilants are (based on Hall 1997*b*: 134-150): [strident], which characterises and delimits the entire class of sibilants with respect to all non-sibilants; [coronal], which is also common to all sibilants, since they are articulated mainly with the front part of the tongue; and of course, [consonantal] as well as [obstruent]. Features that are distinctive within the class of sibilants are: [voice], which distinguishes voiced from voiceless sibilants; [continuant], which distinguishes fricatives ([+cont]) from affricates ([-cont]; cfr. Kehrein 2002); [±anterior], which distinguishes dental and alveolar sibilants (Hall 1997*b*: 38) as well as retracted alveolar (according to me, but not e.g. Adams 1975: 284) and alveopalatal (according to Sagey 1986, cited in Hall 1997*b*: 145, but not Hall 1997*b*: 60 himself), from posterior sibilants such as palatoalveolars, subapical and plain retroflexes; [high], which according to Hall (1997*b*: 147) includes "palatoalveolars, alveolopalatals, palatals, and velars (and high vowels like [i] and [u])"; [±distributed], which distinguishes apical from laminal; and finally, several secondary articulations that we have chosen from the start to disregard in this dissertation (e.g. [spread glottis] for /s^h/ or [constricted glottis] for /s'/). The relevant features are illustrated in fig. CXI, showing their placement in a feature tree. Note that this feature tree

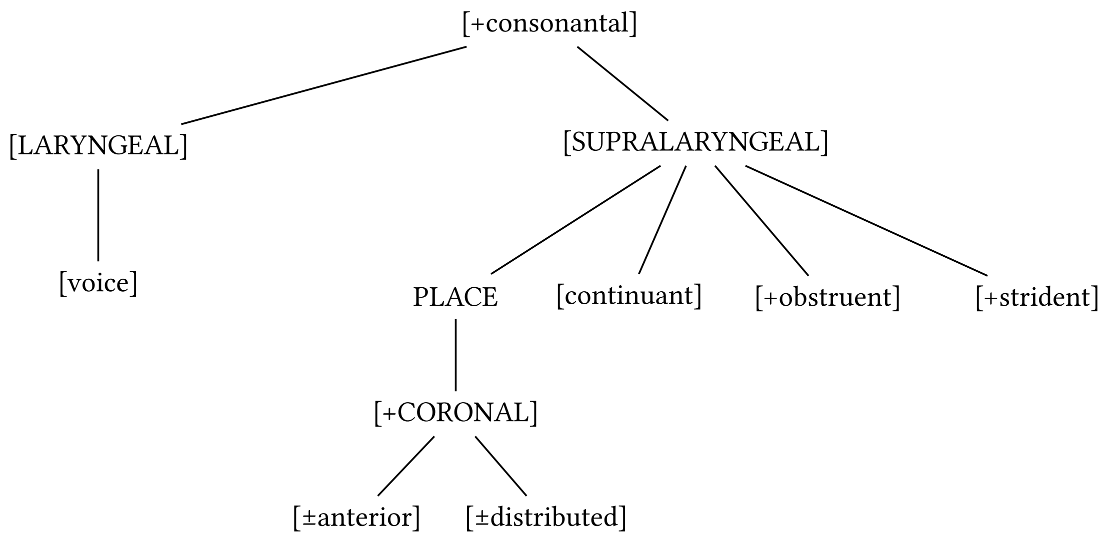


Figure CXI: *Distinctive features relevant for sibilant inventories.*

only includes the features that are relevant for sibilants. Those who are common to all sibilants have a '+' sign. With respect to the 'major' PoA of coronal segments, Clements (2004: 3f.) explains:

> Features make strong predictions about the number and types of *contrasts* a language may have.
> Example: "major place of articulation" in coronal sounds is defined by two features:
> [anterior/posterior]: dental, alveolar vs. postalveolar, retroflex, palatal
> [apical/distributed]: apical (including retroflex) vs. laminal ("extended contact")
>
> These two features define four major categories:

	apico-anterior	lamino-anterior	retroflex	postalveolar/palatal
posterior	-	-	+	+
distributed	-	+	-	+

The relevant features to generate the predicted sibilant inventories are not only [±anterior] and [±distributed], but also [±voice] and [±continuant]. The feature [±high] is not required for this purpose and has been omitted in fig. CXI, since it overlaps with

[±distributed] at least in this case (e.g. [+anterior, +high] = [+anterior, +distributed] = [ɕ]). Table 27 illustrates the DFT typology of sibilant inventories with these features, in the line of Clements (2004: 3f.). Besides the sibilantless possibility, there are 16 rows because 2^4 possibilities with four features. There can logically be any number of active features from 0 to 5, but if [strident] is not active, all the rest is not relevant. Some striking properties of this DFT sibilant typology appear when one examines in detail the predicted inventories, as in table 27.

A first observation is that this model accounts for allophonic variation without any difficulty. For example, if only [strident] is contrastive, allophones can have any value for [voice] (e.g. voiced in between two vowels, voiceless before /p/), [continuant] (e.g. affricate after /n/, fricative in between vowels), [±distributed] (e.g. laminal before laminal sounds, apical before apical sounds) and [±anterior] (e.g. more anterior after /t/, more posterior before /i/). This is reflected in table 27 as ‘s̠ ((t͡)s̻∼(t͡)s̺̠∼(d͡)z̻∼(d͡)z̺̠)’. The allophones between parentheses show that in certain contexts, a sibilant could have a different phonetic value without any implication for phonological contrast.

The other side of the coin with this allophonic variation is however the ‘unspecification’ of the actual phoneme: if only [strident] is distinctive, nothing tells us *a priori* what the main allophone will be, since markedness considerations are not incorporated into the model. Typical DFT accounts for the typology would see /s/ as the realisation of the sibilant phoneme if only [strident] is distinctive (e.g. Hall 1997*b*: 91, though apparently also allowing a single /ʃ/), ignoring the nuance between alveolar and retracted alveolar. But if a language has only [strident], why would it not have [t͡s̻] as main sibilant allophone? In table 27, I have used retracted alveolar sibilants as default value thanks to the knowledge acquired from the OT sibilant typology; but in the perspective of only the ‘pure DFT’ approach, much is said about the nature of the contrasts while little is said about the actual phonetic nature of the contrasting segments (about which DFT most often does not aim to make predictions).

Another question that can be raised is about the dependence of [±distributed] on [±anterior] (as is visible where ‘# of PoAs’ = ‘1 (?)’ in table 27). In principle, the typology predicts feature combinations with distinctive [±distributed] and without [±anterior], i.e. apical/laminal contrasts without place contrasts. However, as Ladefoged (1971: 39) observes (cited in Hall 1997*b*: 42), “apical vs. laminal contrasts are virtually always accompanied by different places of articulation” (Hall 1997*b*: 42). The DFT typology predicts languages with [±distributed] without [±anterior], but these are unattested as such, being rather attested as also having [±anterior]. The other way around, [±anterior] without [±distributed] is possible (e.g. French) and even the most widespread sibilant inventory pattern (recall e.g. fig. XL). One can still interpret the predicted inventories with only [±distributed] as the AB or BC pattern, which would partly clarify this question, but the lack of explanation for why one would then not find distributedness contrasts with e.g. CD instead remains. This brings us back to the problematic lack of markedness representation in the theory.

Another concern is the lack of ‘proper’ 3s inventories in these predictions. The attested 3s are consistently analysed as having both [±anterior] and [±distributed] (e.g.

<table>
<tr><th># of PoAs</th><th># of active features</th><th>Active features</th><th colspan="2">Sibilant inventory
(contrasts and phonetic realisations)</th></tr>
<tr><td>0</td><td>0</td><td>nihil</td><td colspan="2">(no sibilants)</td></tr>
<tr><td rowspan="4">1</td><td>1</td><td>[strident]</td><td colspan="2">s̠ ((t͡)s̪~(t͡)ʂ~(d͡)z̪~(d͡)ʐ)</td></tr>
<tr><td>2</td><td>[strident], [voice]</td><td colspan="2">s̠ ((t͡)s̪~(t͡)ʂ) - z̠ ((d͡)z̪~(d͡)ʐ)</td></tr>
<tr><td>2</td><td>[strident], [continuant]</td><td colspan="2">s̠ (s̪~ʂ~z̪~ʐ) - t͡s̠ (t͡s̪~t͡ʂ~d͡z̪~d͡ʐ)</td></tr>
<tr><td>3</td><td>[strident], [voice],
[continuant]</td><td colspan="2">s̠ (s̪~ʂ) - z̠ (z̪~ʐ) -
t͡s̠ (t͡s̪~t͡ʂ) - d͡z̠ (d͡z̪~d͡ʐ)</td></tr>
<tr><td rowspan="4">2</td><td>2</td><td>[strident], [anterior]</td><td colspan="2">s ((t͡)s~(d͡)z) - ʃ ((t͡)ʃ~(d͡)ʒ)</td></tr>
<tr><td>3</td><td>[strident], [anterior],
[voice]</td><td colspan="2">s ((t͡)s) - z ((d͡)z) -
ʃ ((t͡)ʃ) - ʒ ((d͡)ʒ)</td></tr>
<tr><td>3</td><td>[strident], [anterior],
[continuant]</td><td colspan="2">s (z) - t͡s (d͡z) -
ʃ (ʒ) - t͡ʃ (d͡ʒ)</td></tr>
<tr><td>4</td><td>[strident], [anterior],
[voice], [continuant]</td><td colspan="2">s - t͡s - z - d͡z -
ʃ - t͡ʃ - ʒ - d͡ʒ</td></tr>
<tr><td rowspan="4">1 (?)</td><td>2</td><td>[strident], [distributed]</td><td colspan="2">s̻ ((t͡)s̪~(t͡)ʂ~(d͡)z̪~(d͡)ʐ) - s̺ ((t͡)s̪~(t͡)ʂ~(d͡)z̪~(d͡)ʐ)</td></tr>
<tr><td>3</td><td>[strident], [distributed],
[voice]</td><td colspan="2">s̻ ((t͡)s̪~(t͡)ʂ) - z̻ ((d͡)z̪~(d͡)ʐ) -
s̺ ((t͡)s̪~(t͡)ʂ) - z̺ ((d͡)z̪~(d͡)ʐ)</td></tr>
<tr><td>3</td><td>[strident], [distributed],
[continuant]</td><td colspan="2">s̻ (s̪~ʂ~z̪~ʐ) - t͡s̻ (t͡s̪~t͡ʂ~d͡z̪~d͡ʐ) -
s̺ (s̪~ʂ~z̪~ʐ) - t͡s̺ (t͡s̪~t͡ʂ~d͡z̪~d͡ʐ)</td></tr>
<tr><td>4</td><td>[strident], [distributed],
[voice], [continuant]</td><td colspan="2">s̻ (s̪~ʂ) - z̻ (z̪~ʐ) -
t͡s̻ (t͡s̪~t͡ʂ) - d͡z̻ (d͡z̪~d͡ʐ) -
s̺ (s̪~ʂ) - z̺ (z̪~ʐ) -
t͡s̺ (t͡s̪~t͡ʂ) - d͡z̺ (d͡z̪~d͡ʐ)</td></tr>
<tr><td rowspan="4">4</td><td>3</td><td>[strident], [anterior],
[distributed]</td><td>s̪ ((t͡)s̪~(d͡)z̪) - ɕ ((t͡)ɕ~(d͡)ʑ) -
ṣ ((t͡)ṣ~(d͡)ẓ) - ʂ ((t͡)ʂ~(d͡)ʐ)</td><td>s̪ ((t͡)s̪~(d͡)z̪) - s̻ ((t͡)s̻~(d͡)z̻) -
ʃ̻ ((t͡)ʃ̻~(d͡)ʒ̻) - ʂ ((t͡)ʂ~(d͡)ʐ)</td></tr>
<tr><td>4</td><td>[strident], [anterior],
[distributed], [voice]</td><td>s̪ ((t͡)s̪) - z̪ ((d͡)z̪) -
ɕ ((t͡)ɕ) - ʑ ((d͡)ʑ) -
ṣ ((t͡)ṣ) - ẓ ((d͡)ẓ) -
ʂ ((t͡)ʂ) - ʐ ((d͡)ʐ)</td><td>s̪ ((t͡)s̪) - z̪ ((d͡)z̪) -
s̻ ((t͡)s̻) - z̻ ((d͡)z̻) -
ʃ̻ ((t͡)ʃ̻) - ʒ̻ ((d͡)ʒ̻) -
ʂ ((t͡)ʂ) - ʐ ((d͡)ʐ)</td></tr>
<tr><td>4</td><td>[strident], [anterior],
[distributed], [continuant]</td><td>s̪ (z̪) - t͡s̪ (d͡z̪) -
ɕ (ʑ) - t͡ɕ (d͡ʑ) -
ṣ (ẓ) - t͡ṣ (d͡ẓ) -
ʂ (ʐ) - t͡ʂ (d͡ʐ)</td><td>s̪ (z̪) - t͡s̪ (d͡z̪) -
s̻ (z̻) - t͡s̻ (d͡z̻) -
ʃ̻ (ʒ̻) - t͡ʃ̻ (d͡ʒ̻) -
ʂ (ʐ) - t͡ʂ (d͡ʐ)</td></tr>
<tr><td>5</td><td>[strident], [anterior],
[distributed], [voice],
[continuant]</td><td>s̪ - z̪ - t͡s̪ - d͡z̪ -
ɕ - ʑ - t͡ɕ - d͡ʑ -
ṣ - ẓ - t͡ṣ - d͡ẓ -
ʂ - ʐ - t͡ʂ - d͡ʐ</td><td>s̪ - z̪ - t͡s̪ - d͡z̪ -
s̻ - z̻ - t͡s̻ - d͡z̻ -
ʃ̻ - ʒ̻ - t͡ʃ̻ - d͡ʒ̻ -
ʂ - ʐ - t͡ʂ - d͡ʐ</td></tr>
</table>

Table 27: *Sibilant inventories predicted by Distinctive Feature Theory.*

Basque: [+ant, +dist] vs. [+ant, -dist] vs. [-ant, +dist]; Polish: [+ant, -dist] vs. [+ant, +dist] vs. [-ant, -dist]), but they would then all lack a possible fourth member (e.g. Basque: [-ant, -dist]; Polish: [-ant, +dist]). One would then need to consider each 3s to be in fact a 4s with a gap, a gap that is probably not accidental because of its recurrence at always the same PoA (namely D) rather than at any possible slot. This concern could seemingly be answered by seeing the contrast as e.g. [+ant, +dist] vs. [+ant, -dist] vs. [-ant] instead, i.e. allowing distinctive features to not be specified for all segments where it could. This is however problematic to the extent that the underspecified member of the triple then loses information about its distributedness, so that both Basque and Polish would have a [-ant] segment, implying that it is the same kind of segment. It is not at all the same segment: Basque /x/ is a laminal and slightly retracted [ʃ], while Polish /sz/ is an apical and slightly retracted [ṣ] (or: [ʃ̺]).

In 'pure' DFT, all possible contrasts between sibilants can be represented, but as I have shown, it lacks an integration of markedness to make predictions about the actual phonetic nature of the contrasting segments. This issue will be tackled in the next subsection, with a typology of sibilant inventories in a combination of OT and DFT.

5.5.2 OT and DFT combined

Two possible approaches will be detailed here: one in which the sibilant candidates in OT are specified for a certain feature (i.e. have + or - some feature, binarily), and one in which they can also be underspecified (i.e. [+feature], [-feature] or [0feature]). The underspecified option is sometimes (but not uncontestedly) used in the literature (e.g. Kim 2002) to represent the idea that some segments might be unspecified for one feature although the feature is active elsewhere in the language, even in other similar segments (e.g. Turkish underlying /t/, /d/ or /T/; Kim 2002).

In the first case, the constraints and candidates look precisely as in the basic OT typology of 5.2.2.1, except for the fact that the candidates are no longer simply A, B, C and D but are specified for [±anterior] and [±distributed] as well. The letter corresponding to the PoA is put between brackets to distinguish [A] from Anterior and Apical, respectively. Thus, [C]PL means for example 'C, posterior, laminal'. The Factorial Typology resulting from evaluation is shown in fig. CXII. One observes that this version differs

SibInvFF	FacTyp				
Inputs->	[A]AL	[B]AA	[C]PL	[D]PA	.....
L.1	[A]AL	[B]AA	[A]AL	[B]AA	.....
L.2	[A]AL	[B]AA	[C]PL	[B]AA	.....
L.3	[A]AL	[B]AA	[C]PL	[D]PA	.....
L.4	[B]AA	[B]AA	[B]AA	[B]AA	.....
L.5	[B]AA	[B]AA	[C]PL	[B]AA	.....
L.6	.....	.....	.....	.....	.....

Figure CXII: *Factorial Typology (in OTWorkplace, Prince et al. 2007-2021) of sibilant inventories with features.*

from the basic sibilant typology in the sense that it allows for 'distributedness mappings': for example, [D]PA prefers [B]AA over his immediate neighbour [C]PD because the former is also apical. This is a consequence of f.Ident(Sib) choosing [B]AA because the string has two letters of difference with [D]PA (namely the PoA and the value for [±anterior]), but [C]PD has a different value for [±distributed] in addition to the two differences already mentioned. This version of the typology thus predicts languages in which sibilants can ‘hop’ above a place of articulation to a further one to merge with it because of their identical distributedness value. It remains rather improbable that such a language could exist, in which a [ʃ] is e.g. reinterpreted as /s̺/. For the rest, it does not differ extensively from the basic sibilant typology, as witnessed by their identical Property Analyses.

It has been assumed that retracted alveolar sibilants are apical and (denti-)alveolar sibilants laminal by default. This is in accordance with Hall (1997*b*: 143), who assumes “that dentals are laminal and alveolars apical. See Ladefoged (1968), and Ladefoged and Maddieson (1996), who observe this strong cross-linguistic tendency”. If one allows both options, namely also apical dentals, laminal retracted alveolars etc., one also obtains 6 languages in the Factorial Typology, but with the mapping that we will recognise from the basic OT typology, as illustrated in fig. CXIII. It is worth noticing that the mappings

SibInvFF	FacTyp								
Inputs->	[A]AL	[B]AA	[C]PL	[D]PA	.....	[A]AA	[B]AL	[C]PA	[D]PL
L.1	[A]AL	[B]AA	[C]PL	[D]PA	.....	[A]AA	[B]AL	[C]PA	[D]PL
L.2	[A]AL	[B]AA	[C]PL	[C]PA	.....	[A]AA	[B]AL	[C]PA	[C]PL
L.3	[A]AL	[B]AA	[B]AL	[B]AA	.....	[A]AA	[B]AL	[B]AA	[B]AL
L.4	.....	.....	.....	.....	.....	.....	.....	.....	.....
L.5	[B]AL	[B]AA	[C]PL	[C]PA	.....	[B]AA	[B]AL	[C]PA	[C]PL
L.6	[B]AL	[B]AA	[B]AL	[B]AA	.....	[B]AA	[B]AL	[B]AA	[B]AL

Figure CXIII: *Factorial Typology (in OTWorkplace, Prince et al. 2007-2021) of sibilant inventories with features, this time allowing all [distributed] combinations.*

are precisely the same as in the basic OT typology, although F.IDENT(SIB) is formulated here as a categorical constraint. This is an effect of the [±anterior] feature: D is mapped to C because if it cannot be mapped faithfully, the best remaining faithful option is to become C, with which D shares the ‘P’ of ‘Posterior’. This version thus incorporates in some way the idea that a mapping to a further sibilant (namely one that does not share [±anterior]) is dispreferred. This typology with DFT can thus serve as an additional argument in favour of the basic OT sibilant typology with gradual faithfulness: the same effect arises if the articulatory/acoustic distance or feature specification is encoded to reflect the fact that neighbouring sibilants are more similar to each other than distant sibilants.

A weakness of the analyses provided until here is that one obtains candidates with fully-specified features regardless of the inventory, although the inventory itself in fact determines which features are distinctive. To tackle this problem, a version of the OT + DFT typology without full specification is made as follows: candidates consist of strings with the structure ‘S[]A[]D[]’, in which the brackets can contain ‘0’, ‘+’ or ‘-’. ‘S’ stands for [strident], ‘A’ for [±anterior] and ‘D’ for [±distributed]. For example, ‘S+A+D-’ represents an apical retracted alveolar /s̺/. To avoid dealing with a very large number of rows in OTWorkplace ($3^4 + 1 = 82$ candidates, which means $82 \cdot 82 = 6724$ rows and a hardly readable typology), [voice] and [continuant] have been left aside in this typology. Every possible output is also a possible input and vice-versa. There is only one candidate in which [strident] is not contrastive, namely 'S0A0D0' ([±anterior] and [±distributed] are irrelevant in that case). Fig. CXIV gives an overview of the Violation Tableau. The screenshot is cut at the 21st row, but there are of course 100 rows with candidates (9 possibilities for [±anterior] and [±distributed] + 1 sibilantless possibility). The constraints consist simply of a stringency hierarchy of constraints against specifi-

SibInv	VT.ini		*+, -	func:Plus2	func:Plus3	ffunc:diff
input	output	opt	**Plus1**	**Plus2**	**Plus3**	**Faithfulness**
S0A0D0	S0A0D0		0	0	0	0
	S+A0D0		1	0	0	1
	S+A+D0		2	2	0	2
	S+A-D0		2	2	0	2
	S+A0D+		2	2	0	2
	S+A0D-		2	2	0	2
	S+A+D+		3	3	3	3
	S+A-D-		3	3	3	3
	S+A+D-		3	3	3	3
	S+A-D+		3	3	3	3
S+A0D0	S0A0D0		0	0	0	1
	S+A0D0		1	0	0	0
	S+A+D0		2	2	0	1
	S+A-D0		2	2	0	1
	S+A0D+		2	2	0	1
	S+A0D-		2	2	0	1
	S+A+D+		3	3	3	2
	S+A-D-		3	3	3	2
	S+A+D-		3	3	3	2
	S+A-D+		3	3	3	2
S+A+D0	S0A0D0		0	0	0	2

Figure CXIV: *Violation Tableau (in OTWorkplace, Prince et al. 2007-2021) of sibilant inventories with features, this time allowing for underspecified sibilant features.*

cation vs. one faithfulness constraint. The first constraint against specification (a kind of markedness constraint, since specification *is* markedness in this case) penalises any specification i.e. any '+' or '-'. The second only assigns violations when there are at least two specifications, and the third when there are at least three. This typology yields 4 languages, detailed in fig. CXV. Unsurprisingly, they correspond to the possibilities of having 0, 1, 2 or 3 specification(s) respectively. One obtains one language without sibilants, one with only [strident] active, one which allows two features to be active and one which allows all three. The language with two features allows either to have only [±anterior] or [±distributed], indistinctively; one could also have a constraint against [±anterior] and one against [±distributed] to see that this language would split into two languages.

One will notice that this Factorial Typology matches precisely the sibilant inventories predicted in table 27 (except for the lack of [voice] and [continuant], of course), bringing us back to the DFT typology without markedness encoded against PoAs (A, B, C, D). It

SibInv	FacTyp									
Inputs->	S0A0D0	S+A0D0	S+A+D0	S+A-D0	S+A0D+	S+A0D-	S+A+D+	S+A-D-	S+A+D-	S+A-D+
L.1	S0A0D0	S0A0D0	S0A0D0	S0A0D0	S0A0D0	S0A0D0	S0A0D0	S0A0D0	S0A0D0	S0A0D0
L.2	S0A0D0	S+A0D0	S+A0D0	S+A0D0	S+A0D0	S+A0D0	S+A0D0	S+A0D0	S+A0D0	S+A0D0
L.3	S0A0D0	S+A0D0	S+A+D0	S+A-D0	S+A0D+	S+A0D-	S+A+D0 S+A0D+	S+A-D0 S+A0D-	S+A+D0 S+A0D-	S+A-D0 S+A0D+
L.4	S0A0D0	S+A0D0	S+A+D0	S+A-D0	S+A0D+	S+A0D-	S+A+D+	S+A-D-	S+A+D-	S+A-D+

Figure CXV: *Factorial Typology (in OTWorkplace, Prince et al. 2007-2021) of sibilant inventories with features, this time allowing for underspecified sibilant features.*

thus seems that both variants have their pros and cons (possibility of underspecification vs. integration of phonetic detail through PoA markedness), and can partly account for the attested sibilant inventories although with a few disadvantages listed above. With respect to auditory dispersion, one shall note however that DFT has the advantage that dispersion emerges automatically rather than needing to be added to the model (see Hall 2011).

Modelling within Distinctive Feature Theory, be it with or without Optimality Theory, thus constitute another way of creating a typology of sibilant inventories. Together with Neural Networks, the advantages and disadvantages of the three frameworks (OT, NN, DFT) will be summarised in 6.2, after the comparison of predicted and attested inventories.

6. Conclusion

This chapter concludes with a brief summary of the main findings of this dissertation as follows: section 6.1 compares the logically possible, predicted and attested sibilant inventories, section 6.2 compares the predictions, advantages and disadvantages of the different frameworks used in chapter 5, and section 6.3 lists some more peripheral findings together with an excursus on how this dissertation can contribute to the reconstruction of the Biblical Shibboleth story.

6.1 Comparison of predicted and attested inventories

The sibilant typology presented in section 5.4 of chapter 5, together with the different attested types of sibilant inventories in subsection 3.2.2 of chapter 3, has shown how the rather large diversity of sibilant inventories can be explained and reduced to a small set of synchronically motivated or emergent fundamental traits (4 PoAs at ABCD + the interaction of 8 principles). Besides this aim of reduction to a small set of explanatory traits, the sibilant typology also aims to limit its number of predicted inventories. It must neither overpredict nor underpredict, i.e. not allow any- and everything to be possible if there is evidence that a few patterns are recurring while many are unattested, and allow that which is attested to be possible, respectively. It is obvious that the latter case, underprediction, is much more explicitly and directly problematic, in the sense that tangible evidence can then be found to show that the predictions were inaccurate, whereas overprediction can 'hide' behind the impossibility of demonstrating that that which does not exist does not exist because it is impossible rather than because of chance. Every typological prediction is thus situated somewhere in between the large number of logically possible cases and the small number of attested cases: de Lacy (2011: 18) says that "every theory of phonology predicts many tens of thousands of distinct phonological systems, and only a few thousand have existed and will ever exist", to which Zimmermann (2009: I) adds that "a more restrictive theory should always be favored over a less restrictive theory as long as all attested patterns are derivable". The attested patterns should thus constitute a subset of the predicted patterns, which should constitute a subset of the logically possible patterns, as exemplified in fig. CXVI. Zimmermann (2009)'s claim implies that in fig. CXVI, the sizes of the rectangles for predicted and attested languages should

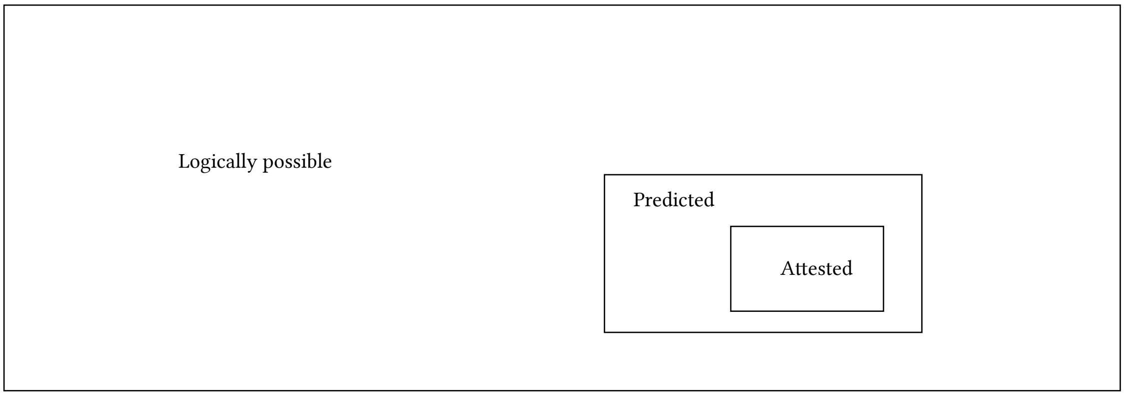


Figure CXVI: *Example of logically possible, predicted and attested sibinvs as subsets.*

be as similar as possible, whereby the smaller rectangle ('attested') must always be fully contained in the larger one ('predicted').

One can now fill such a figure with the proportions between logically possible, predicted and attested inventories obtained in this dissertation. In 5.4.2.2, it has been calculated that in the version of the sibilant typology with 4 PoAs, there are 84849 logically possible and 5326 predicted sibilant inventories. In the SibInv database, there are 75 different patterns when reducing the sibilant inventories to 4 PoAs. This leads to the proportions between logically possible, predicted and attested inventories illustrated in CXVII. A search for duplicate values in the lists of attested and predicted inventories

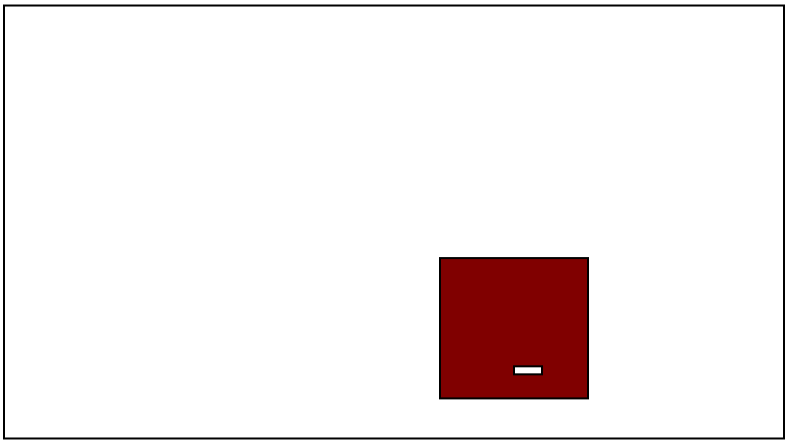


Figure CXVII: *Comparison of logically possible (white), predicted (red) and attested (smaller white) sibilant inventories with 4 PoAs. One pixel corresponds to one language. The numbers have been rounded from 84849 to 84847 (= 391 x 217) and from 5326 to 5325 (= 75 x 71) to obtain integer numbers of pixels to multiply.*

reveals that all attested inventories are also predicted to exist, something which shows that this dissertation has not been underpredicting. It has probably not overpredicted either, in the sense that the red rectangle for predicted inventories is almost 16 times smaller than the logical possibilities, although it is still 71 times larger than the number of attested patterns. However, the number of 71 can be easily explained by looking at the numbers of PoAs from 0 to 4 within the samples: there are 4564 types of 4s that are predicted to be possible by the sibilant typology, and only 3 different types attested in the SibInv database. The larger the number of PoAs, the larger the number of possibilities but also the smaller the number of attested inventories, as was explained in 4.2. The fact that there are 71 times more predicted than attested inventories is thus a consequence of the large number of possible 4s. Figure CXVIII details the proportions between the 5326 predicted and the 75 distinct attested inventories of the SibInv database for each number of PoAs, revealing the discrepancy between the numerous predicted large inventories (3s, 4s) and the rare attested large inventories. If one removes the 4s inventories, the

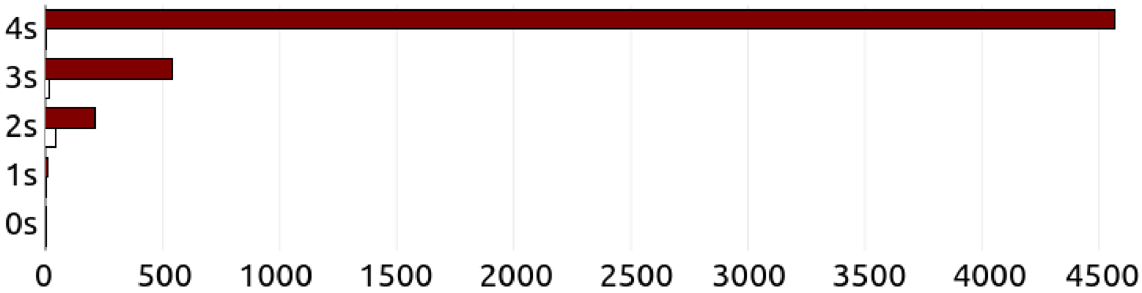


Figure CXVIII: *Number of predicted (red) and attested (white) different types of sibilant inventories reduced to 4 PoAs, per number of PoAs from 0 to 4.*

rectangles for 'predicted' and 'attested' become much more similar in size, with only 10 times more predicted than attested inventories. It is thus clear why the sibilant typology of this dissertation does not overgenerate: the attested sibilant inventories are only a tiny fraction of the possible inventories because larger inventories are rare and imply many possibilities at the same time, as de Lacy (2011: 18) was saying above. Considering the proportions between logically possible, predicted and attested in one frame, as in fig. CXVII, the numbers seem to show that this sibilant typology strikes an optimal balance between the risks of under- and overgeneration.

Turning now to the typological predictions that also include distributedness distinctions and thus extend to 7 PoAs, as calculated in 5.4.2.3, the number of logically possible inventories rises steeply to 4362638065, while the number of predicted inventories is more or less doubled to 10570. The number of different attested patterns does however not change much, becoming 84 rather than 75. The 4 billions of logically possible inventories make it hard to represent the proportions on a simple image; the numbers have been divided by 10570 in fig. CXIX to be able to represent the 10570 predicted languages as 1 pixel on a picture of 741 x 557 pixels (which is approximatively $\frac{4362638065}{10570}$). The large frame in fig. CXIX is thus large like the 4362638065 logically possible inventories where the 10570 predicted ones are contained in one pixel. Furthermore, this figure also represents the proportion between the 10570 predicted inventories and the 84 different attested patterns of the SibInv in the lower right, after zooming to obtain 151 x 70 i.e. 10570 pixels and 9 x 9 i.e. 81 pixels (81 rather than the hardly divisible 83 different attested inventories), respectively. In this case also, all attested languages are predicted to exist.

To conclude, this dissertation posits the existence of 8 principles argued to influence

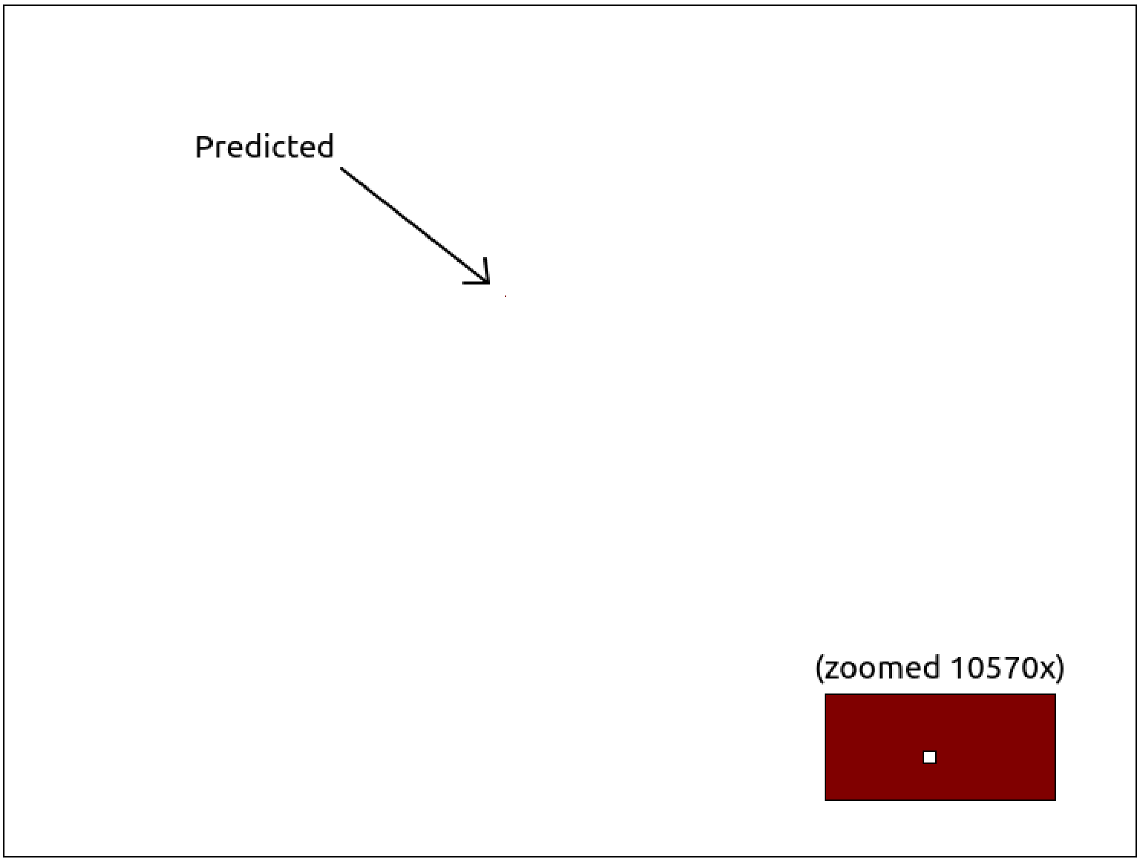


Figure CXIX: *Comparison of logically possible (white), predicted (red) and attested (smaller white) sibilant inventories with 7 PoAs. One pixel corresponds to 10570 languages, except for the part below "zoomed 10570x" where one pixel corresponds to 1 language.*

the structure of sibilant inventories, then creates OT and NN typologies that incorporate these principles and yield 7 fundamental sibilant inventory patterns (sibilantless, B, AB, AC, BC, ABC and ABCD). When adding affricates, voicing, gaps and distributedness distinctions to these predicted patterns, with the assumptions that no gap at a voiceless sibilant or at /s/ can occur, one obtains 10570 predicted inventories. These inventories constitute only 0.00024% of the logical possibilities, yet they match the attested inventories accurately.

6.2 Comparison of the frameworks used

An advantage of all phonological models used in this dissertation (in Optimality Theory, BiPhon Neural Networks and Distinctive Feature Theory) is that they do not appeal to 'inventory constraints', constraints that penalise the auditory contrasts of an inventory or its general structure rather than its single segments (as developed in 5.1.1.1). Nevertheless, a weaker point of the OT typology I proposed is that this has the consequence that the most frequent sibilant inventory, the well-dispersed AC 2s, is not predicted to exist as such. Instead, it surfaces as a 'filtered' variant of the ABC language, either with language-specific cue constraints in BiPhon-OT or with an ungrounded constraint against the absent B. For NN and DFT, this does not constitute a problem from the start, since auditory dispersion is incorporated into these models as an emergent principle (cfr. Boersma et al. 2020 and Hall 2011, respectively).

On the other hand, OT has the advantage of incorporating the principles in a transparent way, as constraints that interact to yield a typology. The ERCs, Hasses, Property Analysis and typohedron shed light on the inner workings of the typology, showing how different rankings can yield one language or another and why. In BiPhon-NN, articulatory markedness constraints such as centralisation and posterior markedness do not interact in language-specific ways (i.e. with one dominating the other in a given language) but rather penalise connections towards the edges and the posterior sibilants invariably. In DFT, principles of articulatory markedness are not integrated and the mapping between contrasts and phonetic realisations unclear (e.g. why an inventory with only one sibilant has the main realisation [s̺] rather than any other PoA is not addressed). In a sense, DFT 'needs' OT to account for markedness and concrete phonetic realisations as much as OT 'needs' NN to account for auditory dispersion. The weak point of DFT for phonetic realisations is however a strong point at the same time, in the sense that it accounts for permissible variation (e.g. a sole sibilant on the continuum can be realised in any way, it will always be perceived as the single existing category). Yet, this holds also for OT and NN, where input-output mappings or connections respectively evidentiate the link between a category and its permissible variation.

An advantage of the DFT model over the other two is that OT and NN do not make any prediction from the start about the number of PoAs that can be contrasted. For OT, I performed a perception experiment in 2.3 to support the modelling with ABCD rather than e.g. ABCDE, which could *a priori* be as valid as ABCD. In NN, the network can be tuned and trained to perceive 5 or 6 distinct PoAs if one wants to. In DFT, the features [±anterior] and [±distributed] by definition limit the range of PoAs to 4 possible values.

Despite all these differences, all three frameworks are able to generate a typology of sibilant inventories in their own way, and are thus considered equivalent in terms of 'efficiency'. The 'pros and cons' detailed here are summarised in table 28.

	OT	NN	DFT
Does not appeal to inventory constraints	+	+	+
Incorporates auditory dispersion	-	+	+
Sheds light on the inner workings and the way in which markedness principles interact	+	+/-	-
Incorporates permissible variation	+	+	+
Admits 4 PoAs *a priori*	-	-	+

Table 28: *Stronger and weaker points of the three frameworks used in this dissertation.*

6.3 Take-home messages

This section summarises, briefly and non-exhaustively, findings with which this dissertation is likely to have contributed to the existing literature.

The sibilant typology makes some phonemic status debates superfluous

The literature is not so infrequently divided on questions such as "Is X a phoneme in language Y?", and it occurs that the sibilant phonemes of a same language are analysed differently by different linguists (e.g. French has no affricates for Jakobson and Lotz 1949, Fougeron and Smith 1993 and Philippart de Foy 2019, but has /t͡ʃ, d͡ʒ/ as marginal phonemes in loanwords according to Goelzer 2005: 54 and Capliez 2016: 83).

The sibilant typology I presented here has the advantage that it predicts a certain number of possible 'slots' for sibilants, which however need not all be occupied. Whether e.g. French has /d͡z/ as a native or marginal phoneme is in fact of little relevance from the perspective of the OT typology: it is predicted to be possible in the inventory of French, and even if that 'allocated slot' is not occupied by a native phoneme, it is predicted to be pronounceable for any speaker of French. One can thus argue that affricates in standard French are or are not phonemes of the core native inventory, in any case their presence is predicted as possible by the inventory grammar.

This further implies that native speakers can use and produce any sibilant in foreign languages without difficulty if and only if the foreign sibilant is predicted to be possible in the native sibilant inventory (Best et al. 2001): for example, a native speaker of French does not struggle with /t͡s, d͡z, t͡ʃ, d͡ʒ/ when learning Italian (both having a well-dispersed 2s), native speakers of English do not struggle with the pronunciation of Mandarin Chinese /t͡s̺/ (which is phonetically very close to the non-phonemic but predicted English /t͡s/) but with that of /t͡ɕ/ and /t͡ʂ/ (Wang and Chen 2020: 11), and native speakers of Hollandic Dutch struggle with the /s - ʃ/ contrast of English (Kwakkel 2008), unlike speakers of Flemish Dutch who have the same /s - ʃ/ contrast as English and have no problem with /ʒ, t͡ʃ, d͡ʒ/ even if these are marginal phonemes in Flemish Dutch. The 'possible slots' predicted for sibilants in a native language (the 'arch-inventory') are thus a direct indicator of what is a possible inventory both in an L1 and L2.

A typology of sibilants needs retracted alveolar as a distinct category

Despite their absence from many typological studies and databases (see 3.1.5.1), retracted alveolar sibilants play a *central* role in the sibilant typology (both in the sense of an important PoA and a central type of sibilant). If typological description wants to be accurate, it is simply required to recognise retracted alveolars as a distinct category to be able to account for non-retroflex 3s and 4s inventories. In the OT and BiPhon-NN models, they play a crucial role for auditory dispersion by constituting an unused category in between A and C in AC languages, where the A and C categories in BiPhon-NN divide for themselves the continuum of B into two roughly equal parts. Ignoring retracted alveolars is thus not only descriptively inaccurate, but also leads to an excessively simplified vision of sibilant PoA contrasts in which the only opposition anterior - posterior overlooks the problematic aspects of auditory dispersion in 2s inventories. Furthermore, overlooking retracted alveolars as a class potentially leads to incorrect understandings of several sound shifts, as argued e.g. in Kokkelmans (2020*a*) for /rs/-clusters (see B.4), Alber et al. (2021) for preconsonantal s-retraction (see B.17) and Kokkelmans (in prep.) for stop palatalisation.

Posterior markedness interacts with centralisation, explaining the patterns B-AB-AC-BC-ABC

A crucial aspect of the typology presented in the dissertation is that auditory dispersion does not only interact with centralisation, as was already the case in e.g. Boersma and Hamann (2008) and Boersma et al. (2020), but also with posterior markedness. Posterior markedness was recognised as a phenomenon by a few linguists (see 4.4.1) until now but not yet integrated together with auditory dispersion and centralisation into a typological model of sibilant inventories. In the OT typology, posterior markedness provokes the appearance of the AB language and the B language in a direct way, in the sense that AB and B would not exist in the Factorial Typology (but would be ABC(D) and BC instead) if it weren't for posterior markedness. In Neural Networks, the addition of a second articulatory node for posterior markedness shifted the 'preferred' or least marked point of the phonetic continuum from 50% to 37.5%. An inventory with a single category would thus end up centred around 37.5% in NN simulations, while two categories exhibit an interplay of repulsion, due to auditory dispersion, and attraction, due to articulatory markedness. Fig. CXX illustrates this interplay with a concrete physical example: two magnetic balls that repel each other and strive to roll towards the lowest point at the same time. In this example, magnetic energy takes over the role of auditory dispersion and gravity that of articulatory markedness. This abstract and simplified representation

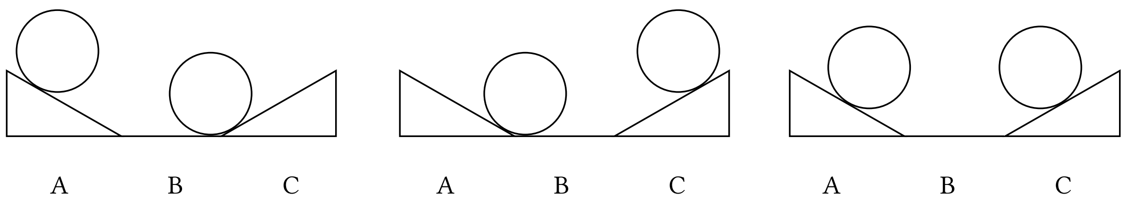


Figure CXX: *Abstract representation of the combined forces of centralisation and posterior markedness vs. auditory dispersion, as gravity and elevation vs. magnetic repulsion.*

of the interactions in the case of two categories on an AudF continuum in NN sheds light on how AB, AC and BC can all equivalently constitute (more or less stable) sibilant inventories. In an AB language, the fact that B is perfectly unmarked explains why A struggles to push it towards C; as a mirror image, C in a BC language struggles to push B forward to A. In an AC language, A and C repeal each other symmetrically so that none of them slides towards B more than the other, and the inventory remains stable. This dissertation has thus integrated in its typology an important aspect of sibilant inventories, posterior markedness, which plays a decisive role for sibilants and could potentially also play a decisive role in coronal inventories in general (compare e.g. denti-alveolar [t] as the default coronal stop realisation in inventories with 1 coronal stop).

Sound shift X always yields phonetic output Y

In 3.3.2, I make the claim and provide evidence that sound shifts that create or delete sibilants yield one precise direct phonetic output, and that at a subsequent stage, this phonetic output is reinterpreted as a phonemic category based on the structure of the inventory. For example, s-retraction in /rs/-clusters with a retracted alveolar [r] always yields an apical retracted alveolar [s̺̠], which can then be reinterpreted by the sibilant inventory in different ways depending on its structure. A well-dispersed /s - ʃ/ language such as Flemish Dutch maps it to /ʃ/ (Kokkelmans 2020*a*), while a non-retroflex 3s such as Western Basque maps it to the pre-existing apical retracted alveolar phoneme /s̺̠/ (Trask 2004: 77). This is an important insight for solving the difficulties posed by *rule telescoping* (Anderson 1981: 521), when a sound shift is thought to be able to yield very diverse outputs simply because the linguistic data has been studied *post hoc* without considering diachronic changes in the inventory.

Voicing distinctions in sibilants correlate with higher numbers of sibilant PoAs

The sibilant inventories of the languages of the world tend to have no voicing distinctions (at most, allophonic voicing instead) if they have 1 or 2 sibilant PoAs. The more PoAs for sibilants an inventory has, the more likely it is to exhibit phonemic voicing contrasts in sibilants. This was observed in 3.2.1.3 and was statistically significant ($p < 0.001$) in each of the UPSID, PHOIBLE and SibInv databases. An inventory that has precisely two sibilant phonemes was found to be more likely to have a PoA contrast (e.g. /s - ʃ/) than to have a voicing contrast (/s̪ - z̪/), in accordance with this universal tendency.

A tentative explanation for this observation was formulated as follows: in smaller inventories, a place contrast between 2 PoAs can provide a large acoustic space to distinguish the two categories in perceptual space (e.g. in its widest possibility, the contrast can be realised as the extremely distinct [s̪ - ʂ̺]). As the number of PoAs increases, the available acoustic space for each category decreases (Boersma and Hamann 2008; Jaggers and Baese-Berk 2019), so that adding a PoA contrast becomes less and less efficient in terms of auditory distance. However, the auditory distance between a voiced and voiceless sibilant (e.g. [s - z]) does not increase or decrease in function of the number of PoAs. In a Northwest Caucasian language with 4 sibilant PoAs, the /s̪ - z̪/ contrast has no reason *a priori* to be any different from the /s - z/ contrast in e.g. any Romance language with 2 sibilant PoAs.

Excursus: the `Shibboleth' mispronunciation as 2s-and-1s-interference

The word *Shibboleth* has become known as a term for a linguistic or cultural trait distinguishing a social group from another, often without the control or intention of the social group in question. For example, if people have a peculiar way of pronouncing /r/

and cannot prevent it, they are likely to be identified by others by means of a *Shibboleth* implying the use of /r/ (e.g. the word 'Brian'). The word *Shibboleth* finds its origin in the following passage of the Bible (King James Bible 1769: Judges 12:4–6):

> 4Then Jephthah gathered together all the men of Gilead, and fought with Ephraim: and the men of Gilead smote Ephraim, because they said, Ye Gileadites are fugitives of Ephraim among the Ephraimites, and among the Manassites.
> 5And the Gileadites took the passages of Jordan before the Ephraimites: and it was so, that when those Ephraimites which were escaped said, Let me go over; that the men of Gilead said unto him, Art thou an Ephraimite? If he said, Nay;
> 6Then said they unto him, Say now Shibboleth: and he said Sibboleth: for he could not frame to pronounce it right. Then they took him, and slew him at the passages of Jordan: and there fell at that time of the Ephraimites forty and two thousand.

The Gileadites, who are visibly able to pronounce and distinguish the words *Sibboleth* and *Shibboleth*, win a war against the Ephraimites and manage to kill them in their escape by identifying them by means of their pronunciation of the word. This passage has been the object of much debate with respect to the phonetic nature of the first segment in these different pronunciations. In their comment on this passage, Hess et al. (2016) note the following:

> The Gileadites plan to identify Ephraimites involves a dialectical *shin/samek* contrast. The origin of the variation is not clear, though the issue seems to be phonetic rather than phonemic. This is probably not a case of divergent development of sibilants in Gileadite and Ephraimite dialects of Hebrew, but simply a case of differentiation in the pronunciation of the same sibilant in these regions. To the Gileadites, the Ephraimite pronunciation of *shin* sounded exactly like a *samek*. Accordingly, whenever the Gileadites demanded an Ephraimite to say "Shibboleth", he would have betrayed his origin by saying "Sibboleth". Extrabiblical evidence for the persistence of this phenomenon is provided by a sixth-century B.C. Ammonite seal found at Tell el-ʿUmeiri bearing the inscription, "[belonging] to Milkom-ʿ ûr, servant of Baʿalyšaʿ." Scholars are agreed [sic] that the Baʿalyšaʿ named here is to be identified with Baʿalis reffered to in Jeremiah 40:14. As in our text, the latter spelling reflects the way a Cisjordanian speaker would have heard and reproduced the Transjordanian *shin*.

To summarise: Ephraimites and Gileadites live on two different sides of the river Jordan, and the Ephraimites are Cisjordanian while the Gileadites are Transjordanian. The Gileadites realise a contrast between *Sibboleth* and *Shibboleth*, whereas the Ephraimites pronounce *Shibboleth* as something that sounds like *Sibboleth*. The Ammonites are Transjordanian like the Gileadites, and the /š/ in their inscription corresponds to the /s/ of

Cisjordanian speakers, precisely as was the case for the Gileadite *Shibboleth* (Hendel 1996: 71).

The literature has long been divided on the question of the phonetic nature of these segments (see e.g. Faber 1992, Hendel 1996, Woodhouse 2003). The most widespread interpretations of the facts consider that it must have been a difference in two sibilant fricatives (or two sibilant affricates, which would not have changed anything, since the key difference is in the sibilant part of the segments). However, the literature misses a fundamental fact: the possible existence of retracted alveolar /s̠/ besides the well-known /s/ and /ʃ/ (for example, Hendel 1996: 72 notes that the solution probably lies "in the phonetic range of the (voiceless) palatal, alveolar, and palato-alveolar affricates and fricatives", adding that "[i]t is difficult to be more precise"). As a hint in favour of a retracted alveolar pronunciation, Woodhouse (2003: 277) observes that in the transcription of Ancient Hebrew words by Egyptian scribes, there was:

> [A] fairly even-handed vacillation between <š> and <s> (...), except in some environments where the sibilant was in contact with a high vowel or a reflex of one of the corresponding semivowels. In these environments the writing of Egyptian <š> was much more heavily favored over <s> than in other environments.

This is strongly reminiscent of the loanword adaptation patterns between well-dispersed /s - ʃ/ contrasts (in this case, Egyptian) and 1s with a single [s̠] (in this case, Ephraim Hebrew) described in this dissertation (see 3.2.2.2). Adams (1975) contains plenty of examples of this kind: in the ears of speakers of a well-dispersed 2s, [s̠] sounds sometimes as /s/ and sometimes as /ʃ/, more often as /ʃ/ in 'postalveolar-friendly contexts' such as in contact with high vowels or glides. In the ears of speakers of a 1s, [s] and [ʃ] might sound different or identical, but articulation will merge them to [s̠] anyway.

There is thus enough evidence to support the following hypothesis: the Transjordanian Gileadites have a sibilant phoneme that the Cisjordanian Ephraimites do not have and cannot pronounce faithfully; assuming that 1) the dialectal differences involved are strictly sibilantic (as do Hess et al. 2016 and much of the literature), that 2) it is very unlikely that a contrast between 4 sibilant PoAs existed in Gileadite at the time (since 4s constitute about 1% of attested modern languages in the segment databases and Proto-Semitic is not reconstructed with more than 2 sibilant PoAs), and that 3) it is highly improbable that any of both languages was sibilantless, only these possibilities are left:

1. The Gileadites have an AB sibinv, the Ephraimites a B, BC or AC sibinv;
2. The Gileadites have a BC sibinv, the Ephraimites a B, AB or AC sibinv;
3. The Gileadites have an AC sibinv, the Ephraimites a B, AB or BC sibinv;
4. The Gileadites have an ABC sibinv, the Ephraimites a B, AB, BC or AC sibinv.

All logically possible combinations with 1s, 2s or 3s are listed in 8.7 in the appendix. The 13 combinations that are compatible with the *Shibboleth* story are still numerous, but evidence from the statistical frequency of the respective inventories helps to discern the most probable explanation. In 8.7, I cross the probabilities for two given types of inventory to exist based on data from the SibInv database, and it turns out that the AC - B combination is by far the most likely based on statistical frequency (11.63%). This is due to the fact that AC and B languages are the most frequent and secondmost frequent types of sibilant inventories in the world, respectively. The secondmost likely combination is ABC - AC, with 4.84% i.e. not even half as likely as the AC - B combination. It is thus most probable that the Gileadites could distinguish [s]*ibboleth* from [ʃ]*ibboleth* but the Ephraimites would have struggled to pronounce sibilants other than [s̺]*ibboleth*, and less than half that probable that Gileadites could distinguish [s]*ibboleth* from [s̺]*ibboleth* and [ʃ]*ibboleth* but the Ephraimites could only articulate [s]*ibboleth* and [ʃ]*ibboleth*. In the former case, the reason for the death of the Ephraimites would have been the difficulty to pronounce a [s] or [ʃ] that Gileadites would perceive as a sufficiently distinct and prototypical [s] or [ʃ], and in the latter, it would have been the inability to pronounce a clear [s̺] perceived as /s̺/.

Although both possibilities are likely, evidence from both the Egyptian scribes and the dialectal varieties of Hebrew might speak in favour of the Gileadite AC - Ephraimite B possibility. Geographically, Egypt is closer to Cisjordania than Transjordania, and if the Cisjordanians had had AC where the Transjordanians had ABC, the number of Egyptian confusions would have been expected to be much lower given that they would share the AC pattern with the Cisjordanians and be much less exposed to Transjordanian /s̺/. If the Cisjordanian Ephraimites had only /s̺/ and Egypt one one side as well as Transjordania on the other side had AC sibilant inventories, the confusion in Egyptian texts would have been expected to be as frequent as it actually was.

Secondly, within Hebrew varieties, sibilant mergers have been frequent: all at least either merge ס (Proto-Semitic */s/) with שׂ (Proto-Semitic */ɬ/), or שׂ (Proto-Semitic */ɬ/) with שׁ (Proto-Semitic */ʃ/ and */θ/; Morag 2007). Some have merged even more sibilants, as Khan et al. (2013: 558) report:

> No distinction was made between the articulation of ס (and שׂ) and שׁ in the northeastern Ashkenazi communities, with some local variations. Also in some Moroccan communities, the articulation of both graphemes lay "between [s] and [š]" (Morag 2007: 558).

All these mergers make it more likely that the Gileadites would have had AC and the Ephraimites B, rather than the former ABC and the latter AC.

To summarise, although it is a difficult task to reconstruct the pronunciation of sibilants from thousands of years ago, this dissertation can contribute to this debate with the recognition of retracted alveolars as a distinct category, insights from the typological diversity and frequency of sibilant inventory patterns, and loanword adaptations witnessed between one type and another. One will thus probably never know with 100% certainty how the Ephraimites and Gileadites pronounced *Shibboleth*, but one does now have strong arguments to make us bet on the AC - B combination, in which any kind of [s/ʃ]*ibboleth* becomes a [s̺]*ibboleth*. To conclude with a humorous note, if the famous *Shibboleth* story were to occur today, e.g. in northern Italy, it would thus be most likely that it would occur as follows, recalling the Veronese Venetian dialogue that was an opening quote of this dissertation:[1]

> *And it came to pass in those days that the Tuscans rose up with an army against the Venetians who were with them in their land, and YHWH gave them their enemies into their hand, and the Venetians fled towards Veneto. And behold, the Po river had swollen full with a lot of water, and there were few bridges left to cross towards the north, that is towards Veneto, and the Tuscans controlled these bridges. Then to anybody who wanted to cross the bridge they said "Are you a Venetian?", and the person would answer "No", and then they would say: "Please say* sciopero*". And the Venetians would say* sopero *and the Tuscans would know it was a Venetian, for he could not say* sciopero*, and they would strike him with a blow so as to kill him. And it happened that forty thousand men among the Venetians died because of this.*

6.4 Conclusion

This dissertation had the aims of 1) centralising existing knowledge about sibilants in a comprehensive account of their phonetic (section 2.1), phonological (section 2.2) and sociolinguistic (section 2.4) nature and behaviour; 2) describing the ways in which sibilants pattern to constitute a range of attested inventories (section 3.2) and evolve within them by means of sound shifts (section 3.3). After this descriptive part, it aimed to explain what had been observed by 3) identifying, arguing in favour of and demonstrating the existence of a set of 8 principles that impact the structures of sibilant inventories (chapter 4); integrate these principles into phonetic-phonological models within several linguistic frameworks, and analyse how accurately these models predicted the existence of a certain number of sibilant inventories (chapter 5). A comparison of the logically possible, predicted and attested sibilant inventories in section 6.1 of this conclusion indicated that the OT + NN sibilant typology proposed in section 5.4 is able to predict correctly all attested types of sibilant inventories, and that it manages to avoid both the risk of overgeneration and that of undergeneration. The main other outcomes of this dissertation have been summarised in this chapter, outcomes that range from the very central aspect that the sibilant typology is not only valuable for its predictive power but also for its explanatory power (i.e. showing transparently *why* sibilant inventories are shaped the way they are), to the rather peripheral argumentation about *Shibboleth* that follows from the central findings of this dissertation.

Synthesising existing and new findings about sibilants and sibilant inventories has not been a simple task, especially considering the wide variety of sibilants and descriptions of sibilants in the literature. This correlated with a rather large number of pages (279), resulting from the choice of comprehensiveness over conciseness. I thus formulate the hope that, even if one will never know everything about sibilants, this dissertation

[1] Other possible modern parallels include Hollandic vs. southern speakers of Dutch, Castilian vs. American speakers of Spanish, Danes vs. Norwegians etc.

contributed to approaching this goal and will allow those who want to approach this goal to find relevant insights more easily. Still much of the particular aspects of sibilants is yet to be explored, from the triggering mechanisms for preconsonantal s-retraction to enhanced methods for the acoustic analysis of sibilant spectra or the reconstruction of historical sibilant inventories. Importantly, the more progress is made in one of these topics, the more it can help progress in other related topics (e.g. a better understanding of s-retraction helps to better understand a historical sibilant inventory). In the gigantic puzzle of sibilants, being able to place one piece of the puzzle contributes to seeing where another should be placed, which is even more efficient when one is aware of the large existing variety of puzzle pieces.

List of Figures

List of Tables

7. References

Abbi, Anvita (2013). *A Grammar of the Great Andamanese Language*. Leiden, New York, Köln. Brill.

Abylkasymova, Mairam and Jumabaeva, Gulaim (1997). *Kyrgyz Language Manual*. Bishkek. Peace Corps Kyrgyzstan.

Adam, Hisham (2013). 'An Acoustical Study of the Fricative /s/ in the Speech of Palestinian-speaking Broca's Asphasics – Preliminary Findings'.

Adams, Douglas Q. (1975). 'The Distribution of Retracted Sibilants in Medieval Europe'. *Language* **51**(2), 282–292.

Adell, Eric J. (2016). 'Dynamic properties of noise: Acoustics of fricative, affricate, and ejective sibilants of Chajul Ixil'. Draft for the dissertation Adell (2019).

Adell, Eric J. (2018). 'Sibilant fricatives of Chajul Ixil: acoustic evidence for contrastive dynamic frication'. Draft for the dissertation Adell (2019).

Adell, Eric J. (2019). The phonetics, phonology, and morphology of Chajul Ixil. PhD thesis. University of Texas at Austin.

Agostiniani, Luciano (2014). The Etruscan Language. *In:* Jean MacIntosh Turfa (ed.): 'The Etruscan World'. London. Routledge. pp. 929–968.

Al-Deaibes, Mutasim (2016). The phonetics and phonology of assimilation and gemination in Rural Jordanian Arabic. PhD thesis. The University of Manitoba.

Alamer, Asma (2018). Phonological Acquisition and development in Arabic-English Bilingual Children. PhD thesis. University of London.

Alber, Birgit (2001). 'Regional Variation and Edges: Glottal Stop Epenthesis and Dissimilation in standard and Southern Varieties of German'. *Zeitschrift für Sprachwissenschaft* **20**(1), 3–41.

Alber, Birgit (2014). 'Microvariation inside typological space'. Presentation at the Old World Conference in Phonology (OCP) 11, Leiden/Amsterdam, 24th January 2014.

Alber, Birgit (2020). 'Vinko – Ein Projekt'. Slides on the VinKo project (Zago et al. 2021) for a lesson of the course *Linguistik des Deutschen/Vergleichende Grammatik*, Free University of Bozen-Bolzano.

Alber, Birgit, DelBusso, Natalie and Prince, Alan (2016). 'From intensional properties to universal support'. *Language* **92**(2), e88–e116.

Alber, Birgit and Kokkelmans, Joachim (in prep.). ‘Germanisch-romanischer Sprachkontakt im Sibilantensystem Tiroler und Trentiner Dialekte’. Abstract for the IGDD 2021 conference in Salzburg.

Alber, Birgit, Kokkelmans, Joachim and Rabanus, Stefan (2021). *S*-retraction in the Alps: Germanic, Romance, Slavic. *In:* Thomas Stolz (ed.): ‘Sprachtypologie und Universalienforschung (STUF - Language Typology and Universals)’. Vol. 74. Berlin/New York. De Gruyter. pp. 17–38.

Alber, Birgit and Meneguzzo, Marta (2016). Germanic and Romance onset clusters - how to account for microvariation. *In:* Ermenegildo Bidese, Federica Cognola and Manuela Caterina Moroni (eds.): ‘Theoretical Approaches to Linguistic Variation’. Vol. 234 of *Linguistik Aktuell/Linguistics Today*. Amsterdam. John Benjamins. pp. 25–52.

Alber, Birgit and Prince, Alan (2016). ‘Outline of Property Theory’. Ms.

Alber, Birgit and Rabanus, Stefan (2017). Die Sibilanten des Zimbrischen: Konservativität durch Sprachkontakt. *In:* Stefan Rabanus (ed.): ‘Deutsch als Minderheitensprache in Italien: Theorie und Empirie kontaktinduzierten Sprachwandels’. Vol. 239-240 of *Germanistische Linguistik*. Hildesheim. Georg Olms Verlag. pp. 19–47.

Albright, Adam (2004). ‘The Emergence of the Marked: Root-Domain Markedness in Lakhota’. Handout of a paper presented at the 78th Annual Meeting of the Linguistic Society of America, Boston.

Alderete, John (2004). Dissimilation as Local Conjunction. *In:* John J. McCarthy (ed.): ‘Optimality Theory in Phonology’. Blackwell Publishing Ltd. pp. 394–406.

Alderete, John and Bradshaw, Mark (2012). ‘Samoan grammar synopsis’. Manuscript, Simon Fraser University.

Alderete, John and Kochetov, Alexei (2017). ‘Integrating sound symbolism with core grammar: The case of expressive palatalization’. *Language* **93**(4), 731–766.

Allen, Dana Lynne (2002). The History of the Sibilants of Peninsular Spanish from the Eleventh to the Sixteenth Centuries. PhD thesis. Queen Mary University of London.

Alvarez Enparantza, José Luis (Txillardegi) (1982). ‘Some acoustic data about the three Basque sibilants’. *Primera Conferencia Internacional Vasca en Norteamérica* **1**, 18–34.

Alvord, Scott, Echávez-Solano, Nelsy and Klee, Carol (2005). ‘La (r) asibilada en el español andino: un estudio socio lingüístico’. *Lexis* **29**(1), 27–45.

Amayreh, Mousa M. and Dyson, Alice T. (1998). ‘The Acquisition of Arabic Consonants’. *Journal of Speech, Language, and Hearing Research* **41**(3), 642–653.

Ambrazas, Vytautas (2006). *Lithuanian grammar*. Vilnius. Baltos lankos.

Anderson, Stephen R. (1981). ‘Why phonology isn’t 'natural'’. *Linguistic Inquiry* **12**(4), 493–539.

Anjos, Ivo, Grilo, Margarida, Ascensão, Mariana, Guimarães, Isabel, Magalhães, João and Cavaco, Sofia (2018). A Model for Sibilant Distortion Detection in Children. *In:* Association for Computing Machinery (ed.): 'Proceedings of the 2018 International Conference on Digital Medicine and Image Processing - DMIP '18'. International Conference Proceeding Series (ICPS). New York. ACM Press.

Anttila, Arto (1995). 'Deriving variation from grammar: A study of Finnish genitives'. *Rutgers Optimality Archive* **63**.

Anttila, Arto and Cho, Young-mee Yu (1998). 'Variation and change in optimality theory'. *Lingua* **104**(1-2), 31–56.

Applebaum, Ayla and Gordon, Matthew (2013). A comparative phonetic study of the Circassian languages. *In:* Chundra Cathcart, Shinae Kang and Clare S. Sandy (eds.): 'Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society: Special Session on Languages of the Caucasus'. pp. 3–17.

Arakin, Vladimir Dmitrievich (1973). *Самоанский язык (The Samoan Language)*. Moscow. Nauka.

Árnason, Kristján (2011). *The phonology of Icelandic and Faroese*. Oxford. Oxford University Press.

Aronson, Howard I. (1990). *Georgian: A Reading Grammar*. 2 edn. Columbus, OH. Slavica.

Arvaniti, Amalia (2010). A (brief) overview of the phonetics and phonology of Cypriot Greek. *In:* Andreas Voskos, Dionysis Goutsos and Amalia Moser (eds.): 'Η ελληνική γλώσσα στην Κύπρο από την αρχαιότητα ως σήμερα (The Greek Language in Cyprus from Antiquity to Today)'. Athens. National and Kapodistrian University of Athens. pp. 107–124.

As-Sammer, Mohammed Ahmed Abdul Sattar (2011). 'Phonological Variation in Modern Standard Arabic: The Case of the Affricate /ʤ/: Oman as a Sample'. *Journal of Basra researches for Human Sciences* **36**(4), 29–57.

Asu, Eva Liina, Ewald, Otto and Schötz, Susanne (2015). 'The realisation of sj-and tj-sounds in Estonian Swedish: some preliminary results'. *Working Papers Lund University* **55**, 23.

Asu, Eva Liina and Teras, Pire (2009). 'Estonian'. *Journal of the International Phonetic Association* **39**(3), 367–372.

Avesani, Cinzia, Galatà, Vincenzo, Best, Catherine, Vayra, Mario, Di Biase, Bruno and Ardolino, Fabio (2017). 'Phonetic details of coronal consonants in the Italian spoken by Italian-Australians from two areas of Veneto'. *Fattori sociali e biologici nella variazione fonetica* **3**, 281–306.

Aviles, Arthur J. (2008). The Phonology and Morphology of the Dar Daju Daju Language. Master's thesis. University of North Dakota.

Baker, Gary Kenneth (2004). Palatal Phenomena in Spanish Phonology. PhD thesis. University of Florida.

Baker, Philip and Kriegel, Sibylle (2013). Mauritian Creole. *In:* Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.): 'The survey of pidgin and creole languages'. Vol. 2. Oxford. Oxford University Press.

Baković, Eric (2006). 'The jug trade'. Blog post on the linguistic blog 'phonoloblog'.

Balise, Raymond R. and Diehl, Randy L. (1994). 'Some Distributional Facts about Fricatives and a Perceptual Explanation'. *Phonetica* **51**(1–3), 99–110.

Ball, Martin J., Howard, Sara J. and Miller, Kirk (2017). 'Revisions to the extIPA chart'. *Journal of the International Phonetic Association* **48**(2), 155–164.

Ball, Martin J., Muller, Nicole and Rutter, Ben (2014). *Phonology for Communication Disorders*. 1 edn. New York. Psychology Press.

Ball, Martin J. and Rahilly, Joan (1999). *Phonetics: The Science of Speech*. London. Routledge.

Baltazani, Mary, Kainada, Evia, Revithiadou, Anthi and Topintzi, Nina (2016). 'Vocoid-driven processes: Palatalization and glide hardening in Greek and its dialects'. *Glossa: a journal of general linguistics* **1**(1).

Baran, Dominika (2014). 'Linguistic practice and identity work: Variation in Taiwan Mandarin at a Taipei County high school'. *Journal of Sociolinguistics* **18**(1), 32–59.

Barlow, Jessica A. and Gierut, Judith A. (1999). 'Optimality Theory in Phonological Acquisition'. *Journal of Speech, Language, and Hearing Research* **42**(6), 1482–1498.

Baroni, Antonio (2014*a*). The invariant in phonology. The role of salience and predictability. PhD thesis. Università degli Studi di Padova.

Baroni, Antonio (2014*b*). 'Strength-based faithfulness and the sibilant /s/ in Italian'. *Yearbook of the Poznań Linguistic Meeting* **1**(1), 29–53.

Basbøll, Hans (2005). *The Phonology of Danish*. The Phonology of the World's Languages. Oxford. Oxford University Press.

Bateman, Nicoleta (2007). A crosslinguistic investigation of palatalization. PhD thesis. Univ. of California San Diego.

Bateman, Nicoleta (2008). Palatalization as Overlap of Articulatory Gestures: Crosslinguistic Evidence. *In:* Sarah Berson, Alex Bratkievich, Daniel Bruhn, Ramon Escamilla, Amy Campbell, Allegra Giovine, Lindsey Newbold, Marta Piqueras-Brunet, Marilola Perez and Russell Rhomieux (eds.): 'Proceedings of the Thirty-Fourth Annual Meeting of the Berkeley Linguistics Society, February 8–10, 2008'. Vol. 34. Berkeley, CA. Berkeley Linguistics Society. pp. 25–37.

Bateman, Nicoleta (2011). 'On the Typology of Palatalization'. *Language and Linguistics Compass* **5**(8), 588–602.

Baxter, William H. (1992). *A Handbook of Old Chinese Phonology*. Vol. 64 of *Trends in Linguistics. Studies and Monographs*. Berlin, Boston. Mouton De Gruyter.

Beale-Rivaya, Yasmine (2010). 'On the Relationship between Mozarabic Sibilants and Andalusian Seseo'. *Journal of Iberian Studies* **14**, 40–56.

Becker, Michael and Potts, Kathryn Flack (2011). The Emergence of the Unmarked. *In:* Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.): 'The Blackwell Companion to Phonology'. Oxford. Wiley Online Library. pp. 1–17.

Beeler, Madison S. (1970). 'Sibilant Harmony in Chumash'. *International Journal of American Linguistics* **36**(1), 14–17.

Beers, Mieke (2003). 'Klankproductieproblemen: een fonologische benadering'. *Stem-, Spraak- en Taalpathologie* **11**(4), 245–259.

Beguš, Gašper (2017). 'The ruki-rule in Vedic'. Handout of a presentation given at the 29th Annual UCLA Indo-European Conference (WeCIEC).

Bell-Berti, Fredericka (1975). 'Control of pharyngeal cavity size for English voiced and voiceless stops'. *Journal of the Acoustical Society of America* **57**, 456–461.

Belloni, Silvano (2006). *Grammatica Veneta*. Padova. Esedra Editrice. 2nd edition.

Benet, Ariadna, Cortés, Susana and Lleó, Conxita (2012). Devoicing of sibilants as a segmental cue to the influence of Spanish onto current Catalan phonology. *In:* Kurt Braunmüller and Christoph Gabriel (eds.): 'Multilingual Individuals and Multilingual Societies'. number 13 *In:* 'Hamburg Studies on Multilingualism'. Amsterdam. John Benjamins Publishing Company. pp. 391–404.

Bennett, William G. (2015). *The Phonology of Consonants: Harmony, Dissimilation and Correspondence*. Vol. 147. Cambridge. Cambridge University Press.

Bennett, William G. and Braver, Aaron (2020). 'Different speakers, different grammars: Productivity and representation of Xhosa labial palatalization'. *Phonological Data and Analysis* **2**(6).

Bennett, William G. and DelBusso, Natalie (2018). 'The typological effects of ABC constraint definitions'. *Phonology* **35**(1), 1–37.

Benware, Wilbur Alan (1996). 'Processual Change and Phonetic Analogy: Early New High German <s> <sch>'. *American Journal of Germanic Linguistics and Literatures* **8**(2), 265–287.

Berger, Hermann (1998). *Die Burushaski-Sprache von Hunza und Nager (Teil I: Grammatik)*. Neuindische Studien 13. Wiesbaden. Harrassowitz Verlag.

Berger, Jakob (1913). *Die Laute der Mundarten des St. Galler Rheintals und der angrenzenden vorarlbergischen Gebiete.* Vol. 3 of *Beiträge zur Schweizerdeutschen Grammatik.* Frauenfeld. Huber.

Berkson, Kelly Harper (2013). 'Optionality and locality: Evidence from Navajo sibilant harmony'. *Laboratory Phonology* **4**(2), 287–337.

Berns, Janine (2013). Friction between phonetics and phonology: The status of affricates. PhD thesis. Radboud Universiteit Nijmegen.

Berns, Janine (2014). 'A typological sketch of affricates'. *Linguistic Typology* **18**(3).

Bessett, Ryan M. and Colina, Sonia (2017). 'Spanish 'Depalatalization': the Synchronic, Diachronic and Perception Perspectives'. *Borealis – An International Journal of Hispanic Linguistics* **6**(1), 223.

Best, Catherine T., McRoberts, Gerald W. and Goodell, Elizabeth (2001). 'Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system'. *The Journal of the Acoustical Society of America* **109**(2), 775–794.

Best, Catherine T., McRoberts, Gerald W. and Sithole, Nomathemba M. (1988). 'Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants'. *Journal of Experimental Psychology: Human Perception and Performance* **14**(3), 345–360.

Bhat, Darbhe N.S. (1978). A General Study of Palatalization. *In:* Joseph H. Greenberg (ed.): 'Universals of Human Language'. Vol. 2. Stanford. Stanford University Press. pp. 47–92.

Billings, Curtis J., Penman, Tina M., Ellis, Emily M., Baltzell, Lucas S. and McMillan, Garnett P. (2016). 'Phoneme and Word Scoring in Speech-in-Noise Audiometry'. *American Journal of Audiology* **25**, 75–83.

Bingham, Hiram (1847). *A residence of twenty-one years in the Sandwich Islands; or, The civil, religious, and political history of those islands: comprising a particular view of the missionary operations connected with the introduction and progress of Christianity and civilization among the Hawaiian people.* Hartford, Connecticut. H. Huntington.

Bladon, Anthony, Clark, Christopher and Mickey, Katrina (1987). 'Production and perception of sibilant fricatives: Shona data'. *Journal of the International Phonetic Association* **17**(1), 39–65.

Blevins, Juliette (2004). *Evolutionary Phonology: The Emergence of Sound Patterns.* Cambridge. Cambridge University Press.

Blevins, Juliette. (2007). 'A Long Lost Sister of Proto-Austronesian?: Proto-Ongan, Mother of Jarawa and Onge of the Andaman Islands'. *Oceanic Linguistics* **46**(1), 154–198.

Blevins, Juliette and Harrison, Sheldon P (1999). 'Trimoraic feet in Gilbertese'. *Oceanic Linguistics* **38**(2), 203–230.

Boersma, Paul (1997*a*). 'How we learn Variation, Optionality, and Probability'. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* **21**, 43–58.

Boersma, Paul (1997*b*). 'Inventories in Functional Phonology'. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* **21**, 59–90.

Boersma, Paul (1998). Functional Phonology: Formalizing the interactions between articulatory and perceptual drives. PhD thesis. Amsterdam. University of Amsterdam.

Boersma, Paul (2007). 'OT Learning 2.4. Evaluation'. Entry in the Praat search manual (Boersma and Weenink 2019).

Boersma, Paul (2009). Cue constraints and their interactions in phonological perception and production. *In:* Paul Boersma and Silke Hamann (eds.): 'Phonology in Perception'. Phonology and Phonetics [PP] 15. Berlin, Boston. De Gruyter Mouton. pp. 55–110.

Boersma, Paul (2011*a*). 'Intro 3.1. Viewing a spectrogram'. Entry in the Praat search manual (Boersma and Weenink 2019).

Boersma, Paul (2011*b*). A programme for bidirectional phonology and phonetics and their acquisition and evolution. *In:* Anton Benz and Jason Mattausch (eds.): 'Bidirectional Optimality Theory'. Vol. 180 of *Linguistik Aktuell/Linguistics Today*. Amsterdam. John Benjamins Publishing Company. pp. 33–72.

Boersma, Paul (2017). 'Formulas 5. Mathematical functions'. Entry in the Praat search manual (Boersma and Weenink 2019).

Boersma, Paul, Benders, Titia and Seinhorst, Klaas (2020). 'Neural network models for phonology and phonetics'. *Journal of Language Modelling* **8**(1), 103–177.

Boersma, Paul and Chládková, Kateřina (2013). 'Detecting categorical perception in continuous discrimination data'. *Speech Communication* **55**(1), 33–39.

Boersma, Paul and Hamann, Silke (2006). Sibilant inventories in bidirectional phonology and phonetics. Handout of a talk at the 3rd Old World Conference in Phonology (OCP), Budapest.

Boersma, Paul and Hamann, Silke (2008). 'The evolution of auditory dispersion in bidirectional constraint grammars'. *Phonology* **25**(02), 217–270.

Boersma, Paul and Hamann, Silke (2009). Loanword adaptation as first-language phonological perception. *In:* Andrea Calabrese and W. Leo Wetzels (eds.): 'Loan Phonology'. Current Issues in Linguistic Theory 307. John Benjamins Publishing Company. pp. 11–58.

Boersma, Paul and Hayes, Bruce (2001). 'Empirical Tests of the Gradual Learning Algorithm'. *Linguistic Inquiry* **32**(1), 45–86.

Boersma, Paul and Weenink, David (2019). 'Praat: doing phonetics by computer'.

Bonet, Eulàlia and Lloret, Maria-Rosa (2018). 'Fricative–affricate alternations in Catalan'. *Probus* **30**(2), 215–249.

Bosch Galcerán, Laura (1987). *Avaluació del desenvolupament fonològic en nens catalanoparlants de 3 a 7 anys*. Barcelona. Promociones Publicaciones Universitarias.

Botma, Bert and van 't Veer, Marijn (2013). 'A fraction too much friction: The phonological status of voiced fricatives'. *Linguistics in the Netherlands 2013* **30**, 46–60.

Bradley, Travis G. and Delforge, Ann Marie (2006). Systemic Contrast and the Diachrony of Spanish Sibilant Voicing. *In:* Gess Randall and Deborah Arteaga (eds.): 'Historical Romance Linguistics'. Amsterdam. John Benjamins Publishing Company. pp. 19–52.

Brandão de Carvalho, Joaquim (2008). Western Romance. *In:* Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral (eds.): 'Lenition and Fortition'. Studies in Generative Grammar 99. Berlin. Mouton de Gruyter. pp. 207–233.

Braune, Wilhelm (1911). *Althochdeutsche Grammatik*. Vol. V of *Sammlung kurzer Grammatiken germanischer Dialekte*. 3 edn. Halle an der Saale. Niemeyer.

Briggs, Lucy Therina (2007). Dialectal variation in the Aymara language of Bolivia and Peru. PhD thesis. UMI Ann Arbor.

Bright, William (1978). Karok. *In:* Robert F. Heizer (ed.): 'California'. Vol. 8 of *Handbook of North American Indians*. Washington. Smithsonian Institution. pp. 180–189.

Brosseau-Lapré, Françoise and Rvachew, Susan (2018). *Introduction to Speech Sound Disorders*. San Diego, CA. Plural Publishing.

Brown, Jason (2016). 'Laryngeal assimilation, markedness and typology'. *Phonology* **33**(3), 393–423.

Brown, Jason Camy (2008). Theoretical aspects of Gitksan phonology. PhD thesis. University of British Columbia.

Bruhn, Daniel (2006). 'The Phonetic Inventory of Mong Leng'.

Buizza, Emanuela and Plug, Leendert (2012). 'Lenition, fortition and the status of plosive affrication: the case of spontaneous RP English /t/'. *Phonology* **29**(1), 1–38.

Buk, Solomija, Mačutek, Ján and Rovenchak, Andrij (2008). 'Some properties of the Ukrainian writing system'.

Bukmaier, Véronique, Harrington, Jonathan and Kleber, Felicitas (2014). 'An analysis of post-vocalic /s-ʃ/ neutralization in Augsburg German: evidence for a gradient sound change'. *Frontiers in Psychology* **5**.

Bukmaier, Véronique and Harrington, Jonathan (2016). ‘The articulatory and acoustic characteristics of Polish sibilants and their consequences for diachronic change’. *Journal of the International Phonetic Association* **46**(3), 311–329.

Bundgaard-Nielsen, Rikke L., Baker, Brett J., Kroos, Christian H., Harvey, Mark and Best, Catherine T. (2015). ‘Discrimination of Multiple Coronal Stop Contrasts in Wubuy (Australia): A Natural Referent Consonant Account’. *PLOS ONE* **10**(12), e0142054.

Bybee, Joan L. (2001). *Phonology and Language Use*. Cambridge. Cambridge University Press.

Bybee, Joan L. (2007). Diachronic Linguistics. *In:* Dirk Geeraerts and Hubert Cuyckens (eds.): ‘The Oxford handbook of Cognitive Linguistics’. Oxford. Oxford University Press.

Bøe, Per Sigmund S. (2015). ‘Sj-lyden på Vestlandet: 100 år med endringer: en historisk-dialektologisk undersøkelse av talemålsmateriale fra ti steder på Vestlandet’. *Maal og Minne* **1**, 37–60.

Cairns, Charles E. and Feinstein, Mark H. (1982). ‘Markedness and the Theory of Syllable Structure’. *Linguistic Inquiry* **13**(2), 193–225.

Calamai, Silvia (2016). Tuscan between standard and vernacular: A sociophonetic perspective. *In:* Massimo Cerruti, Claudia Crocco and Stefania Marzo (eds.): ‘Towards a new standard: Theoretical and empirical studies on the restandardization of Italian’. Berlin, New York. De Gruyter. pp. 213–241.

Canalis, Stefano (2017). ‘The voiced and voiceless outcomes of intervocalic -SJ- in Old Tuscan’. *Quaderni di Linguistica e Studi Orientali* **3**, 157–182.

Canfield, D. Lincoln (1981). *Spanish pronunciation in the Americas*. University of Chicago Press.

Capliez, Marc (2016). Acquisition and learning of English phonology by French speakers: on the roles of segments and suprasegments. PhD thesis. Université Lille 3.

Catford, John C. and UC Berkeley (2020). ‘Ubykh’.

Catford, John Cunnison (1977). *Fundamental Problems in Phonetics*. Indiana University Press.

Ćavar, Małgorzata (2011). Merger of the place contrast in the posterior sibilants in Croatian. *In:* Peter Kosta and Lilia Schürcks (eds.): ‘Formalization of Grammar in Slavic Languages. Potsdam Linguistic Investigations’. Frankfurt am Main. Peter Lang. pp. 29–41.

Celata, Chiara (2006). Analisi dei processi di retroflessione delle liquide in area romanza, con dati sperimentali dal còrso e dal siciliano. Tesi di Perfezionamento non pubblicata.

Cercignani, Fausto (1983). ‘The Development of /k/ and /sk/ in Old English (To the Memory of Giorgio Dolfini)’. *The Journal of English and Germanic Philology* **82**(3), 313–323.

Chafe, Wallace (2015). *A Grammar of the Seneca Language*. Berkeley. University of California Press.

Chiochetti, Armin (2016). Fonetiska och fonologiska ljudförändringar i ladinska i Val di Fassa. Master's thesis. Uppsala University.

Chirikba, Viacheslav A. (2003). *Abkhaz*. Vol. 119 of *Languages of the World/Materials*. München, Newcastle. Lincom Europa.

Chitoran, Ioana (2002). *The Phonology of Romanian*. Berlin, New York. De Gruyter.

Chládková, Kateřina (2014). Finding Phonological Features in Perception. PhD thesis. Amsterdam. University of Amsterdam.

Cho, Young-mee Yu (1998). Language change as reranking of constraints. *In:* 'Historical Linguistics 1995'. John Benjamins Publishing Company. p. 45.

Cho, Young-mee Yu (2016). Korean Phonetics and Phonology. *In:* Mark Aronoff (ed.): 'Oxford Research Encyclopedia of Linguistics'. Oxford. Oxford University Press.

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA. MIT Press.

Chomsky, Noam (2001). 'Beyond explanatory adequacy'. *MIT Occasional Papers in Linguistics* **20**.

Chomsky, Noam and Halle, Morris (1968). *The Sound Pattern of English*. New York. Harper and Row.

Chuang, Yu-Ying and Fon, Janice (2017). 'On the dialectal variations of voiced sibilant /dz/ in Taiwan Min young speakers'. *Lingua Sinica* **3**(1), 1–34.

Clark, Larry (1998). *Turkmen Reference Grammar*. Vol. 34 of *Turcologica*. Wiesbaden. Otto Harrassowitz Verlag.

Clayards, Meghan, Niebuhr, Oliver and Gaskell, M. Gareth (2014). 'The time course of auditory and language-specific mechanisms in compensation for sibilant assimilation'. *Attention, Perception, & Psychophysics* **77**(1), 311–328.

Clements, G. Nick (1985). 'The geometry of phonological features'. *Phonology Yearbook* **2**(1), 225–252.

Clements, G. Nick (1990). The role of the sonority cycle in core syllabification. *In:* 'Papers in Laboratory Phonology'. Cambridge. Cambridge University Press. pp. 283–333.

Clements, G. Nick (1999). Affricates as noncontoured stops. *In:* Osamu Fujimura, Brian D. Joseph and Bohumil Palek (eds.): 'Proceedings of LP'98'. Prague. Karolinum Press. pp. 271–299.

Clements, G. Nick (2006). Does sonority have a phonetic basis? Comments on the chapter by Bert Vaux. *In:* Eric Raimy and Charles E. Cairns (eds.): 'Contemporary Views on Architecture and Representations in Phonological Theory'. Current Studies in Linguistics. Cambridge, MA. MIT Press.

Clements, George N (2003). 'Feature economy in sound systems'. *Phonology* **20**(3), 287–333.

Clements, George N. (2004). 'Features and Sound Inventories'. Handout from the Symposium on Phonological Theory: Representations and Architecture (CUNY).

Clements, George N. (2009). 'The role of features in phonological inventories'. *Contemporary views on architecture and representations in phonological theory* pp. 19–68.

Coetzee, Andries W (2002). 'Between-language frequency effects in phonological theory'. *Rutgers Optimality Archive* **973**.

Colantoni, Laura and Kochetov, Alexei (2011). An Articulatory Study of Sibilant Fricatives in Two Spanish Varieties. *In:* Scott M. Alvord (ed.): 'Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology'. Somerville, MA. Cascadilla Proceedings Project. pp. 84–97.

Colarusso, John (1999). 'A North West Caucasian Reader: phonemicized, glossed, and translated texts in East Circassian (Kabardian), West Circassian (Bzhedukh), Ubykh, Abaza (Tapanta), Abzhwi Abkhaz and Bzyb Abkhaz'.

Collins, Beverley and Mees, Inger (1999). *The phonetics of English and Dutch*. Brill.

Conrad, François (2018). Der Zusammenfall von /ʃ/ > /ɕ/ < /ç/ im Luxemburgischen. *In:* 'Proceedings of the Conference on Phonetics & Phonology in German-speaking countries'. Vol. 13. pp. 29–32.

Creissels, Denis (1999). Remarks on the Sound Correspondences Between Proto-Bantu and Tswana (S31), with a Particular Attention to Problems Involving* j (or* y),* i and Sequences* NC. *In:* Jean-Marie Hombert and Larry M. Hyman (eds.): 'Bantu Historical Linguistics: Theoretical and Empirical Perspectives'. Vol. 99 of *CSLI Lecture Notes*. Stanford, CA. Center for the Study of Language and Information.

Cresci, Michela (2014). The Sound Patterns Of Camuno: Description And Explanation In Evolutionary Phonology. PhD thesis. City University of New York (CUNY).

Cristià, Alejandrina, McGuire, Grant L., Seidl, Amanda and Francis, Alexander L. (2011). 'Effects of the distribution of acoustic cues on infants' perception of sibilants'. *Journal of Phonetics* **39**(3), 388–402.

Crystal, David (2008). *A dictionary of Linguistics and Phonetics*. The Language Library. 6 edn. Oxford. Wiley-Blackwell.

Cunningham, Margaret C. (1969). ‘A Description of the Yugumbir Dialect of Bandjalang’. *University of Queensland Papers (Arts Series)* **1**(8), 69–122.

Czaplicki, Bartlłomiej, Żygis, Marzena, Pape, Daniel and Jesus, Luis M. T. (2016). ‘Acoustic Evidence of New Sibilants in the Pronunciation of Young Polish Women’. *Poznan Studies in Contemporary Linguistics* **52**(1), 1–42.

Dal Negro, Silvia (2004). *The Decay of a Language: The Case of a German Dialect in the Italian Alps*. Bern. Peter Lang CH.

Dalbor, John B. (1980). ‘Observations on Present-Day Seseo and Ceceo in Southern Spain’. *Hispania* **63**.

Dankovičová, Jana (1997). ‘Czech’. *Journal of the International Phonetic Association* **27**(1-2), 77–80.

Dannenbring, Gary L. (1980). ‘Perceptual Discrimination of Whispered Phoneme Pairs’. *Perceptual and Motor Skills* **51**(3), 979–985.

Dart, Sarah N. (1991). Articulatory and acoustic properties of apical and laminal articulations. PhD thesis. UCLA: Department of Linguistics.

Dart, Sarah N. (1993). ‘Phonetic Properties of O'odham Stop and Fricative Contrasts’. *International Journal of American Linguistics* **59**(1), 16–37.

Dart, Sarah N. (1998). ‘Comparing French and English coronal consonant articulation’. *Journal of Phonetics* **26**(1), 71–94.

Davis, Stuart (1990). ‘Italian onset structure and the distribution of *il* and *lo*’. *Linguistics* **28**(1), 43–55.

de Lacy, Paul V. (2002). The Formal Expression of Markedness. PhD thesis. University of Massachusetts Amhurst.

de Lacy, Paul V. (2004). ‘Markedness conflation in Optimality Theory’. *Phonology* **21**(2), 145–199.

de Lacy, Paul V. (2006). *Markedness: Reduction and Preservation in Phonology*. number 112 *In:* ‘Cambridge Studies in Linguistics’. Cambridge. Cambridge University Press.

de Lacy, Paul V. (2011). Markedness and Faithfulness Constraints. *In:* Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.): ‘The Blackwell Companion to Phonology’. Oxford. Wiley Online Library. pp. 1–22.

de Oliveira, Sanderson Castro Soares (2009). Preliminares sobre a fonética e a fonologia da língua falada pelo primeiro grupo de indios korúbo recémcontatados. Master’s thesis. Universidade de Brasília.

de Rijk, Rudolf P. G. (2008). *Standard Basque: A progressive grammar*. number 44 *In:* ‘Current studies in linguistics’. Cambridge, MA and London, England. The MIT Press.

De Schutter, Georges (1999). Het Antwerps: Een schets van zijn evolutie tussen 1898 en 1998. *In:* Joep Kruijsen and Nicoline van der Sijs (eds.): 'Honderd jaar stadstaal'. Amsterdam/Antwerp. Contact. pp. 301–315.

de Vriendt, Sera (2006). 'Mannelijke Persoonsnamen in het Brussels Vlaams'. *Taal en Tongval Themanummer* **19**, 231–243.

De Wulf, Chris and Goossens, Jan en Taeldeman, Johan (2005). 'Fonologische Atlas van de Nederlandse Dialecten (FAND)'. Language Atlas.

Del Saz, María (2019). 'Perception of speaker age and speaker origin in a sound change in progress: The case of /s/-aspiration in Andalusian Spanish'. *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* pp. 760–764.

DelBusso, Natalie R. (2018). Typological structure and properties of property theory. PhD thesis. Rutgers University.

Denton, Jeannette Marshall (2003). 'Reconstructing the articulation of Early Germanic *r'. *Diachronica, International Journal for Historical Linguistics* **20**(1), 11–43.

Dinnsen, Daniel A., Chin, Steven B., Elbert, Mary and Powell, Thomas W. (1990). 'Some Constraints on Functionally Disordered Phonologies'. *Journal of Speech, Language, and Hearing Research* **33**(1), 28–37.

Dinnsen, Daniel A. and Farris-Trimble, Ashley W. (2009). 'Developmental shifts in phonological strength relations'. *Phonological strength* pp. 113–149.

Dixit, R. Prakash and Hoffman, Paul R. (2004). 'Articulatory characteristics of fricatives and affricates in Hindi: an electropalatographic study'. *Journal of the International Phonetic Association* **34**(2), 141–159.

Dixon, Robert M.W. (2002). *Australian Languages: Their Nature and Development*. number 1 *In:* 'Cambridge Language Surveys'. Cambridge. Cambridge University Press.

Dols, Nicolau and Mansell, Richard (2017). *Catalan: an essential grammar*. 1 edn. London. Routledge.

Donaldson, Bruce (1983). *Dutch: A linguistic history of Holland and Belgium*. Leiden. Martinus Nijhoff.

Drake, Richard L., Vogl, A. Wayne and Mitchell, Adam W. M. (2009). *Gray's Anatomy for Students E-Book: With STUDENT CONSULT Online Access*. 2 edn. Philadelphia, PA. Churchill Livingstone Elsevier.

Duanmu, San (2000). *The phonology of Standard Chinese*. Oxford. Oxford University Press.

Dubois, Sylvie and Horvath, Barbara M. (1998). 'Let's tink about dat: Interdental fricatives in Cajun English'. *Language Variation and Change* **10**(3), 245–261.

Dufour, Sophie, Kriegel, Sibylle, Alleesaib, Muhsina and Nguyen, NoÃ«l (2014). ‘The perception of the French /s/-/ʃ/ contrast in early Creole-French bilinguals’. *Frontiers in Psychology* **5**.

Dunbar, Ewan and Dupoux, Emmanuel (2016). ‘Geometric Constraints on Human Speech Sound Inventories’. *Frontiers in Psychology* **7**.

Dunn, Michael John (1999). A Grammar of Chukchi. PhD thesis. Australian National University.

Dyson, Alice Tanner (1988). ‘Phonetic Inventories of 2- and 3-Year-Old Children’. *Journal of Speech and Hearing Disorders* **53**(1), 89–93.

Döhler, Christian (2018). *A Grammar of Komnzo*. Vol. 22 of *Studies in Diversity Linguistics*. Berlin. Language Science Press.

Eckman, Fred and Iverson, Gregory K. (2012). ‘The Role of Native Language Phonology in the Production of L2 Contrasts’. *Studies in Second Language Acquisition* **35**(1), 67–92.

Edwards, Jan and Beckman, Mary E. (2008). ‘Some Cross-Linguistic Evidence for Modulation of Implicational Universals by Language-Specific Frequency Effects in Phonological Development’. *Language Learning and Development* **4**(2), 122–156.

Elbert, Samuel H. (1964). ‘Hawaiian reflexes of Proto-Malayo-Polynesian and Proto-Polynesian reconstructed forms’. *The Journal of the Polynesian Society* **73**(4), 399–410.

Elhija, Dua'a Abu and Davis, Stuart (2016). On the status of derived affricates in Arabic dialects. *In:* Youssef A. Haddad and Eric Potsdam (eds.): ‘Perspectives on Arabic Linguistics XXVIII: Papers from the Annual Symposium on Arabic Linguistics, Gainesville, Florida’. Vol. 4 of *Studies in Arabic Linguistics*. John Benjamins. pp. 89–104.

Elmiger, Daniel, Barmaz, Janine and Pannatier, Gisèle (2013). ‘Cours de patois: En patois d’Évolène’. Online course. Avec le soutien de la Fondation du patois.

Emeneau, Murray B. (1984). *Toda Grammar and Texts*. Vol. 155 of *Memoirs of the American Philosophical Society*. Philadelphia. American Philosophical Society.

Emmitte, Aaron (2013). Un cadjin qui dzit bon dieu!: assibilation and affrication in three generations of Cajun male speakers. PhD thesis. Louisiana State University and Agricultural and Mechanical College. LSU Doctoral Dissertations 2513.

Epps, Patience (2010). Linguistic Typology and Language Documentation. *In:* Jae Jung Song (ed.): ‘The Oxford Handbook of Linguistic Typology’. Oxford. Oxford University Press. pp. 634–649.

Everett, Caleb (2018). ‘The similar rates of occurrence of consonants across the world's languages: A quantitative analysis of phonetically transcribed word lists’. *Language Sciences* **69**, 125–135.

Everett, Daniel L. and Kern, Barbara (2002). *Wari'*. London. Routledge.

Evers, Vincent, Reetz, Henning and Lahiri, Aditi (1998). 'Crosslinguistic acoustic categorization of sibilants independent of phonological status'. *Journal of Phonetics* **26**(4), 345–370.

Ewald, Otto (2015). 'A spectral analysis of the backing of Afrikaans /s/ in the consonant cluster /rs/'. *Proceedings of Fonetik 2015* .

Ezeizabarrena, María José and Alegría, Alaitz (2015). Early coda production in bilingual Spanish and Basque. *In:* 'The Acquisition of Spanish in Understudied Language Pairings'. John Benjamins Publishing Company. pp. 75–104.

Faber, Alice (1990). Interpretation of Orthographic Forms. *In:* Philip Baldi (ed.): 'Linguistic Change and Reconstruction Methodology'. Berlin, New York. De Gruyter Mouton.

Faber, Alice (1992). 'Second harvest: 'shibboleth' revisited (yet again)'. *Journal of Semitic Studies* **XXXVII**(1), 1–10.

Fant, Gunnar Carl Michael (1960). *Acoustic Theory of Speech Production*. 's Gravenhage. Mouton.

Favre, Christophe and Balet, Zacharie (1960). *Lexique du Parler de Savièse*. Vol. 71 of *Romanica Helvetica*. Bern. A. Francke.

Fenwick, Rhona S. H. (2011). *A Grammar of Ubykh*. Vol. 19 of *LINCOM Studies in Caucasian Linguistics*. Munich. LINCOM.

Ferguson, Charles A. (1969). 'The /g/ in Syrian Arabic: Filling a Gap in a Phonological Pattern'. *Word* **25**(1-3), 114–119.

Ferguson, Charles A. (1990). 'From esses to aitches: Identifying Pathways of Diachronic Change'. *Studies in Typology and Diachrony for Joseph H. Greenberg* pp. 59–78.

Ferreira-Silva, Audinéia, Pacheco, Vera and Cagliari, Luiz Carlos (2015). 'Statistical descriptors in the characterization of some Brazilian Portuguese fricatives: analysis of spectral moments'. *Acta Scientiarum. Language and Culture* **37**(4), 371–379.

Flego, Stefon M. and Berkson, Kelly Harper (2020). 'A Phonetic Illustration of the Sound System of Icelandic'. *IULC Working Papers* **20**(1).

Fleischer, David, Wagner, Michael and Clayards, Meghan (2013). A following sibilant increases the ambiguity of a sibilant continuum. *In:* 'Proceedings of Meetings on Acoustics'. Vol. 19. San Diego, CA. ASA.

Flemming, Edward S. (1995). Auditory Representations in Phonology. PhD thesis. University of California, Los Angeles. Published 2002, London and New York: Routledge.

Flemming, Edward S. (1996). 'Evidence for constraints on contrast: The dispersion theory of contrast'. *UCLA Working Papers in Phonology* **1**, 86–106.

Flemming, Edward S. (2002). *Auditory Representations in Phonology*. London, New York. Routledge.

Flemming, Edward S. (2004). Contrast and perceptual distinctiveness. *In:* Bruce Hayes, Robert Kirchner and Donca Steriade (eds.): ‘Phonetically Based Phonology’. Cambridge. Cambridge University Press. pp. 232–276.

Flemming, Edward S. (2008*a*). Speech Perception and Phonological Contrast. *In:* ‘The Handbook of Speech Perception’. Blackwell Publishing Ltd. pp. 156–181.

Flemming, Edward S. (2008*b*). ‘Speech Perception in Phonology’.

Flemming, Edward S. (2018). ‘Systemic markedness in sibilant inventories’. Poster presented at the Annual Meeting on Phonology (AMP) 2018, UCSD.

Fletcher, Samuel G. (1989). ‘Palatometric Specification of Stop, Affricate, and Sibilant Sounds’. *Journal of Speech, Language, and Hearing Research* **32**, 736–748.

Flynn, Darin (2012). ‘Phonology: the Distinctive Features of Speech Sounds’.

Forrest, Karen, Weismer, Gary, Milenkovic, Paul and Dougall, Ronald N. (1988). ‘Statistical analysis of word-initial voiceless obstruents: Preliminary data’. *The Journal of the Acoustical Society of America* **84**(1), 115–123.

Fougeron, Cécile and Smith, Caroline L. (1993). ‘French’. *Journal of the International Phonetic Association* **23**(2), 73–76.

Fox, Annette V. and Dodd, Barbara J. (1999). ‘The phonological acquisition of German’. *Sprache Stimme Gehoer* .

Fox, Robert Allen and Nissen, Shawn L. (2005). ‘Sex-Related Acoustic Changes in Voiceless English Fricatives’. *Journal of Speech, Language, and Hearing Research* **48**(4), 753–765.

Francis, Naomi, Ma, Jamie and Gick, Bryan (2012). ‘Aerotactile acuity as a predictor of sibilant contrast’. *Canadian Acoustics* **40**(3), 22–23.

Franklin, Karl and Franklin, Joice (1962). ‘Kewa I: Phonological Asymmetry’. *Anthropological Linguistics* **4**(7), 29–37.

Freeman, Aaron (2014). ‘The Geography of Dorsal Consonants in Syrian Arabic’. Handout, University of Pennsylvania.

Fuchs, Susanne (2005). ‘Articulatory correlates of the voicing contrast in alveolar obstruent production in German’. *ZAS Papers in Linguistics* **41**.

Fuchs, Susanne and Toda, Martine (2010). Do differences in male versus female /s/ reflect biological or sociophonetic factors?. *In:* Susanne Fuchs, Martine Toda and Marzena Żygis (eds.): ‘Turbulent Sounds’. Berlin. De Gruyter Mouton. pp. 281–302.

Fuhrmeister, Pamela (2020). Neural Correlates of Native-language Speech Perception and Non-native Speech Sound Learning. PhD thesis. University of Connecticut.

Gabbard, Kevin M. (2010). 'A Phonological Analysis of Somali and the Guttural Consonants'. Bachelor's Thesis. The Ohio State University.

Gafos, Adamantios I. (1999). *The Articulatory Basis of Locality in Phonology*. Outstanding Dissertations in Linguistics. New York. Garland.

Galmés De Fuentes, Álvaro (1962). *Las sibilantes en la Romania*. Madrid. Gredos.

Galović, Filip (2013). 'Štokavsko i čakavsko: iz jezika splitskoga dijalekatskoga pjesništva (Chakavian and Shtokavian: lingustic characteristics of poetry of local speech of Split)'. *Čakavska rič: Polugodišnjak za proučavanje čakavske riječi* **41**(1-2), 121–138.

Garbacz, Piotr (2008). 'Älvdalska – ett mindre känt nordiskt språk'.

García, Alison (2013). Allophonic variation in the Spanish sibilant fricative. PhD thesis. University of Wisconsin-Milwaukee.

Gerrits, Ellen and Schouten, Marten E. H. (2004). 'Categorical perception depends on the discrimination task'. *Perception & Psychophysics* **66**(3), 363–376.

Gerrits, Petra A. M. (2001). The categorisation of speech sounds by adults and children. PhD thesis. Utrecht University.

Gilles, Peter and Trouvain, Jürgen (2013). 'Luxembourgish'. *Journal of the International Phonetic Association* **43**(1), 67–74.

Gnanadesikan, Amalia (2004). Markedness and faithfulness constraints in child phonology. *In:* 'Constraints in Phonological Acquisition'. Cambridge. Cambridge University Press. pp. 73–108.

Goad, Heather (2012). 'sC Clusters are (almost always) coda-initial'. *The Linguistic Review* **29**(3), 335–373.

Goelzer, Abigail (2005). 'Variations in French Phonetics, Phonology, Morphology, and Syntax on the Island of Jersey'. *LSO Working Papers in Linguistics* **5**, 51–66.

Gogolewski, Stanisław (2001). Chapter 4: Dialectology in Poland, 1873-1997. *In:* E.F. Konrad Koerner and Aleksander Szwedek (eds.): 'Towards a History of Linguistics in Poland: From the early beginnings to the end of the 20th century'. Studies in the History of the Language Sciences, 102. Amsterdam. John Benjamins Publishing Company. pp. 123–146.

Goldman, Ronald, Fristoe, Macalyne and Williams, Kathleen T. (2000). *Goldman-Fristoe test of articulation*. 2 edn. Minneapolis, MN. Pearson.

Golston, Chris and Yang, Phong (2001). White Hmong loanword phonology. *In:* Caroline Féry, Anton Dubach Green and Ruben van de Vijver (eds.): 'Proceedings of HILP'. 5. Potsdam. Potsdam University. pp. 40–57.

González, Hebe (2003). A Typology of Stops in South American Indian Languages. *In:* 'Proceedings of the Conference on Indigenous Languages of Latin America'. University of Texas at Austin. pp. 1–20.

Gordon, Matthew K. (2016). *Phonological Typology*. Oxford Surveys in Phonology and Phonetics. Oxford. Oxford University Press.

Goss-Grubbs, David (2007). 'Onset Cluster Coalescence in Child Language. The case of Henry, 3;8'. Ling 452: June 6, 2007.

Goswami, Usha (2012). Phonological Representation. *In:* Norbert M. Seel (ed.): 'Encyclopedia of the Sciences of Learning'. New York. Springer. pp. 2625–2627.

Grammont, Maurice (1933). *Traité de phonétique*. Paris. Charles Delagrave.

Green, Jordan R. and Wang, Yu-Tsai (2003). 'Tongue-surface movement patterns during speech and swallowing'. *The Journal of the Acoustical Society of America* **113**(5), 2820–2833.

Green, Melanie (2006). Levels of Adequacy, Observational, Descriptive, Explanatory. *In:* Keith Brown (ed.): 'Encyclopedia of Language & Linguistics'. 2 edn. Amsterdam. Elsevier. pp. 49–51.

Greenberg, Joseph H (1963). 'Some universals of grammar with particular reference to the order of meaningful elements'. *Universals of language* **2**, 73–113.

Greenberg, Joseph H (1965). 'Some Generalizations concerning Initial and Final Consonant Sequences'. *Linguistics* **3**(18).

Greenberg, Joseph H (1966). *Universals of Language*. 2 edn. Cambridge, MA. M.I.T. Press.

Grimaldi, Mirko (2018). The phonetics-phonology relationship in the neurobiology of language. *In:* Roberto Petrosino, Pietro Cerrone and Harry van der Hulst (eds.): 'From Sounds to Structures'. De Gruyter. pp. 65–103.

Grønnum, Nina (2005). *Fonetik og Fonologi: Almen og Dansk*. Copenhagen. Akademisk Forlag.

Grossman, Eitan, Eisen, Elad, Nikolaev, Dmitry and Moran, Steven (2019). 'How different were phonological distributions?'. Presentation at SLE 52, Leipzig.

Grunwell, Pamela (1982). *Clinical phonology*. London. Croom Helm.

Gruppo 'Amis du patois' (2007). *Dizionario del dialetto francoprovenzale di Hône, Valle d'Aosta*. Aosta. Comune di Hône/Le Château.

Guerrero, Jairo (2019). Amin SO9RAT: a Moroccan Rapper from the North-Eastern Region. *In:* Catherine Miller, Alexandrine Barontini, Marie-Aimée Germanos, Jairo Guerrero and Christophe Pereira (eds.): 'Proceedings of the 12th International Conference of AIDA held in Marseille from May 30th to June 2nd 2017'. Studies on Arabic Dialectology and Sociolinguistics. Institut de recherches et d'études sur les mondes arabes et musulmans.

Gussenhoven, C. and Jacobs, H. (2017). *Understanding Phonology*. Understanding Language. Taylor & Francis.

Gussenhoven, Carlos and Aarts, Flor (1999). 'The dialect of Maastricht'. *Journal of the International Phonetic Association* **29**(2), 155–166.

Haiman, John and Benincà, Paola (1992). *The Rhaeto-Romance Languages*. Croom Helm Romance Linguistics Series. London. Routledge.

Hall, Daniel Currie (2011). 'Phonological contrast and its phonetic enhancement: dispersedness without dispersion'. *Phonology* **28**(1), 1–54.

Hall, Daniel Currie and Mielke, Jeff (2011). Distinctive Features. *In:* 'Oxford Bibliographies'. Oxford. Oxford University Press (OUP).

Hall, T. Alan (2001). *Distinctive Feature Theory*. Vol. 2. Berlin. Walter de Gruyter.

Hall, Tracy A. (2002). 'Against extrasyllabic consonants in German and English'. *Phonology* **19**(01), 33–75.

Hall, Tracy A. (2007). *Segmental Features*. Cambridge Handbooks in Language and Linguistics. Cambridge. Cambridge University Press. p. 311–334.

Hall, Tracy A. and Żygis, Marzena (2010). An overview of the phonology of obstruents. *In:* Susanne Fuchs, Martine Toda and Marzena Żygis (eds.): 'Turbulent Sounds: An Interdisciplinary Guide'. Berlin. De Gruyter Mouton. pp. 1–36.

Hall, Tracy Alan (1997*a*). 'The historical development of retroflex consonants in Indo-Aryan'. *Lingua* **102**(4), 203–221.

Hall, Tracy Alan (1997*b*). *The Phonology of Coronals*. Amsterdam. John Benjamins.

Hall, Tracy Alan (2000). 'Syllabically conditioned coalescence and deletion in Zoque: an optimality-theoretic approach'. *Linguistics* **38**.

Hall, Tracy Alan (2008). 'Middle High German [rs]>[rʃ] as height dissimilation'. *The Journal of Comparative Germanic Linguistics* **11**(3), 213–248.

Hall, Tracy Alan (2012). 'The Representation of Affricates in Cimbrian German'. *Journal of Germanic Linguistics* **24**(1), 1–22.

Hall, Tracy Alan (2013). 'Alveolopalatalization in Central German as Markedness Reduction'. *Transactions of the Philological Society* **112**(2), 143–166.

Hall, Tracy Alan and Hamann, Silke (2006). 'Towards a typology of stop assibilation'. *Linguistics* **44**(6), 1195–1236.

Hall, Tracy Alan, Hamann, Silke and Żygis, Marzena (2006). 'The phonetic motivation for phonological stop assibilation'. *Journal of the International Phonetic Association* **36**(1), 59–81.

Hall, Tracy Alan and Scott, John H. G. (2007). 'Inflectional paradigms have a base: evidence from s-Dissimilation in Southern German dialects'. *Morphology* **17**(1), 151–178.

Hamann, Silke (2003). The Phonetics and Phonology of Retroflexes. PhD thesis. Utrecht. University of Utrecht.

Hamann, Silke (2004). 'Retroflex fricatives in Slavic languages'. *Journal of the International Phonetic Association* **34**(1), 53–67.

Hannahs, S. J. (2013). *The Phonology of Welsh*. Oxford. Oxford University Press.

Hansson, Gunnar Ólafur (2003). 'Laryngeal licensing and laryngeal neutralization in Faroese and Icelandic'. *Nordic Journal of Linguistics* **26**(1), 45–79.

Hansson, Gunnar Ólafur (2010). *Consonant harmony: Long-distance Interaction in Phonology*. Vol. 145. University of California Press.

Hansson, Gunnar Ólafur (2001*a*). The phonologization of production constraints: Evidence from consonant harmony. *In:* 'Chicago Linguistic Society'. Vol. 37. p. 187.

Hansson, Gunnar Ólafur (2001*b*). Theoretical and Typological Issues in Consonant Harmony. PhD thesis. University of California, Berkeley.

Hanulíková, Adriana and Hamann, Silke (2010). 'Slovak'. *Journal of the International Phonetic Association* **40**(3), 373–378.

Hardcastle, William J. (1976). *Physiology of Speech Production: Introduction for Speech Scientists*. Academic Press Inc.

Hardman, Martha J. (2001). *Aymara*. Vol. 35 of *Lincom Studies in Native American Linguistics*. Munich. Lincom Europa.

Harlow, Ray (1996). *Māori*. Vol. 20 of *Languages of the World/Materials*. München, Newcastle. Lincom Europa.

Harms, Phillip Lee (1984). Fonología del epena pedee (saija). *In:* Linda Gawthorne, Grace Hensarling, Frances Gralow, Dorothy Cook, Carolyn Muller de Young, Jack Keels, Phillip Harms and Riena de Kondo (eds.): 'Sistemas fonológicas de idiomas colombianos'. Vol. 5. Lomalinda. Editorial Townsend.

Harrington, Jonathan (2010). *Phonetic Analysis of Speech Corpora*. Oxford. Wiley-Blackwell.

Harris, Katherine Safford (1958). ‘Cues for the discrimination of American English fricatives in spoken syllables’. *Language and Speech* **1**(1), 1–7.

Harris, Zellig (1942). ‘The phonemes of Moroccan Arabic’. *Journal of the American Oriental Society* **62**(4), 309–318.

Hary, Benjamin (1996). The Ǧīm/Gīm in Colloquial Urban Egyptian Arabic. *In:* Shlomo Izre'el and Shlomo Raz (eds.): ‘Israel Oriental Studies’. Vol. XVI of *Studies in Modern Semitic Languages*. Leiden, New York, Köln. Brill. pp. 154–168.

Hayes, Bruce (1999). Phonetically Driven Phonology. *In:* Michael Darnell, Edith A. Moravcsik, Michael Noonan, Frederick J. Newmeyer and Kathleen Wheatley (eds.): ‘Studies in Language Companion Series’. 41. John Benjamins Publishing Company. pp. 243–285.

Hayes, Bruce and Steriade, Donca (2004). Introduction: the phonetic bases of phonological Markedness. *In:* Bruce Hayes, Robert Kirchner and Donca Steriade (eds.): ‘Phonetically Based Phonology’. Cambridge University Press. pp. 1–33.

Heath, Jeffrey (2015). A Grammar of Tebul Ure (Dogon, Mali). PhD thesis. University of Michigan.

Hedrick, Mark S. and Ohde, Ralph N. (1993). ‘Effect of relative amplitude of frication on perception of place of articulation’. *The Journal of the Acoustical Society of America* **94**(4), 2005–2026.

Hendel, Ronald S. (1996). ‘Sibilants and šibbōlet (Judges 12:6)’. *Bulletin of the American Schools of Oriental Research* **301**, 69–75.

Henke, Eric, Kaisse, Ellen M. and Wright, Richard (2012). Is the Sonority Sequencing Principle an epiphenomenon?. *In:* Steve Parker (ed.): ‘The Sonority Controversy’. Vol. 18. Berlin. Mouton de Gruyter. pp. 65–100.

Hennequin, Guillaume (2017). ‘Neural networks subtract and conquer’. *eLife* **6**.

Hermo del Teso, Lucia (2017). A preliminary study of language mixture between Asturian and Spanish. Master’s thesis. University of Montana.

Herrick, Dylan (2002). Catalan phonology. *In:* Caroline R. Wiltshire and Joaquim Camps (eds.): ‘Romance Phonology and Variation: Selected papers from the 30th Linguistic Symposium on Romance Languages, Gainesville, Florida, February 2000’. Current Issues in Linguistic Theory 217. Amsterdam. John Benjamins Publishing Company. pp. 69–83.

Hess, Richard, Block, Daniel I. and Manor, Dale W. (2016). *Joshua, Judges, and Ruth*. Grand Rapids, Michigan. Zondervan Academic.

Hewitt, George (2013). *The Abkhazians: A Handbook*. Caucasus World: Peoples of the Caucasus. London. Routledge.

Hickey, Raymond (ed.) (2003). *Motives for Language Change*. Cambridge University Press.

Hoard, James E. (1975). The New Phonological Paradigm. *In:* Didier L. Goyvaerts and Geoffrey K. Pullum (eds.): 'Essays on the Sound Pattern of English'. number 1 *In:* 'Studies in the Sciences of Language Series'. Ghent. E. Story-Scientia. pp. 21–61.

Hock, Hans Henrich (2009). *Principles of historical linguistics*. Walter de Gruyter.

Hoey, Elliott Michael (2013). Grammatical Sketch of Turkmen. PhD thesis. University of California, Santa Barbara.

Höfler, Otto (1957). 'Die zweite Lautverschiebung bei Ostgermanen und Westgermanen'. *Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB)* **1957**(79), 161–350.

Hogg, Richard M (1996). Tertiary stress in Old English: Some reflections on explanatory inadequacy. *In:* 'English Historical Linguistics 1994: Papers from the 8th International Conference on English Historical Linguistics (8 ICEHL, Edinburgh, 19 23 September 1994)'. Vol. 135. John Benjamins Publishing. p. 3.

Holt, D. Eric (2007). Optimality Theory and language change in Spanish. *In:* Fernando Martínez-Gil and Sonia Colina (eds.): 'Linguistik Aktuell/Linguistics Today'. 99. Amsterdam. John Benjamins Publishing Company. pp. 378–398.

Honeybone, Patrick (2005). Diachronic evidence in segmental phonology: the case of obstruent laryngeal specifications. *In:* Marc van Oostendorp and Jeroen van de Weijer (eds.): 'The Internal Organization of Phonological Segments'. Berlin. Mouton De Gruyter. pp. 317–352.

Honeybone, Patrick (2008). What and where?. *In:* Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral (eds.): 'Lenition, weakening and consonantal strength: tracing concepts through the history of phonology'. Studies in Generative Grammar 99. Berlin. Mouton de Gruyter. pp. 9–92.

Howson, Phil (2015). An acoustic examination of the three-way sibilant contrast in Lower Sorbian. *In:* 'Proceedings of INTERSPEECH 2015, Dresden'. pp. 2670–2674.

Hu, Fang (2008). The Three Sibilants in Standard Chinese. *In:* Rudoph Sock, Susanne Fuchs and Yves Laprie (eds.): 'Proceedings of the 8th International Seminar on Speech Production (ISSP)'. Strasbourg. Université March Bloch. pp. 105–108.

Hua, Zhu and Dodd, Barbara (2006). *Phonological development and disorders in children: A multilingual perspective*. Vol. 8. Multilingual Matters.

Hualde, José Ignacio (1988). Affricates are not contour segments. *In:* Hagit Borer (ed.): 'Proceedings of the 7th West Coast Conference on Formal Linguistics'. pp. 143–157.

Hualde, José Ignacio (2004). *Basque Phonology*. London. Routledge.

Hualde, José Ignacio (2005). *The Sounds of Spanish*. Cambridge. Cambridge University Press.

Huang, Hui-chuan J. (2019). 'Glide Strengthening in Atayal: Sonority Dispersion and Similarity Avoidance'. Presentation given at the sixteenth Old World Conference in Phonology (OCP XVI) at the University of Verona.

Hunziker, Daniel A., Hunziker, Elisabeth and Eaton, Helen (2008). 'A description of the phonology of the Sandawe language'. *SIL Electronic Working Papers 2008–004* .

Hyllested, Adam (2014). Word Exchange at the Gates of Europe: Five Millennia of Language Contact. PhD thesis. Københavns Universitet, Det Humanistiske Fakultet.

Hyman, Larry M. (2008). 'Universals in Phonology'. *The Linguistic Review* **25**(1-2), 83–137.

Ibrahim, Suleiman I. O. (2019). Typology and Syntactical Description of Daju Language. PhD thesis. Sudan University of Science and Technology.

International Phonetic Association (2015). 'IPA Chart'. Available under a Creative Commons Attribution-Sharealike 3.0 Unported License.

Iskarous, Khalil (2017). 'The relation between the continuous and the discrete: A note on the first principles of speech dynamics'. *Journal of Phonetics* **64**, 8–20.

Iskarous, Khalil, Shadle, Christine H. and Proctor, Michael I. (2011). 'Articulatory–acoustic kinematics: The production of American English /s/'. *The Journal of the Acoustical Society of America* **129**(2), 944–954.

Itō, Junko and Mester, Armin (2003). Systemic Markedness and Faithfulness. *In:* Jonathan E. Cihlar, David W. Kaiser, Irene Kimbara and Amy Franklin (eds.): 'Proceedings from the Annual Meeting of the Chicago Linguistic Society'. Vol. 39. Chicago. Chicago Linguistic Society. pp. 665–689.

Jacobs, Haike and Berns, Janine (2013). Perception, Production and Markedness in Sound Change: French Velar Palatalization. *In:* Deborah L. Arteaga (ed.): 'Research on Old French: The State of the Art'. New York. Springer. pp. 107–122.

Jaggers, Zachary and Baese-Berk, Melissa (2019). 'Moments of moments: Acoustic phonetic character and within-category variability of the Basque three-sibilant contrast'. *Proceedings of the International Congress on Phonetic Sciences* pp. 1828–1832.

Jakobson, Roman (1972). *Kindersprache, Aphasie und allgemeine Lautgesetze*. Frankfurt am Main. Suhrkamp.

Jakobson, Roman, Fant, Gunnar and Halle, Morris (1952). *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates*. Cambridge, MA. The MIT Press.

Jakobson, Roman and Halle, Morris (1956). *Fundamentals of Language*. Janua Linguarum. Series Minor 1. Berlin, Boston. De Gruyter Mouton.

Jakobson, Roman and Lotz, John (1949). 'Notes on the French Phonemic Pattern'. *WORD* **5**(2), 151–158.

Jakobson, Roman Ossipovich (1968). *Child Language: Aphasia and Phonological Universals (translated by Allan R. Keiler)*. Vol. 72 of *Janua Linguarum, Series Minor*. Berlin, Boston. De Gruyter Mouton.

Janda, Richard D. and Joseph, Brian D. (2003). Reconsidering the canons of sound-change. *In:* 'Historical Linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics, Melbourne, 13-17 August 2001'. Vol. 237. John Benjamins Publishing. p. 205.

Jannedy, Stefanie and Weirich, Melanie (2016). The acoustics of fricative contrasts in two German dialects. *In:* Christoph Draxler and Felicitas Kleber (eds.): 'Proceedings of the Conference on Phonetics and Phonology in German-speaking countries (P und P 12)'. Munich. Ludwig-Maximilian University. pp. 70–73.

Jannedy, Stefanie and Weirich, Melanie (2017). 'Spectral moments vs discrete cosine transformation coefficients: Evaluation of acoustic measures distinguishing two merging German fricatives'. *The Journal of the Acoustical Society of America* **142**(1), 395–405.

Janse, Mark (2002). 'Review article: Aspects of Pontic grammar'. *Journal of Greek Linguistics* **3**(1), 203–231.

Jany, Carmen, Gordon, Matthew, Nash, Carlos M. and Takara, Nobutaka (2007). How Universal is the Sonority Hierarchy?: A Cross-Linguistic Acoustic Study. *In:* Jürgen Trouvain (ed.): 'Proceedings of the 16th International Congress of Phonetic Sciences ICPhS XVI, 6-10 August 2007, Saarbrücken'. Saarbrücken.

Jarkey, Nerida (2015). *Serial Verbs in White Hmong*. Leiden, New York, Köln. Brill.

Jarosz, Gaja (2006). Richness of the Base and Probabilistic Unsupervised Learning in Optimality Theory. *In:* 'Proceedings of the Eighth Meeting of the ACL Special Interest Group on Computational Phonology at HLT-NAACL 2006'. New York. Association for Computational Linguistics. pp. 50–59.

JASP Team (2021). 'JASP (Version 0.11.1)'. [Computer software].

Jassem, Wiktor (1995). 'The Acoustic Parameters of Polish Voiceless Fricatives: An Analysis of Variance'. *Phonetica* **52**(3), 251–258.

Jekiel, Mateusz (2012). The evolution of English dental fricatives: variation and change. Master's thesis. Adam Mickiewicz University.

Jocz, Lechosław (2015). *Studien zum obersorbischen Konsonantismus*. Szczecin. Volumina.

Johnson, Keith (2003). *Acoustic and Auditory Phonetics*. 2 edn. Oxford. Wiley-Blackwell.

Jones, Daniel (1956). *An outline of English phonetics*. 8 edn. Cambridge. Heffer.

Jongman, Allard, Wang, Yue and Kim, Brian H. (2003). 'Contributions of Semantic and Facial Information to Perception of Nonsibilant Fricatives'. *Journal of Speech, Language, and Hearing Research* **46**(6), 1367–1377.

Jongman, Allard, Wayland, Ratree and Wong, Serena (2000). 'Acoustic characteristics of English fricatives'. *The Journal of the Acoustical Society of America* **108**(3), 1252–1263.

Joos, Martin (1952). 'The Medieval Sibilants'. *Language* **28**(2), 222.

Joos, Martin (1957). *Readings in Linguistics: The Development Of Descriptive Linguistics In America Since 1925*. Washington. American Council of Learned Societies.

Juneau, Marcel (1972). *Contribution à l'histoire de la prononciation française au Québec: étude des graphies des documents d'archives*. Presses de l'Université Laval.

Kawahara, Shigeto, Godoy, Mahayana C. and Kumagai, Gakuji (2020). 'Do Sibilants Fly? Evidence from a Sound Symbolic Pattern in Pokémon Names'. *Open Linguistics* **6**(1), 386–400.

Keating, Patricia A. (1991). Coronal Places of Articulation. *In:* Carole Paradis and Jean-François Prunet (eds.): 'The special status of coronals: Internal and external evidence'. Amsterdam. Elsevier. pp. 29–48.

Kehrein, Wolfgang (2002). *Phonological Representation and Phonetic Phasing*. Linguistische Arbeiten 466. Berlin, Boston. De Gruyter.

Kemp, Renee Lorraine (2011). The Perception of German Dorsal Fricatives by Native Speakers of English. Master's thesis. University of Georgia.

Kent, Raymond D. and Moll, Kenneth L. (1969). 'Vocal-tract characteristics of the stop cognates'. *Journal of the Acoustical Society of America* **46**, 1549–1555.

Khan, Geoffrey, Bolozky, Shmuel, Fassberg, Steven, Rendsburg, Gary A., Rubin, Aaron D., Schwarzwald, Ora R. and Zewi, Tamar (2013). *Encyclopedia of Hebrew Language and Linguistics*. Vol. 3. Leiden, Boston. Brill.

Khan, Sameer ud Dowla (2010). 'Bengali (Bangladeshi Standard)'. *Journal of the International Phonetic Association* **40**(2), 221–225.

Kim, Hyunsoon (1999). 'The Place of Articulation of Korean Affricates Revisited'. *Journal of East Asian Linguistics* **8**(4), 313–347.

Kim, Hyunsoon (2004). 'Stroboscopic-Cine MRI Data on Korean Coronal Plosives and Affricates: Implications for Their Place of Articulation as Alveolar'. *Phonetica* **61**(4), 234–251.

Kim, Yuni (2002). Phonological Features: Privative or Equipollent?. Master's thesis. Harvard University.

King, Duane (1975). A Grammar and Dictionary of the Cherokee Language. PhD thesis. University of Georgia.

King James Bible (1769). *King James Bible*. Oxford. Benjamin Blayney.

Kinney, John (2004). Common, Uncommon and Rare Phonological Disorders: An OT Perspective. Master's thesis. Fairfax, VA. George Mason University.

Kirby, James and Yu, Alan C. L. (2007). 'Hidden knowledge of syllable gap wellformedness'. Presentation given at ExpOT (Experimental Approaches to Optimality Theory).

Kirchner, Robert (2001). Phonological Contrast and Articulatory Effort. *In:* Linda Lombardi (ed.): 'Segmental Phonology in Optimality Theory'. Cambridge. Cambridge University Press. pp. 79–117.

Kleine, Ane (2003). 'Standard Yiddish'. *Journal of the International Phonetic Association* **33**(2), 261–265.

Kluender, Keith R. and Walsh, Margaret A. (1992). 'Amplitude rise time and the perception of the voiceless affricate/fricative distinction'. *Perception & Psychophysics* **51**(4), 328–333.

Ko, Insung (2013). The Articulation of Korean Coronal Obstruents: Data from Heritage Speakers and Second Language Learners. PhD thesis. University of Hawaii.

Kobayashi, Masato (2004). *Historical Phonology of Old Indo-Aryan Consonants*. number 42 *In:* 'Study of Languages and Cultures of Asia and Africa Monograph Series'. Tokyo. Research Institute for Languages and Cultures of Asia and Africa.

Kochetov, Alexei (2006). 'The role of social factors in the dynamics of sound change: A case study of a Russian dialect'. *Language Variation and Change* **18**(01).

Kochetov, Alexei (2011). Palatalization. *In:* Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.): 'The Blackwell Companion to Phonology'. Oxford. Wiley Online Library. pp. 1–25.

Kochetov, Alexei (2017). 'Acoustics of Russian voiceless sibilant fricatives'. *Journal of the International Phonetic Association* **47**(3), 321–348.

Kochetov, Alexei and Alderete, John (2011). 'Patterns and scales of expressive palatalization: Experimental evidence from Japanese'. *Canadian Journal of Linguistics/Revue canadienne de linguistique* **56**(3), 345–376.

Koenig, Laura L., Shadle, Christine H., Preston, Jonathan L. and Mooshammer, Christine R. (2013). 'Toward Improved Spectral Measures of /s/: Results From Adolescents'. *Journal of Speech, Language, and Hearing Research* **56**(4), 1175–1189.

Koivulehto, Jorma (1986). Die sieverssche Regel im Lichte der germanisch-finnischen Lehnbeziehungen. *In:* 'Germanic Dialects'. John Benjamins Publishing Company. p. 249.

Kokkelmans, Joachim (2017). Die Achillesfersche einer Lautverschiebung im Vollzug. Postalveolarisierung, Retroflektierung und r-Schwund bei /rs/ in flämischen Umgangssprachen und den germanischen Sprachen, sowie deren Beziehung zu den Ausspracheartenvon /r/. Master's thesis. Ghent University.

Kokkelmans, Joachim (2019*a*). Phonology I: Segments and Distinctive Features. Presentation given at the Winter School in Linguistics in Verona.

Kokkelmans, Joachim (2019*b*). 'A typological model of sibilant inventories and the principles which shape them'. Poster presented at the 27th Manchester Phonology Meeting (MfM).

Kokkelmans, Joachim (2020*a*). Middle High German and modern Flemish s-retraction in /rs/-clusters. *In:* Gunther De Vogelaer, Dietha Koster and Torsten Leuschner (eds.): 'German and Dutch in Contrast: Synchronic, Diachronic and Psycholinguistic Perspectives'. Vol. 11 of *Konvergenz und Divergenz*. Berlin/New York. De Gruyter. pp. 213–238.

Kokkelmans, Joachim (2020*b*). Postalveolaire klanken in het dialect van Eys en hun belang voor de taalkunde: Wat e paar wöët op et plat òs kenne vertelle övver algemeng klankwette. *In:* Pierre Bakkes, Peter Klein and Ton van de Wijngaard (eds.): 'Veldeke Jaarboek 2019'. Roermond. Vereniging Veldeke Limburg. pp. 149–173.

Kokkelmans, Joachim (in prep.). 'Sibilant outputs of Romance stop palatalisation explained: diachronic evidence and BiPhon-NN'. Ms., Free University of Bozen-Bolzano.

Kosa, Loredana Andreea (2010). Sibilant harmony: investigating the connection between typology and learnability. *In:* 'Proceedings of the 2010 annual conference of the Canadian Linguistic Association'.

Kossmann, Maarten G. and Stroomer, Harry J. (1997). Berber phonology. *In:* Alan S. Kaye and Peter T. Daniels (eds.): 'Phonologies of Asia and Africa (including the Caucasus)'. Vol. 1. Winona Lake, IN. Eisenbrauns. pp. 461–476.

Kostakis, Andrew (2017). 'Extrametricality in Latin'. *Transactions of the Philological Society* **115**(3), 336–364.

Koul, Omkar Nath (2008). *Modern Hindi Grammar*. Centreville, VA. Dunwoody Press.

Krämer, Martin and Urek, Olga (2016*a*). 'The Naturalness of Palatalization'. *Nordlyd* **42**(1), i.

Krämer, Martin and Urek, Olga (2016*b*). 'Perspectives on palatalization'. *Glossa: a journal of general linguistics* **1**(31), 1–17.

Kristoffersen, Gjert (2000). *The phonology of Norwegian*. Oxford. Oxford University Press.

Krueger, Breanna I. (2019). 'Eligibility and Speech Sound Disorders: Assessment of Social Impact'. *Perspectives of the ASHA Special Interest Groups* **4**(1), 85–90.

Kubozono, Haruo (2015). *Handbook of Japanese Phonetics and Phonology*. Vol. 2 of *Handbook of Japanese Language and Linguistics*. Berlin, Boston. De Gruyter Mouton.

Kuki, Hiroshi (1969). Tuamotuan Phonology. Master's thesis. University of Hawaii.

Kümmel, Martin (2007). *Konsonantenwandel: Bausteine zu einer Typologie des Lautwandels und ihre Konsequenzen für die vergleichende Rekonstruktion*. Wiesbaden. Reichert Verlag.

Kwakkel, Vincent (2008). Sibilants in Dutch and English: typologies, learning tasks and language change. Master's thesis. Utrecht University.

Labrune, Laurence (2012). *The phonology of Japanese*. Oxford. Oxford University Press.

LaCharité, Darlène (1993). The Internal Structure of Affricates. PhD thesis. University of Ottawa.

Ladefoged, Peter (1968). *A Phonetic Study of West African Languages*. Cambridge. Cambridge University Press. 2nd edition.

Ladefoged, Peter (1971). *Preliminaries to linguistic phonetics*. Chicago. University of Chicago Press.

Ladefoged, Peter and Maddieson, Ian (1996). *The Sounds of the World's Languages*. Oxford. Blackwell.

Ladefoged, Peter and Wu, Zongji (1984). 'Places of articulation: an investigation of Pekingese fricatives and affricates'. *Journal of Phonetics* **12**(3), 267–278.

Lahrouchi, Mohamed (2018). 'The Amazigh influence on Moroccan Arabic: Phonological and morphological borrowing'. *International Journal of Arabic Linguistics* **4**(1), 39–58.

Lamont, Andrew (2015). Progressive place assimilation in Optimality Theory. Master's thesis. Eastern Michigan University.

Lance, Donald M. and Howie, Stephen M. (1997). 'Spectrographic analysis of English phonemes and allophones'. *American Pronunciation* pp. 267–344.

Lange, Klaus-Peter (2007). 'Der galloromanische Ursprung der zweiten Lautverschiebung'. *Rask* **25**, 3–24.

Lassettre, Paul and Donegan, Patricia (1998). 'Perception in Optimality Theory: The Frugality of the Base'. *Annual Meeting of the Berkeley Linguistics Society* **24**(1), 346.

Laufer, Asher (1990). 'Hebrew'. *Journal of the International Phonetic Association* **20**(2), 40–43.

Leben, William (1973). Suprasegmental Phonology. PhD thesis. Massachussets Institute of Technology.

Ledgeway, Adam and Maiden, Martin (eds.) (2016). *The Oxford Guide to the Romance Languages*. Oxford. Oxford University Press.

Lee, Joo Kyeong (2000). 'A Functional Approach to American English Palatalization'. *The Journal of English Language and Literature* **46**(4), 989–1002.

Lee-Kim, Sang-Im, Kawahara, Shigeto and Lee, Seunghun J. (2014). 'The 'Whistled' Fricative in Xitsonga: Its Articulation and Acoustics'. *Phonetica* **71**(1), 50–81.

Lee, Wai-Sum and Zee, Eric (2010). 'Articulatory characteristics of the coronal stop, affricate, and fricative in Cantonese'. *Journal of Chinese Linguistics* **38**(2), 336–372.

Lemieux, Geneviève (2011). 'Le développement de la prononciation'. Online document. accessed on 18.02.2020.

Lev, Michael, Stark, Tammy and Chang, Will (2012). 'Phonological inventory of Abipon'. The South American Phonological Inventory Database, version 1.1.3. Berkeley: University of California: Survey of California and Other Indian Languages Digital Resource.

Li, Fangfang, Edwards, Jan and Beckman, Mary E. (2009). 'Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers'. *Journal of Phonetics* **37**(1), 111–124.

Li, Fangfang and Munson, Benjamin (2016). 'The Development of Voiceless Sibilant Fricatives in Putonghua-Speaking Children'. *Journal of Speech, Language, and Hearing Research* **59**(4), 699–712.

Li, Fangfang, Rendall, Drew, Vasey, Paul L., Kinsman, Melissa, Ward-Sutherland, Amanda and Diano, Giancarlo (2016). 'The development of sex/gender-specific /s/ and its relationship to gender identity in children and adolescents'. *Journal of Phonetics* **57**, 59–70.

Li, Feipeng, Trevino, Andrea, Menon, Anjali and Allen, Jont B. (2012). 'A psychoacoustic method for studying the necessary and sufficient perceptual cues of American English fricative consonants in noise'. *Journal of the Acoustic Society of America* **132**(4), 2663–2675.

Li, Mingxing and Zhang, Jie (2017). 'Perceptual distinctiveness between dental and palatal sibilants in different vowel contexts and its implications for phonological contrasts'. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* **8**(1), 1–27.

Li, Wen-Chao (1996). 'Sum: interdental fricatives'. LINGUIST List 7.859.

Li, Wen-Chao (1999). *A diachronically-motivated segmental phonology of Mandarin Chinese*. Vol. 37. Bern / Frankfurt am Main. Peter Lang.

Liem, Quincy C. (2019). 'Does the gay accent exist? An acoustic comparison between homosexual and heterosexual Dutch male millennials'. Bachelor's Thesis. University of Amsterdam.

Liljencrants, Johan and Lindblom, Björn (1972). 'Numerical simulation of vowel quality systems: The role of perceptual contrast'. *Language* **48**(4), 839–862.

Lin, Yen-Hwei (1997). 'Syllabic and moraic structures in Piro'. *Phonology* **14**(3), 403–436.

Lindblad, Per and Lundqvist, Sture (1994). 'The production of Swedish sibilants - An EPG analysis'. *Working Papers of the Department of Linguistics and Phonetics at Lund University* **43**.

Lindblom, Björn (1999). Emergent phonology. *In:* 'Annual Meeting of the Berkeley Linguistics Society'. Vol. 25. pp. 195–209.

Linville, Sue Ellen (1998). 'Acoustic Correlates of Perceived versus Actual Sexual Orientation in Men's Speech'. *Folia Phoniatrica et Logopaedica* **50**(1), 35–48.

Lisker, Leigh (2001). Hearing the Polish sibilants [s š ś]: phonetic and auditory judgements. *In:* Nina Grønnum and Jørgen Rischel (eds.): 'To honour Eli Fischer-Jørgensen'. Vol. 31 of *Travaux du Cercle Linguistique de Copenhague*. Copenhagen. C.A. Reitzel. pp. 226–238.

Liu, Zirui (2018). 'Phonetics of Southern Welsh Stress'. *UCL Working Papers in Linguistics* **30**, 1–34.

Lloyd, Paul M. (1987). *From Latin to Spanish. Vol. I: Historical Phonology and Morphology of the Spanish Language*. Vol. 173 of *Memoirs of the American Philosophical Society*. Philadelphia, PA. American Philosophical Society.

Lombardi, Linda (1990). 'The nonlinear organization of the affricate'. *Natural Language and Linguistic Theory* **8**(3), 375–425.

Lombardi, Linda (1991). Laryngeal features and laryngeal neutralization. PhD thesis. University of Massachusetts, Amherst.

Lombardi, Linda (1995). 'Laryngeal features and privativity'. *The Linguistic Review* **12**(1), 35–59.

Lombardi, Linda (1996). Restrictions on direction of voicing assimilation: an OT account. *In:* 'University of Maryland Working Papers in Linguistics 3'. College. pp. 89–115.

Lope, Juan M. Blanch (2004). *Cuestiones de filologia mexicana / Mexican philology Issues (Spanish Edition)*. Universidad Nacional Autonoma.

Lorenzetti, Luca (2018). On the Palatalization of /s/ + Consonant in Some Dialects of Middle and Southern Italy. *In:* Roberta D'Alessandro and Diego Pescarini (eds.): 'Advances in Italian Dialectology'. Brill. pp. 109–120.

Lucas, Christopher and Čéplö, Slavomír (2020). Maltese. *In:* Christopher Lucas and Stefano Manfredi (eds.): ‘Arabic and contact-induced change’. Berlin. Language Science Press. pp. 265–302.

Lucazin, Mikael (2010). ‘Utkast till ortografi över skånska språket med morfologi och ordlista’.

Macaulay, Monica Ann (1996). *A grammar of Chalcatongo Mixtec*. Berkeley. University of California Press.

Mack, Sara and Munson, Benjamin (2012). ‘The influence of /s/ quality on ratings of men's sexual orientation: Explicit and implicit measures of the ‘gay lisp’ stereotype’. *Journal of Phonetics* **40**(1), 198–212.

Mackie, Scott and Mielke, Jeff (2011). Feature economy in natural, random, and synthetic inventories. *In:* G. Nick Clements and Rachid Ridouane (eds.): ‘Where Do Phonological Features Come From?: Cognitive, physical and developmental bases of distinctive speech categories’. Language Faculty and Beyond 6. Amsterdam. John Benjamins Publishing Company. pp. 43–63.

Maddieson, Ian (1984). *Pattern of Sounds*. Cambridge. Cambridge University Press.

Maddieson, Ian (1991). ‘Commentary on ‘Reading waveforms’’. *Journal of the International Phonetic Association* **21**(2), 89–91.

Maddieson, Ian (1995). Gestural economy. *In:* Kjell Elenius and Peter Branderud (eds.): ‘Proceedings of the 13th International Congress of Phonetic Sciences’. Vol. 4. Stockholm. KTH and Stockholm University Press. pp. 574–577.

Maddieson, Ian (2013*a*). Absence of Common Consonants. *In:* Matthew S. Dryer and Martin Haspelmath (eds.): ‘The World Atlas of Language Structures Online’. Leipzig. Max Planck Institute for Evolutionary Anthropology.

Maddieson, Ian (2013*b*). Voicing in Plosives and Fricatives. *In:* Matthew S. Dryer and Martin Haspelmath (eds.): ‘The World Atlas of Language Structures Online’. Leipzig. Max Planck Institute for Evolutionary Anthropology.

Maddieson, Ian, Flavier, Sébastien, Marsico, Egidio and Pellegrino, François (2014-2021). ‘LAPSyD: Lyon-Albuquerque Phonological Systems Databases, Version 1.1.’.

Malécot, Andre (1974). ‘Frequency of Occurrence of French Phonemes and Consonant Clusters’. *Phonetica* **29**(3), 158–170.

Manolessou, Io and Pantelidis, Nikolaos (2012). ‘Velar fronting in Modern Greek dialects’. *Modern Greek Dialects and Linguistics Theory* **5**(1), 272–286.

Marcato, Gianna and Ursini, Flavia (1998). *Dialetti veneti: grammatica e storia*. Padova. Unipress.

Marotta, Giovanna (2005). Il consonantismo romano. Processi fonologici e aspetti acustici. *In:* F. Albano Leoni and R. Giordano (eds.): 'Italiano parlato. Analisi di un dialogo'. Napoli. Liguori. pp. 1–24.

Marotta, Giovanna (2008). Lenition in Tuscan Italian (Gorgia Toscana). *In:* Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral (eds.): 'Lenition and Fortition'. Studies in Generative Grammar 99. Berlin. Mouton de Gruyter. pp. 235–271.

Martinet, André (1952). 'Celtic Lenition and Western Romance Consonants'. *Language* **28**(2), 192–217.

Martinet, André (1955). *Économie des changements phonétiques*. Berne. Éditions A. Francke.

Martinet, André (1968). Phonetics and linguistic evolution. *In:* Bertil Malmberg (ed.): 'Manual of phonetics'. 2 edn. Amsterdam. L. Kaiser. pp. 252–273.

Martínez-Celdrán, Eugenio, Fernández-Planas, Ana María and Carrera-Sabaté, Josefina (2003). 'Castilian Spanish'. *Journal of the International Phonetic Association* **33**(2), 255–259.

Mascaró, Joan (2004). External allomorphy as emergence of the unmarked. *In:* John J. McCarthy (ed.): 'Optimality Theory in phonology: A reader'. Oxford. Blackwell. pp. 513–522.

Mathieu, Joseph (2001). 'Étude du patois de Saint-Véran'.

Matras, Yaron (2002). *Romani*. Cambridge. Cambridge University Press.

Mazzaro, Natalia (2011). Experimental Approaches to Sound Variation: a Sociophonetic Study of Labial and Velar Fricatives and Approximants in Argentine Spanish. PhD thesis. University of Toronto.

Maïonchi-Pino, Norbert, Taki, Yasuyuki, Magnan, Annie, Yokoyama, Satoru, Écalle, Jean, Takahashi, Kei, Hashizume, Hiroshi and Kawashima, Ryuta (2014). 'Sonority-related markedness drives the misperception of unattested onset clusters in French listeners'. *L'Année psychologique* **115**(02), 197–222.

McCarthy, John J. (1981). 'A Prosodic Theory of Nonconcatenative Morphology'. *Linguistic Inquiry* **12**, 373–418.

McCarthy, John J. (2001). *A Thematic Guide to Optimality Theory*. Cambridge. Cambridge University Press.

McCarthy, John J. (2007). 'Consonant harmony via correspondence: Evidence from Chumash'. *Papers in Optimality Theory III (University of Massachusetts Occasional Papers in Linguistics)* **21**.

McCarthy, John J. and Prince, Alan S. (1994). The emergence of the unmarked: Optimality in prosodic morphology. *In:* Mercè Gonzàlez (ed.): 'Proceedings of the North East Linguistic Society'. Vol. 24. Amherst, MA. GLSA Publications. pp. 333–379. Available on Rutgers Optimality Archive, ROA-13.

McCarthy, John J. and Prince, Alan S. (1995). Faithfulness and reduplicative identity. *In:* Jill N. Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.): 'University of Massachusetts Occasional Papers in Linguistics'. Vol. 18. Amherst, MA. GLSA Publications. pp. 249–384.

McFarland, David H., Baum, Shari R. and Chabot, Caroline (1996). 'Speech compensation to structural modifications of the oral cavity'. *The Journal of the Acoustical Society of America* **100**(2), 1093–1104.

McGowan, Richard S. (1992). 'Tongue-tip trills and vocal-tract wall compliance'. *The Journal of the Acoustical Society of America* **91**(5), 2903–2910.

McManus, Hope (2016). Stress Parallels in Modern OT. PhD thesis. New Brunswick (NJ). Rutgers University. Available on Rutgers Optimality Archive, ROA-1295.

Meneguzzo, Marta (2017). Consonant Clusters and Sonority in the Germanic and Romance Varieties of Northern Italy. PhD thesis. University of Verona.

Mengarda, Elias José (1996). Aquisição do dialeto vêneto no contexto familiar catarinense. Master's thesis. Federal University of Santa Catarina.

Merchant, Nazarré and Krämer, Martin (2018). The Holographic Principle: Typological Analysis Using Lower Dimensions. *In:* Gillian Gallagher, Maria Gouskova and Sora Heng Yin (eds.): 'Proceedings of the Annual Meetings on Phonology'. Washington, DC. Linguistic Society of America.

Merchant, Nazarré and Prince, Alan (2017). The Mother of All Tableaux. Technical report. Eckerd College, The ThOT Lab.

Miccio, Adele W. and Ingrisano, Dennis R. (2000). 'The Acquisition of Fricatives and Affricates: Evidence From a Disordered Phonological System'. *American Journal of Speech-Language Pathology* **9**(3), 214–229.

Miconi, Thomas (2017). 'Biologically plausible learning in recurrent neural networks reproduces neural dynamics observed during cognitive tasks'. *eLife* **6**.

Mielke, Jeff (2007). 'P-base'.

Mielke, Jeff (2012). 'A phonetically based metric of sound similarity'. *Lingua* **122**(2), 145–163.

Mitani, Shigeki, Kitama, Toshihiro and Sato, Yu (2006). 'Voiceless affricate/fricative distinction by frication duration and amplitude rise slope'. *The Journal of the Acoustical Society of America* **120**(3), 1600–1607.

Mohanan, Karuvannur P (1993). Fields of Attraction in Phonology. *In:* John A Goldsmith (ed.): 'The last phonological rule: Reflections on constraints and derivations'. Chicago, London. The University of Chicago Press. pp. 61–116.

Møller, Henrik B. and Pedersen, Christian S. (2004). 'Hearing at low and infrasonic frequencies'. *Noise & health* **6**(23), 37–57.

Montgomery-Anderson, Brad (2008). A Reference Grammar of Oklahoma Cherokee. PhD thesis. University of Kansas.

Moore, Brian C. J. and Glasberg, Brian R. (1983). 'Suggested formulae for calculating auditory-filter bandwidths and excitation patterns'. *The Journal of the Acoustical Society of America* **74**.

Moore, John (2005). The Development of Spanish Sibilants. Course handout.

Morag, Shelomo (2007). Pronunciations of Hebrew. *In:* 'Encyclopaedia Judaica'. 2 edn. Detroit. Thomson Dale.

Moran, Steven and McCloy, Daniel (eds.) (2019). *PHOIBLE 2.0*. Jena. Max Planck Institute for the Science of Human History.

Morén, Bruce (2006). 'Consonant–vowel interactions in Serbian: Features, representations and constraint interactions'. *Lingua* **116**(8), 1198–1244.

Morin, Yves-Charles (2002). 'Les premiers immigrants et la prononciation du français au Québec'. *Revue québécoise de linguistique* **31**(1), 39–78.

Munshi, Sadaf (2006). Jammu and Kashmir Burushaski: Language, Language Contact, and Change. PhD thesis. University of Texas at Austin.

Munson, Benjamin, McDonald, Elizabeth C., DeBoe, Nancy L. and White, Aubrey R. (2006). 'The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech'. *Journal of Phonetics* **34**(2), 202–240.

Muxika-Loitzate, Oihane (2017). 'Sibilant Merger in the Variety of Basque Spoken in Amorebieta-Etxano'. *Languages* **2**(4), 25.

Muxika-Loitzate, Oihane (2019). 'The impact of degree of bilingualism on the Basque sibilant Merger'. Ph.D. project description.

Müller, Eric M. and Brown, William Samuel Jr. (1980). Variations in the supraglottal air pressure waveform and their articulatory interpretation. *In:* Norman Lass (ed.): 'Speech and Language: Advances in Basic Research and Practice'. Vol. 4. Madison. Academic Press Inc.. pp. 318–389.

Najlis, Elena L. (1966). *Lengua abipona*. Buenos Aires. Centro de Estudios Lingüísticos, Univ. de Buenos Aires, Fac. de Filosofía y Letra.

Nartey, Jonas N. Akpanglo (1979). 'A study in phonemic universals - especially concerning fricatives and stops'. *UCLA Working Papers in Phonetics* **46**.

Ngom, Fallou (2006). Loanwords in the Senegalese Speech Community: Their Linguistics Features and Sociolinguistic Significance. *In:* Piotr Nowak and Paweł Nowakowski (eds.): 'Język, Komunikacja, Informacja (Language, Communication, Information)'. Poznań. Sorus Publishers. pp. 103–113.

Niebuhr, Oliver, Lancia, Leonardo and Meunier, Christine (2008). On Place Assimilation in French Sibilant Sequences. *In:* '8th International Seminar on Speech Production'. INRIA. pp. 221–224.

Niebuhr, Oliver and Meunier, Christine (2011). 'The Phonetic Manifestation of French /s#ʃ/ and /ʃ#s/ Sequences in Different Vowel Contexts: On the Occurrence and the Domain of Sibilant Assimilation'. *Phonetica* **68**(3), 133–160.

Nikolaev, Dmitry and Grossman, Eitan (2018). 'Areal sound change and the distributional typology of affricate richness in Eurasia'. *Studies in Language* **42**(3), 562–599.

Nikolaev, Dmitry, Nikulin, Andrey and Kukhto, Anton (2015). 'The database of Eurasian phonological inventories'.

Nirgianaki, Elina (2014). 'Acoustic characteristics of Greek fricatives'. *The Journal of the Acoustical Society of America* **135**(5), 2964–2976.

Nittrouer, Susan (1995). 'Children learn separate aspects of speech production at different rates: Evidence from spectral moments'. *The Journal of the Acoustical Society of America* **97**(1), 520–530.

Nogita, Akitsugu (2016). 'The /s/-/ʃ/ confusion by Japanese ESL learners in grapheme-phoneme correspondence: bias towards [s] and <s>'. *Working Papers of the Linguistics Circle of the University of Victoria* **26**(1), 45–57.

Norton, Rictor (1997). *The Myth of the Modern Homosexual: Queer History and the Search for Cultural Unity*. London. Cassell.

Nowak, Paweł M. (2006). 'The role of vowel transitions and frication noise in the perception of Polish sibilants'. *Journal of Phonetics* **34**(2), 139–152.

Nozaki, Kazunori, Nakamura, Masanori, Takimoto, Haruka and Wada, Shigeo (2011). 'Effect of expiratory flow rate on the acoustic characteristics of sibilant /s/'. *Journal of Computational Science* **3**(5), 298–305.

Nutterville, Catherine (1934). Speech defects as a school problem. Master's thesis. University of Montana.

Nygaard, Lynne C., Sommers, Mitchell S. and Pisoni, David B. (1995). 'Effects of stimulus variability on perception and representation of spoken words in memory'. *Perception & Psychophysics* **57**(7), 989–1001.

Oberly, Stacey (2008). 'An optimality-theoretic analysis of Navajo sibilant harmony'. *Coyote Papers: Working Papers in Linguistics, Linguistic Theory at the University of Arizona* .

Ohala, John J. (1983). The Origin of Sound Patterns in Vocal Tract Constraints. *In:* Peter F. MacNeilage (ed.): 'The Production of Speech'. New York. Springer. pp. 189–216.

Ohala, John J. (2011). 'Accommodation to the aerodynamic voicing constraint and its phonological relevance'. *International Congress of Phonetic Sciences [ICPhS] XVII, Hong Kong* pp. 64–67.

Ohala, John J. and Solé, Maria-Josep (2010). Turbulence and Phonology. *In:* Susanne Fuchs, Martine Toda and Marzena Żygis (eds.): 'Turbulent sounds. An interdisciplinary guide'. Vol. 21. Berlin. Mouton de Gruyter. pp. 37–97.

Omar, Margaret K. and Frith, James R. (1975). *Saudi Arabic: Urban Hijazi Dialect, Basic Course*. Foreign Service Institute, Department of State.

O'Neill, Timothy (2015). The Phonology of Betsimisaraka Malagasy. PhD thesis. University of Delaware.

Ooijevaar, Etske S. M. (2011). Cue weighting in the perception of Dutch sibilants. Master's thesis. University of Amsterdam.

Orel, Vladimir (2000). *A Concise Historical Grammar of the Albanian Language: Reconstruction of Proto-Albanian*. Leiden. Brill.

Ota, Mitsuhiko (2015). L1 phonology: phonological development. *In:* Haruo Kubozono (ed.): 'Handbook of Japanese Phonetics and Phonology'. Vol. 2. Berlin, Boston. De Gruyter Mouton.

Padgett, Jaye (2003). The Emergence of Contrastive Palatalization in Russian. *In:* D. Eric Holt (ed.): 'Optimality Theory and Language Change'. Amsterdam. Springer Netherlands. pp. 307–335.

Padgett, Jaye and Żygis, Marzena (2007). 'The evolution of sibilants in Polish and Russian'. *Journal of Slavic linguistics* **15**(2), 291–324.

Pajusalu, Karl (2012). 'Phonological innovations of the Southern Finnic languages'. *Riho Grünthal and Petri Kallio. A Linguistic Map of Prehistoric Northern Europe* pp. 201–224.

Pan, Hsin-Ying, Utsugi, Akira and Yamazaki, Shuichi (2004). 'An acoustic phonetic study of voiceless alveolo-palatal fricatives in Japanese, Korean and Chinese'. *Journal of General Linguistics* 7, 1–27.

Parise, Cesare V., Knorre, Katharina and Ernst, Marc O. (2014). 'Natural auditory scene statistics shapes human spatial hearing'. *Proceedings of the National Academy of Sciences* **111**(16), 6104–6108.

Parker, Stephen G. (2002). Quantifying the Sonority Hierarchy. PhD thesis. University of Massachusetts Amherst.

Passy, Paul (1890). *Étude sur les changements phonétiques et leurs caractères généraux.* Paris. Firmin-Didot.

Pavlík, Radoslav (2004). 'Slovenské hlásky a medzinárodná fonetická abeceda'. *Jazykovedný časopis* **55**, 87–109.

Pedersen, Holger (1895). 'Das indogermanische s im Slavischen.'. *Indogermanische Forschungen* **5**(1), 33–87.

Peeke, M. Catherine (1973). *Preliminary Grammar of Auca.* Norman, OK. Summer Institute of Linguistics.

Perkell, Joseph S. (1969). *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study.* Cambridge, MA. MIT Press.

Pétursson, Magnus (1971). 'Etude de la réalisation des consonnes islandaises þ, ð, s, dans la prononciation d'un sujet islandais à partir de la radiocinématographie'. *Phonetica* **23**(4), 203–216.

Peust, Carsten (2008). 'On consonant frequency in Egyptian and other languages'. *Lingua Aegyptia* **16**, 105–134.

Pharao, Nicolai, Maegaard, Marie, Spindler Møller, Janus and Kristiansen, Tore (2014). 'Indexical meanings of [s+] among Copenhagen youth: Social perception of a phonetic variant in different prosodic contexts'. *Language in Society* **43**(1), 1–31.

Philippa, Marlies, Debrabandere, Frans, Quak, Arend and van der Sijs, Nicoline (2005). *Etymologisch Woordenboek van het Nederlands, deel 2: F t/m Ka.* Amsterdam. Amsterdam University Press.

Philippart de Foy, Marie (2019). French phonetic and phonological development in simultaneous bilingual acquisition: Individual trajectories and cross-linguistic effects in French-Italian/Arabic/Mandarin bilingual preschool children. PhD thesis. Université de Mons.

Picard, Marc (1987). *An Introduction to the Comparative Phonetics of English and French in North America.* Studies in the Sciences of Language Series. John Benjamins Publishing Company.

Piñeros, Carlos-Eduardo (2016). 'Exceptional nasal-stop inventories'. *Catalan Journal of Linguistics* **15**, 67.

Pintér, Gábor (2015). The emergence of new consonant contrasts. *In:* Haruo Kubozono (ed.): 'Handbook of Japanese Phonetics and Phonology'. Berlin. De Gruyter Mouton. chapter 3, pp. 121–165.

Pintér, Gábor (2014). 'Sibilant representation using MFCCs and GMMs'. *Bulletin of School of Languages and Communication* **11**, 72–84.

Plank, Frans and Filimonova, Elena (2000). 'The Universals Archive: A Brief Introduction for Prospective Users'. *STUF - Language Typology and Universals* **53**(1).

Poirier, Claude (1994). 'Les causes de la variation géolinguistique du français en Amérique du Nord'. *C. Poirier (réd.) Langue, Espace, Société-Les variétés du français en Amérique du Nord, Sainte-Foy, Les Presses de l'Université Laval* pp. 69–98.

Pont, Arnau, Guasch, Oriol, Baiges, Joan, Codina, Ramon and van Hirtum, Annemie (2018). 'Computational aeroacoustics to identify sound sources in the generation of sibilant /s/'. *International Journal for Numerical Methods in Biomedical Engineering* **35**(1), e3153.

Pope, Mildred Katharine (1934). *From Latin to Modern French with Especial Consideration of Anglo-Norman*. Manchester. Manchester University Press.

Poser, William J. (2004). 'On The Status of Chumash Sibilant Harmony'. Ms., University of Pennsylvania, Philadelphia.

Posner, Rebecca (1996). *The Romance Languages*. Cambridge. Cambridge University Press.

Pouplier, Marianne, Hoole, Philip and Scobbie, James M. (2011). 'Investigating the asymmetry of English sibilant assimilation: Acoustic and EPG data'. *Laboratory Phonology* **2**(1), 1–33.

Prince, Alan (1999). 'Paninian Relations'. Handout of a talk given on 17.03.1999 at the University of Marburg.

Prince, Alan (2014). 'Testing the boundaries: aspects of typological structure in OT'. Presentation held on 17.10.2014 at Stanford University.

Prince, Alan and Smolensky, Paul (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford. Blackwell Publishing Ltd.

Prince, Alan, Tesar, Bruce and Merchant, Nazarré (2007-2021). 'OTWorkplace'.

Privitera, Joseph F. (1998). *Basic Sicilian: A Brief Reference Grammar*. Lewiston, NY. Edwin Mellen Press.

Proctor, Michael I., Shadle, Christine H. and Iskarous, Khalil (2010). 'Pharyngeal articulation in the production of voiced and voiceless fricatives'. *Journal of the Acoustical Society of America* **127**(3), 1507–1518.

Pultrová, Lucie (2013). 'On the Phonetic Nature of the Latin R'. *Eruditio Antiqua* **5**, 21–29.

Puppel, Stanisław, Nawrocka-Fisiak, Jadwiga and Krassowska, Halina (1977). *A handbook of Polish pronunciation*. Warsaw. Państwowe Wydawnictwo Naukowe.

Quaresima, Enrico (1957). *Pronuncia trentina e pronuncia normale [sic] della lingua italiana*. Trento. Temi.

Rabanus, Stefan (2019). 'Tedesco, Versione 2 (11.12.2019, 15:14)'. In: *Korpus im Text*, Serie A, 13187.

Rácz, Péter and Shepácz, András (2013). 'The perception of high frequency sibilants in Hungarian male speech'. *Acta Linguistica Hungarica* **60**(4), 457–468.

Ramírez Martínez, Marta (2017). 'Velar Palatalization: Catalan, Spanish and Bilingualism'.

Rataj, Vlastimil (2005). La influencia del Quechua en el Español andino. Master's thesis. Masaryk University.

Rausch, Paul and Baiamonte, Salvatore (2017). 'Proposta di normalizzazione ortografica comune della lingua siciliana per le varietà parlate sull'isola di Sicilia, arcipelaghi ed isole satelliti, e nell'area di Reggio Calabria'.

Recasens, Daniel (2011). 'Velar and dental stop consonant softening in Romance'. *Diachronica* **28**(2), 186–224.

Recasens, Daniel (2012). 'The phonetic implementation of underlying and epenthetic stops in word final clusters in Valencian Catalan'. *Journal of the International Phonetic Association* **42**(1), 65–90.

Recasens, Daniel (2013). 'On the articulatory classification of (alveolo)palatal consonants'. *Journal of the International Phonetic Association* **43**(1), 1–22.

Recasens, Daniel (2014). *Coarticulation and sound change in Romance*. Vol. 329 of *Current Issues in Linguistic Theory*. Amsterdam. John Benjamins Publishing Company.

Recasens, Daniel (2018). *The Production of Consonant Clusters: Implications for Phonology and Sound Change*. Berlin, Boston. De Gruyter Mouton.

Recasens, Daniel and Espinosa, Aina (2007). 'An electropalatographic and acoustic study of affricates and fricatives in two Catalan dialects'. *Journal of the International Phonetic Association* **37**(2), 143–172.

Recasens, Daniel and Espinosa, Aina (2009). 'Acoustics and perception of velar softening for unaspirated stops'. *Journal of Phonetics* **37**(2), 189–211.

Recasens, Daniel and Rodríguez, Clara (2017). 'Lingual Articulation and Coarticulation for Catalan Consonants and Vowels: An Ultrasound Study'. *Phonetica* **74**(3), 125–156.

Remacle, Louis (1992). *La différenciation dialectale en Belgique romane avant 1600*. Vol. 256. Liège. Librairie Droz.

Repp, Bruno H., Liberman, Alvin M., Eccardt, Thomas and Pesetsky, David (1978). 'Perceptual integration of acoustic cues for stop, fricative, and affricate manner.'. *Journal of Experimental Psychology: Human Perception and Performance* **4**(4), 621–637.

Rice, Keren (1996). 'Default variability: The coronal-velar relationship'. *Natural Language and Linguistic Theory* **14**(3), 493–543.

Rice, Keren (2007). Markedness in phonology. *In:* Paul de Lacy (ed.): 'The Cambridge Handbook of Phonology'. Cambridge. Cambridge University Press. pp. 79–98.

Richtsmeier, Peter T. (2010). 'Child phoneme errors are not substitutions'. Available at https://twpl.library.utoronto.ca/index.php/twpl/article/view/6889 (accessed 16.02.2020).

Ridouane, Rachid (2013). *Words without vowels: Phonetic and phonological evidence from Tashlhiyt Berber.* Universitätsbibliothek Johann Christian Senckenberg.

Riggle, Jason (2010). 'Sampling Rankings'. *Rutgers Optimality Archive* **1075**.

Riney, Timothy James, Takagi, Naoyuki, Ota, Kaori and Uchida, Yoko (2007). 'The intermediate degree of VOT in Japanese initial voiceless stops'. *Journal of Phonetics* **35**(3), 439–443.

Riordan, Carol J. (1980). 'Larynx height during English stop consonants'. *Journal of Phonetics* **8**, 353–360.

Roberts, Philip J. (2012). 'Latin Rhotacism: A Case Study in the Life cycle of Phonological Processes'. *Transactions of the Philological Society* **110**(1), 80–93.

Robinson, Stuart (2006). 'The Phoneme Inventory of the Aita Dialect of Rotokas'. *Oceanic Linguistics* **45**(1), 206–209.

Rohlfs, Gerhard (1966). *Grammatica storica dell"italiano e dei suoi dialetti: Fonetica.* Turin. Einaudi.

Romani, Cristina, Galuzzi, Claudia, Guariglia, Cecilia and Goslin, Jeremy (2017). 'Comparing phoneme frequency, age of acquisition, and loss in aphasia: Implications for phonological universals'. *Cognitive Neuropsychology* **34**(7-8), 449–471.

Ronén, Júda (2019). 'A typological study of lateral fricatives'. Manuscript, University of Jerusalem.

Ronjat, Jules (1932). *Grammaire istorique des parlers provençaux modernes.* Vol. 2 of *Grammaire istorique des parlers provençaux modernes.* Montpellier. Société des Langues Romanes.

Rose, Sharon and Walker, Rachel (2004). 'A Typology of Consonant Agreement as Correspondence'. *Language* **80**(3), 475–531.

Rosner, B. S. and Pickering, J. B. (1994). *Vowel Perception and Production*. Oxford. Oxford University Press.

Round, Erich, Hunter, Jessica and Bowern, Claire (2011). 'Reappraising the Effects of Language Contact in the Torres Strait'. *Journal of Language Contact* **4**(1), 106–140.

Rowley, Anthony R. (2017). *Liacht as de sproch: Grammatica della lingua mòchena/Grammatik des Deutsch-Fersentalerischen*. 2 edn. Palai en Bersntol. Bersntoler Kulturinstitut.

Rubach, Jerzy (1993). 'Skeletal versus moraic representations in Slovak'. *Natural Language & Linguistic Theory* **11**(4), 625–653.

Rubach, Jerzy (1994). 'Affricates as strident stops in Polish'. *Linguistic Inquiry* **25**, 119–143.

Rubach, Jerzy (2007). 'Feature Geometry from the Perspective of Polish, Russian, and Ukrainian'. *Linguistic Inquiry* **38**(1), 85–138.

Ruch, Hanna (2018). 'Perception of speaker age and speaker origin in a sound change in progress: The case of /s/-aspiration in Andalusian Spanish'. *Journal of Linguistic Geography* **6**(1), 40–55.

Rupela, Vani, Manjula, R. and Velleman, Shelley L. (2010). 'Phonological processes in Kannada-speaking adolescents with Down syndrome'. *Clinical Linguistics & Phonetics* **24**(6), 431–450.

Rutter, Lawrence Kenji (2001). 'A Reevaluation of Proto-Polynesian *h'. *Oceanic Linguistics* **40**(1), 20–32.

Sagey, Elizabeth Caroline (1986). The Representation of Features and Relations in Non-Linear Phonology. PhD thesis. Massachussets Institute of Technology.

Saib, Jilali (1974). 'Gemination and spirantization in Berber: diachrony and synchrony'. *Studies in African Linguistics* **5**(1), 1.

Salminen, Tapani (2007). 'Notes on Forest Nenets phonology'. *Suomalais-Ugrilaisen Seuran Toimituksia= Mémoires de la Société Finno-Ougrienne* **253**, 349–372.

Samuels, Michael Louis (1975). *Linguistic evolution: with special reference to English*. Vol. 5. Cambridge. Cambridge University Press.

Sapir, Yair (2005). Elfdalian, the Vernacular of Övdaln. *In:* Gunnar Nyström (ed.): 'Rapport från Fuost konferesn um övdalskų: Första konferensen om älvdalska'. Uppsala. Institutionen för nordiska språk.

Schafer, Robin (1995). 'Headedness in the representation of affricates'. *The Linguistic Review* **12**(1).

Scheiner, Andreas (1895). Die Mundart der Siebenbürger Sachsen. *In:* Alfred Kirchhoff (ed.): 'Beiträge zur Siedelungs- und Volkskunde der Siebenbürger Sachsen'. Forschungen zur Deutschen Landes- und Volkskunde. Stuttgart. J. Engelhorn. pp. 125–190.

Schmid, Heinrich (1956). 'Über Randgebiete und Sprachgrenzen'. *Vox Romanica* **15**(2), 19.

Schmid, Stephan (2011). 'An acoustic analysis of palatal obstruents in two Romance varieties'.

Schütz, Albert J. (1994). *The voices of Eden: A history of Hawaiian language studies.* University of Hawaii Press.

Sebregts, Koen (2014). *The Sociophonetics and Phonology of Dutch r.* Utrecht. LOT, Netherlands Graduate School of Linguistics.

Seeber, Elisabeth (2017). Regionalsprachliche Ansätze bei Dialektsprechern im Ahrntal (Südtirol). Master's thesis. Verona. University of Verona.

Ségéral, Philippe and Scheer, Tobias (2008). Positional factors in lenition and fortition. *In:* Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral (eds.): 'Lenition and Fortition'. Studies in Generative Grammar 99. Berlin. Mouton de Gruyter. pp. 131–172.

Seinhorst, Klaas (2012). The evolution of auditory dispersion in symmetric neural nets. Master's thesis. University of Amsterdam.

Seinhorst, Klaas and Ooijevaar, Etske S. M. (2011). 'The auditory dispersion of Dutch sibilants'. Presentation held at the Workshop *Windows on language change and learnability*, University of Amsterdam.

Shadle, Christine H. (1985). The acoustics of fricative consonants. PhD thesis. MIT.

Shadle, Christine H. (1991). 'The effect of geometry on source mechanisms of fricative consonants'. *Journal of Phonetics* **19**(3-4), 409–424.

Shamei, Arian and Gick, Bryan (2019). 'The Effects of Outer Space on Vowel Space'. *Canadian Acoustics - Acoustique canadienne* **47**(3), 110–111.

Shaw, Patricia A. (1991). Consonant harmony systems: The special status of coronal harmony. *In:* Carole Paradis and Jean-François Prunet (eds.): 'The Special Status of Coronals: Internal and External Evidence'. Amsterdam. Elsevier. pp. 125–157.

Shinn, Philip Charles (1986). A Cross-Language Investigation of the Stop, Affricate and Fricative Manners of Articulation. PhD thesis. Brown University.

Shklovsky, Kirill (2005). Person Marking in Petalcingo Tzeltal. Master's thesis. Reed College.

Shoemaker, Ellenor (2016). 'The Development of Perceptual Sensitivity to Polish Sibilants at First Exposure'. *Proceedings of the Annual Meetings on Phonology* **2**.

Shosted, Ryan K. (2006). 'Just Put Your Lips Together and Blow? The Whistled Fricatives of Southern Bantu'. *UC Berkeley PhonLab Annual Report* **2**.

Shosted, Ryan K. (2011). Articulatory and acoustic characteristics of whistled fricatives in Changana. *In:* 'Selected Proceedings of the 40th Annual Conference on African Linguistics: African Languages and Linguistics Today'. Cascadilla Somerville, MA. pp. 119–29.

Sieber, Inna Arnoldovna (2020). Assibilation Typology and Variation in Sibilant Systems. PhD thesis. Lomonosov Moscow State University.

Silverman, Daniel, Blankenship, Barbara, Kirk, Paul and Ladefoged, Peter (2015). 'Phonetic Structures in Jalapa Mazatec'. *Anthropological Linguistics* **37**(1), 70–88.

Šimáčková, Šárka, Podlipský, Václav Jonáš and Chládková, Kateřina (2012). 'Czech spoken in Bohemia and Moravia'. *Journal of the International Phonetic Association* **42**(2), 225–232.

Siptár, Péter and Torkenczy, Miklós (2000). *The phonology of Hungarian.* Oxford. Oxford University Press.

Skjevik Jacobsen, Anna (2015). "Vi ungdomma sir jo sånn /ʃ/ i stedet for /ç/". Master's thesis. UiT (Norges Arktiske Universitet).

Smit, Ann Bosma (1993). 'Phonologic Error Distributions in the Iowa-Nebraska Articulation Norms Project'. *Journal of Speech, Language, and Hearing Research* **36**(3), 533–547.

Smith, Alexander H. (1925). 'Some Place-Names and the Etymology of 'She''. *The Review of English Studies* **1**(4), 437–440.

Smith, Norval (2000). Dependency Theory Meets OT: A Proposal for a New Approach to Segmental Structure. *In:* Joost Dekkers, Frank van der Leeuw and Jeroen van de Weijer (eds.): 'Optimality Theory: Phonology, Syntax, and Acquisition'. Oxford. Oxford University Press. pp. 234–278.

Smolensky, Paul (1996). 'On the comprehension/production dilemma in child language'. *Linguistic Inquiry* **37**, 720–731.

Söderholm, Eira (2017). *Kvensk grammatikk.* Oslo. Cappelen Damm Akademisk/NOASP.

Solé, Maria-Josep (2002). Assimilatory processes and aerodynamic factors. *In:* Carlos Gussenhoven and Natasha Warner (eds.): 'Laboratory Phonology'. Vol. 7. New York. Mouton De Gruyter. pp. 351–386.

Solhaug, Tor Håvard (2010). Retroflexion in Norwegian. Master's thesis. Universitetet i Tromsø (UiT).

Song, Jae Yung, Demuth, Katherine, Evans, Karen and Shattuck-Hufnagel, Stefanie (2013). 'Durational cues to fricative codas in 2-year-olds' American English: Voicing and morphemic factors'. *The Journal of the Acoustical Society of America* **133**(5), 2931–2946.

Spinu, Laura, Kochetov, Alexei and Lilley, Jason (2018). 'Acoustic classification of Russian plain and palatalized sibilant fricatives: Spectral vs. cepstral measures'. *Speech Communication* **100**, 41–45.

Stampe, David (1969). The acquisition of phonetic representation. *In:* R. I. Binnick, Alice Davidson, Georgia M. Green and J. L. Morgan (eds.): 'Papers from the Fifth Regional Meeting of the Chicago Linguistic Society'. Chicago, IL. Dept. of Linguistics, University of Chicago.

Stampe, David (1979). *A Dissertation in Natural Phonology (Ph.D. thesis).* New York. Garland.

Stankiewicz, Edward (1986). *Polish Mazurzenie and the Serbo-Croatian Palatals.* Berlin / New York. De Gruyter Mouton. pp. 105–112.

Stausland Johnsen, Sverre (2012). 'A diachronic account of phonological unnaturalness'. *Phonology* **29**(3), 505–531.

Steeman, Sander (2012). A grammar of Sandawe: A Khoisan language of Tanzania. PhD thesis. Leiden University.

Steever, Sanford B. (ed.) (2019). *The Dravidian languages.* London. Routledge.

Stemberger, Joseph and Bernhardt, Barbara (1997). *Handbook of Phonological Development.* Leiden. Brill.

Steriade, Donca (1994). Complex onsets as single segments: The Mazateco pattern. *In:* Jennifer Cole and Charles Kisseberth (eds.): 'Perspectives in Phonology'. Stanford, CA. CLSI. pp. 203–291.

Stevens, Karen, Keyser, Samuel Jay and Kawasaki, Haruhisa (1986). Toward a Phonetic and Phonological Theory of Redundant Features. *In:* Joseph S. Perkell and Dennis H. Klatt (eds.): 'Invariance and Variability in Speech Processes'. New Jersey. Lawrence Erlbaum Associates. pp. 426–463.

Stevens, Kenneth N. (1971). 'Airflow and Turbulence Noise for Fricative and Stop Consonants: Static Considerations'. *The Journal of the Acoustical Society of America* **50**(4B), 1180–1192.

Stevens, Mary and Harrington, Jonathan (2016). 'The phonetic origins of /s/-retraction: Acoustic and perceptual evidence from Australian English'. *Journal of Phonetics* **58**, 118–134.

Stolz, Thomas and Levkovych, Nataliya (2019). 'The phonological landscape of Europe in areal-typological perspective'. Presentation at ALT 2019, Pavia.

Stuart-Smith, Jane (2007). Empirical evidence for gendered speech production: /s/ in Glaswegian. *In:* Jennifer Cole and José Ignacio Hualde (eds.): 'Laboratory Phonology'. number 9 *In:* 'Phonology and phonetics'. New York. Mouton de Gruyter. pp. 65–86.

Stuart-Smith, Jane (2020). 'Changing perspectives on /s/ and gender over time in Glasgow'. *Linguistics Vanguard* **6**.

Subiabre Ubilla, Paulina Beatriz (2015). A phonetic variationist study on Chilean speakers of English as a foreign language. PhD thesis. University of Glasgow.

Suomi, Kari, Toivanen, Juhani and Ylitalo, Riikka (2008). *Finnish Sound Structure: Phonetics, Phonology, Phonotactics and Prosody.* Vol. 9 of *Studia Humaniora Ouluensia.* Oulu. Oulu University Press.

Šuštaršič, Rastislav, Komar, Smiljana and Petek, Bojan (1995). 'Slovene'. *Journal of the International Phonetic Association* **25**(2), 86–90.

Svantesson, Jan-Olof (2006). Khalkha. *In:* Juha Janhunen (ed.): 'The Mongolic Languages'. London. Routledge. pp. 154–176.

Svirsky, Mario A., Stevens, Kenneth N., Matthies, Melanie L. and Perkell, Joseph S. (1997). 'Tongue surface deformation during obstruent stop consonants'. *Journal of the Acoustical Society of America* **102**, 562–571.

Szamryk, Konrad Kazimierz (2014). Język rękopiśmiennych kazań Krzysztofa Kluka. PhD thesis. Uniwersytet w Białymstoku, Wydział Filologiczny.

Szigetvári, Péter (2008). What and where?. *In:* Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral (eds.): 'Lenition and Fortition'. Studies in Generative Grammar 99. Berlin. Mouton de Gruyter. pp. 93–129.

Tabain, Marija (1998). 'Non-Sibilant Fricatives in English: Spectral Information above 10 kHz'. *Phonetica* **55**(3), 107–130.

Taeldeman, Johan (2006). 'Verscherping van de fricatieven in de Nederlandse dialecten'. *Taal en Tongval* **58**(1), 167–192.

Tawerilmang, Anthony F. and Sohn, Ho-Min (1984). 'Proto-Oceanic reflexes in Woleaian'. *Studies in Micronesian Linguistics* pp. 171–199.

Telfer, Corey S. (2006). Coronalization as Assibilation. Master's thesis. University of Calgary.

Tent, Jan (1993). 'Phonetic symmetry in sound systems'. *Symmetry: Culture and Science* **4**(4), 345–368.

Teodorsson, Sven-Tage (1979). 'On the Pronunciation of Ancient Greek zeta'. *Lingua* **47**(4), 323–332.

Teselkin, Avenir S. and Echols, John M. (1972). *Old Javanese (Kawi)*. Ithaca, NY. Cornell University Press.

Thaut, Michael H. (2014). Musical echoic memory training (MEM). *In:* Michael H. Thaut and Volker Hömberg (eds.): 'Handbook of neurologic music therapy'. Oxford. Oxford University Press. pp. 311–313.

Themistocleous, Charalambos, Savva, Angelandria and Aristodemou, Andrie (2016). Effects of Stress on Fricatives: Evidence from Standard Modern Greek. *In:* 'Interspeech 2016'. ISCA. pp. 1026–1029.

Theraphan, L.-Thongkum (1992). 'A preliminary reconstruction of Proto-Lakkja (Cha Shan Yao)'. *Mon-Khmer Studies* **20**, 57–89.

Tifrit, Ali and Voeltzel, Laurence (2016). 'Revis(it)ing French palatalization'. *Glossa: a journal of general linguistics* **1**(1), 10.

Toda, Martine (2009). Étude articulatoire et acoustique des fricatives sibilantes. PhD thesis. Paris. Paris 3.

Toda, Martine, Maeda, Shinji and Honda, Kiyoshi (2010). Formant-cavity affiliation in sibilant fricatives. *In:* 'Turbulent Sounds'. Berlin, Boston. De Gruyter Mouton. pp. 343–374.

Tracy, Erik C., Bainter, Sierra A. and Satariano, Nicholas P. (2015). 'Judgments of self-identified gay and heterosexual male speakers: Which phonemes are most salient in determining sexual orientation?'. *Journal of Phonetics* **52**, 13–25.

Trask, Robert Lawrence (1996). *Historical linguistics*. Oxford. Oxford University Press.

Trask, Robert Lawrence (2004). *A Dictionary of Phonetics and Phonology*. London. Routledge.

Trubetzkoy, Nikolaï Sergeyevich (1939). *Grundzüge der phonologie*. Göttingen. Van den Hoeck & Ruprecht.

Tsuchida, Ayako (2001). 'Japanese Vowel Devoicing: Cases of Consecutive Devoicing Environments'. *Journal of East Asian Linguistics* **10**(3), 225–245.

Tyroller, Hans (2003). *Grammatische Beschreibung des Zimbrischen von Lusern*. Wiesbaden. Franz Steiner Verlag.

Ueda, Isao (1996). Segmental acquisition and feature specification in Japanese. *In:* Barbara Bernhardt, John Gilbert and David Ingram (eds.): 'Proceedings of the UBS International Conference on Phonological Acquisition'. Somerville, MA. Cascadilla Press. pp. 15–27.

Ulfsbjorninn, Shanti (2017). Bogus clusters and lenition in Tuscan Italian. *In:* ‘Language Faculty and Beyond’. John Benjamins Publishing Company. pp. 278–296.

van Borsel, John (1988). ‘An analysis of the speech of five down's syndrome adolescents’. *Journal of Communication Disorders* **21**(5), 409–421.

Van de Velde, Mark L.O. (2008). Chapter 2: Phonology. *In:* Mark L.O. Van de Velde (ed.): ‘A Grammar of Eton’. Berlin/New York. Mouton de Gruyter. pp. 8–68.

van de Weijer, Jeroen (1996). *Segmental Structure and Complex Segments.* Tübingen. Niemeyer Verlag.

van de Weijer, Jeroen (2017). ‘Emergent Phonological Constraints: The acquisition of *COMPLEX in English’. *Acta Linguistica Academica. An International Journal of Linguistics (Until 2016 Acta Linguistica Hungarica)* **64**(1), 153–165.

van de Weijer, Jeroen and Hinskens, Frans (2004). ‘On segmental complexity: Affricates and patterns of segmental modification in consonant inventories’. *Linguistics in the Netherlands* **21**, 217–228.

van der Hoek, Michel (2010). Palatalization in West Germanic. PhD thesis. University of Minnesota.

van der Sijs, Nicoline (2001). *Chronologisch woordenboek. De ouderdom en herkomst van onze woorden en betekenissen.* Amsterdam. Uitgeverij LJ Veen.

van Leussen, Jan-Willem and Vondenhoff, Maaike (2008). ‘A self-organizational approach to the gradual learning algorithm in modelling the evolution of auditory contrast’. Unpublished student paper for the course ’Evolution of Speech’, University of Amsterdam.

van Oostendorp, Marc (2005*a*). ‘Obligatory Contour Principle’.

van Oostendorp, Marc (2005*b*). ‘The Theory of Faithfulness’. Unpublished manuscript.

van ’t Veer, Burght Marinus (2015). Building a phonological inventory: feature co-occurrence constraints in acquisition. PhD thesis. Utrecht. Leiden University.

van Wyk, E. B., Odendal, F. F. and Nkatini, N. L. (1989). ‘Fonologiese sisteme van Afrikaans en Tsonga: 'n Vergelyking’. *South African Journal of Linguistics* 7(1), 38–45.

Vanvik, Arne (1972). ‘A phonetic-phonemic analysis of Standard Eastern Norwegian’. *Norwegian Journal of Linguistics* **26**, 119–164.

Vaux, Bert and Samuels, Bridget (2005). ‘Laryngeal markedness and aspiration’. *Phonology* **22**(3), 395–436.

Vaux, Bert and Samuels, Bridget (2006). ‘Explaining vowel systems: dispersion theory vs natural selection’. Previous unpublished version of Vaux and Samuels (2015).

Vaux, Bert and Samuels, Bridget (2015). 'Explaining vowel systems: dispersion theory vs natural selection'. *The Linguistic Review* **32**(3), 573–599.

Verhoeven, Jo (2006). 'Klanksysteem en Prosodie in het Belgisch-Limburgs Dialect van Hamont'. *Taal en Tongval* **58**, 230–249.

Vermeer, Willem (2012). 'Why Baba-Yaga? Substratal phonetics and restoration of velars subject to the Progressive Palatalization in Russian/Belorussian and adjacent areas (appr. 600–900 CE)'. *Suomalais-ugrilaisen seuran toimituksia* **266**, 345––370.

Versteegh, Kees, Eid, Mushira, Elgibali, Alaa, Woidich, Manfred and Zaborski, Andrzej (2005). *Encyclopedia of Arabic Language and Linguistics.* Vol. 1. Leiden. Brill.

Versteegh, Kees, Eid, Mushira, Elgibali, Alaa, Woidich, Manfred and Zaborski, Andrzej (2007). *Encyclopedia of Arabic Language and Linguistics.* Vol. 3. Leiden. Brill.

Vietti, Alessandro (2019). Phonological Variation and Change in Italian. *In:* Mark Aronoff (ed.): 'Oxford Research Encyclopedia of Linguistics'. Oxford. Oxford University Press.

Vihman, Marilyn May (1996). *Phonological Development.* Cambridge, MA. Blackwell.

Vijūnas, Aurelijus (2010). 'The Proto-Indo-European Sibilant /s/'. *Historical Linguistics* **123**(1), 40–55.

Vogt, Hans (1963). *Dictionnaire de la langue oubykh.* Oslo. Universitetsforlaget i Oslo.

Voll, Rebecca (2017). A grammar of Mundabli: a Bantoid (Yemne-Kimbi) language of Cameroon. PhD thesis. Leiden University.

Vrieland, Seán D. (2014). 'The pronunciation of hj- in Modern Faroese'. *Arkiv för nordisk filologi* **129**, 233–244.

Vulić, Sanja (2008). 'Sociolingvsitička situacija u gradu Splitu i okolici'. *Miasto w kulturze chorwackiej/Urbano u hrvatskoj kulturi* pp. 329–338.

Walker, Rachel, Byrd, Dani, Lee, Sungbok, DeFreitas, Celeste and Mpiranya, Fidele (2006). 'The articulation of consonants in Kinyarwanda's sibilant harmony'. *The Journal of the Acoustical Society of America* **120**(5), 3376–3376.

Wang, Xinchun and Chen, Jidong (2020). 'The Acquisition of Mandarin Consonants by English Learners: The Relationship between Perception and Production'. *Languages* **5**(2), 20.

Westbury, John R. (1983). 'Enlargement of the supraglottal cavity and its relation to stop consonant voicing'. *Journal of the Acoustical Society of America* **73**, 1322–1336.

Whitehead, Benedicte Nielsen (2012). *The Sound of Indo-European: Phonetics, Phonemics, and Morphophonemics.* Vol. 4. Museum Tusculanum Press.

Whitley, M. Stanley (2003). 'Rhotic representation: problems and proposals'. *Journal of the International Phonetic Association* **33**(1), 81–86.

Widdison, Kirk A. (1997). 'Phonetic explanations for sibilant patterns in Spanish'. *Lingua* **102**(4), 253–264.

Wiese, Richard (1986). 'Zur Theorie der Silbe'. *Studium Linguistik* **20**, 1–15.

Wiese, Richard (1991). 'Was ist extrasilbisch im Deutschen und warum?'. *Zeitschrift für Sprachwissenschaft* **10**(1), 112–133.

Wiese, Richard (1996). *The Phonology of German.* The Phonology of the World's Languages. Oxford. Clarendon Press.

Wikimedia: Rodolfo Hermans (2008). 'File:Negative and positive skew diagrams (English).svg'.

Wikimedia: Rovdyr (2007*a*). 'File:Pl-czysta.ogg'.

Wikimedia: Rovdyr (2007*b*). 'File:Pl-trzysta.ogg'.

Wissing, Daan, Pienaar, Wikus and van Niekerk, Daniel (2015). 'Palatalisation of /s/ in Afrikaans'. *Stellenbosch Papers In Linguistics Plus* **48**(0).

Woidich, Manfred and Zack, Liesbeth (2009). The g/ǧ-question in Egyptian Arabic revisited. *In:* Enam Al-Wer and Rudolf de Jong (eds.): 'Arabic Dialectology'. Vol. 53 of *Studies in Semitic Languages and Linguistics.* Leiden, New York, Köln. Brill. pp. 41–60.

Wolf, Oskar (1871). *Sprache und Ohr: akustisch-physiologische und pathologische Studien.* Braunschweig. Friedrich Vieweg und Söhne.

Woodhouse, Robert (2003). 'The Biblical Shibboleth Story in the Light of Late Egyptian Perceptions of Semitic Sibilants: Reconciling Divergent Views'. *Journal of the American Oriental Society* **123**(2), 271.

Worren, Dagfinn (1986). 'sandwich'. In: *Nynorskordboka (The Nynorsk Dictionary).*

Wottawa, Jane, Adda-Decker, Martine and Isel, Frédéric (2016). Putting German [ʃ] and [ç] in Two Different Boxes: Native German vs L2 German of French Learners. *In:* 'Interspeech 2016'. ISCA.

Wrench, Alan A. (1995). 'Analysis of Fricatives Using Multiple Centres of Gravity'. Ms. University of Edinburgh.

Wright, Joseph (1910). *Grammar of the Gothic Language.* Oxford. Clarendon Press.

Wright, Richard (2004). A review of perceptual cues and cue robustness. *In:* Bruce Hayes, Robert Kirchner and Donca Steriade (eds.): 'Phonetically Based Phonology'. Cambridge. Cambridge University Press. pp. 34–57.

Yanushevskaya, Irena and Bunčić, Daniel (2015). 'Russian'. *Journal of the International Phonetic Association* **45**(2), 221–228.

Yaseen, Maha S. and Mahadin, Radwan S. (2018). 'An Optimality Analysis and Treatment of Phonological Disorders in the Speech of Jordanian Children: A Case Study'. *International Journal of Linguistics* **10**(5), 1–20.

Yip, Moira (1988). 'The Obligatory Contour Principle and Phonological Rules: A Loss of Identity'. *Linguistic Inquiry* **19**.

Yoshida, Shohei (2001). An Element-Based Analysis of Affrication in Japanese. *In:* Jeroen Weijer and Tetsuo Nishihara (eds.): 'Issues in Japanese Phonology and Morphology'. Vol. 51. Walter de Gruyter. p. 193.

Youngberg, Connor (2013). Affrication in Japanese and Quebec French. Handout of a talk given at the Manchester Phonology Meeting (MfM) 2013.

Zago, Mauro, Rabanus, Stefan, Zamparelli, Roberto, Alber, Birgit, Padovan, Andrea, Cordin, Patrizia, Tomaselli, Alessandra, Kruijt, Anne and Chiarello, Fabrizio (2021). 'VinKo (Varietà in Contatto)'. Joint project of the universities of Verona, Trento and Bozen.

Zamboni, Alberto (1988). Italienisch: Areallinguistik. [Kapitel IV a:] Venezien. *In:* Günter Holtus, Michael Metzeltin and Christian Schmitt (eds.): 'Lexikon der romanistischen Linguistik (LRL)'. Vol. 4. Tübingen. Niemeyer Verlag. pp. 517–538.

Zampaulo, Andre (2019). *Palatal Sound Change in the Romance Languages*. Oxford Studies in Diachronic and Historical Linguistics. Oxford. Oxford University Press.

Zanobini, Mirella, Viterbori, Paola and Saraceno, Francesca (2012). 'Phonology and Language Development in Italian Children: An Analysis of Production and Accuracy'. *Journal of Speech, Language, and Hearing Research* **55**(1), 16–31.

Zariquiey, Roberto (2015). *Bosquejo gramatical de la lengua Iskonawa*. Lima/Boston. Latinoamericana Editores/CELACP/Revista de Crítica Literaria Latinoamericana.

Zellou, Georgia (2013). 'Consonant harmony in Moroccan Arabic: Similarity and incomplete neutralization'. *The Journal of the Acoustical Society of America* **133**(5), 3572–3572.

Zharkova, Natalia, Hardcastle, William J. and Gibbon, Fiona E. (2018). 'The dynamics of voiceless sibilant fricative production in children between 7 and 13 years old: An ultrasound and acoustic study'. *The Journal of the Acoustical Society of America* **144**(3), 1454–1466.

Zimmermann, Eva (2009). Metathesis without reordering. Master's thesis. Leipzig University.

Zmarich, C. and Bonifacio, S. (2005). 'Phonetic inventories in Italian children aged 18-27 months: A longitudinal study'. *Proceedings of Interspeech 2005 - Eurospeech* pp. 757–760.

Zsiga, Elizabeth C. (1995). ‘An acoustic and electropalatographic study of lexical and postlexical palatalization in American English’. *Phonology and phonetic evidence: Papers in laboratory phonology IV* **4**, 282–302.

Żygis, Marzena (2003). ‘Phonetic and phonological aspects of Slavic sibilant fricatives’. *ZAS Papers in Linguistics* **32**, 175–213.

Żygis, Marzena (2008). ‘On the avoidance of voiced sibilant affricates’. *ZAS Papers in Linguistics* **49**, 23–45.

Żygis, Marzena, Fuchs, Susanne and Koenig, Laura L (2012). ‘Phonetic explanations for the infrequency of voiced sibilant affricates across languages’. *Laboratory Phonology* **3**(2), 299–336.

Żygis, Marzena and Hamann, Silke (2003). Perceptual and acoustic cues of Polish coronal fricatives. *In:* Maria-Josep Solé, Daniel Recasens and J. Romero (eds.): ‘Proceedings of the 15th International Congress of Phonetic Sciences’. Barcelona. Causal Productions.

Żygis, Marzena, Jaskuła, Marek, Pape, Daniel and Koenig, Laura L (2019). ‘Production of Polish Sibilants in the Process of Language Acquisition’. *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* pp. 1510–1514.

Żygis, Marzena and Padgett, Jaye (2010). ‘A perceptual study of Polish fricatives, and its implications for historical sound change’. *Journal of Phonetics* **38**(2), 207–226.

Żygis, Marzena, Pape, Daniel and Jesus, Luis M. T. (2012). ‘(Non-)retroflex Slavic affricates and their motivation: Evidence from Czech and Polish’. *Journal of the International Phonetic Association* **42**(3), 281–329.

Żygis, Marzena, Pape, Daniel, Jesus, Luis M. T. and Jaskuła, Marek (2015). The effects of intonation on acoustic properties of fricatives. *In:* ‘Proceedings of the 18th International Congress of Phonetic Sciences ICPhS, Glasgow, UK’. Glasgow.

8. Appendix

8.1 Illustrations

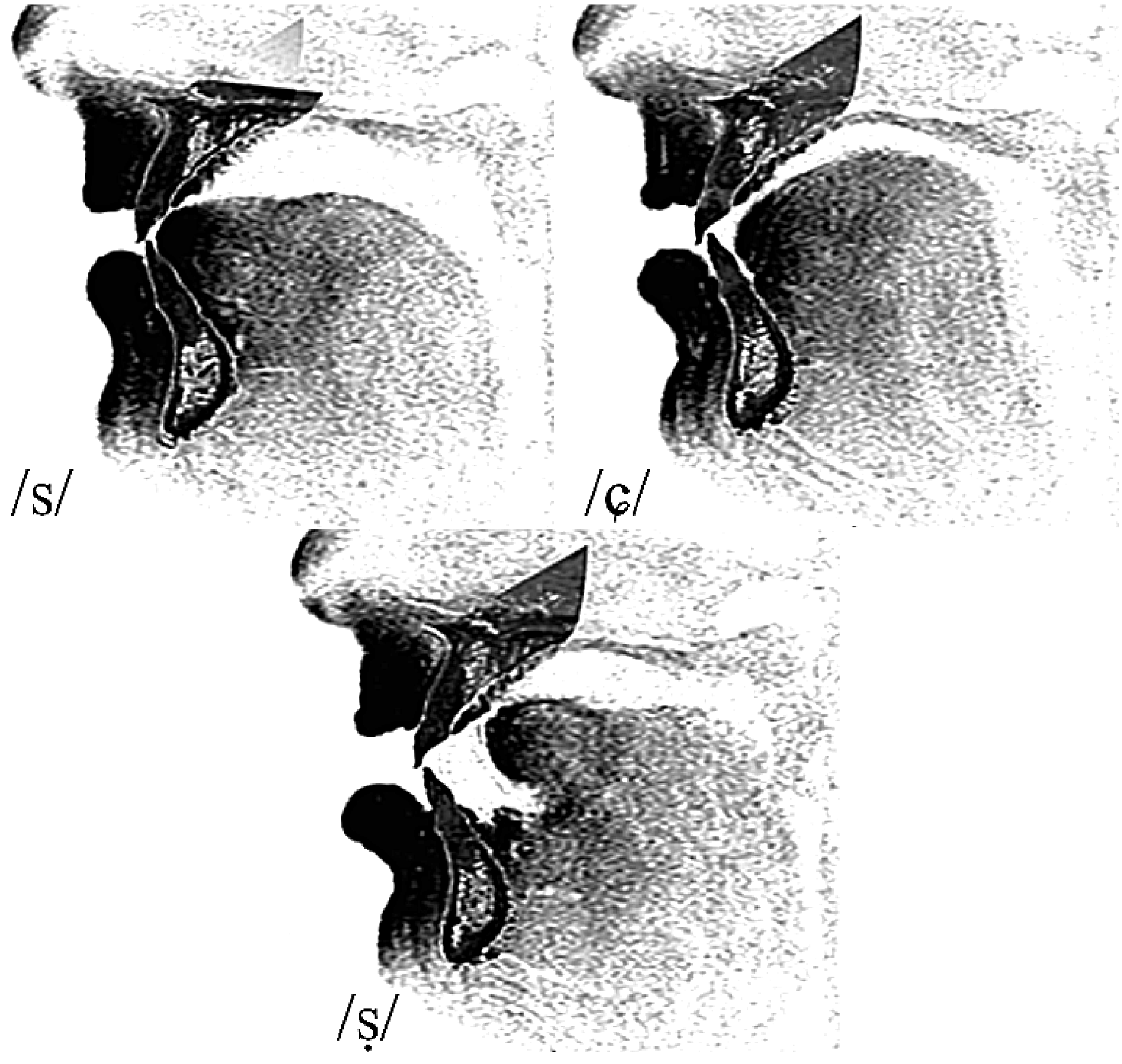


Figure CXXI: *MRI representation of the three sibilant fricatives of Polish, taken and modified from Toda et al. (2010: 360).*

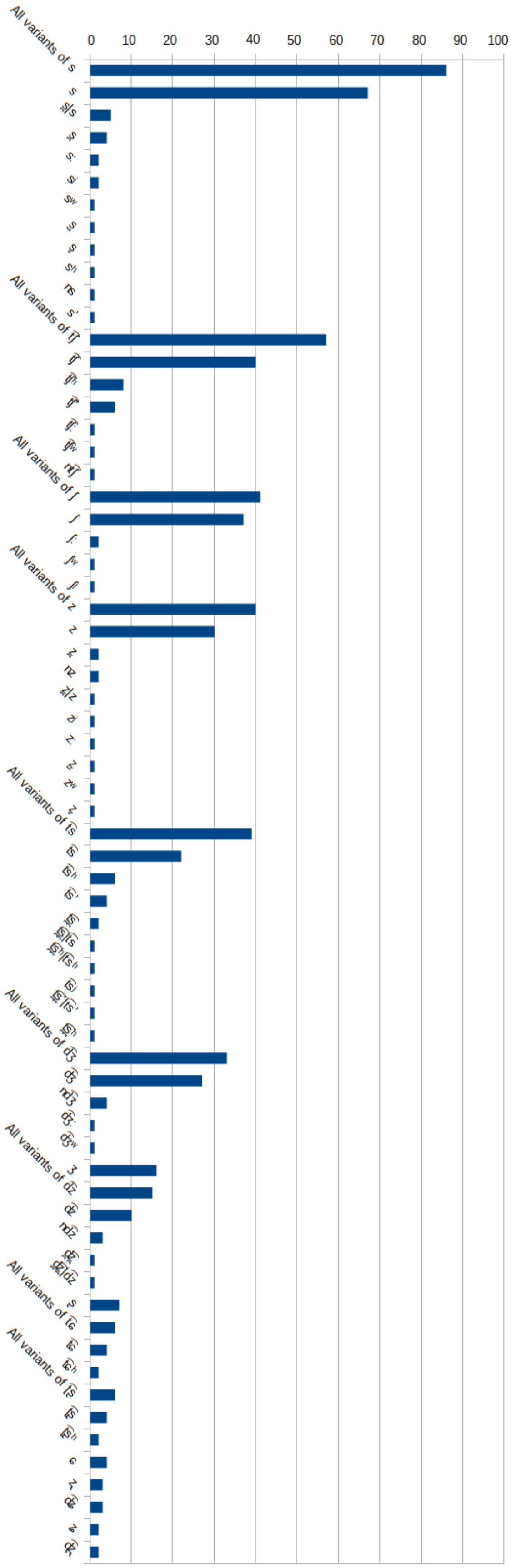


Figure CXXII: *Typological frequency of single sibilant phonemes in the PHOIBLE database (Moran and McCloy 2019). Sibilants with a frequency below 1% have been omitted.*

8.2 Motivations for SibInv.GEN and SibInv.CON

In this section, I motivate the use of GEN and CON as defined in 5.2.1 rather than with other potential definitions. The subsection then focuses on the comparison of the SibInv typologies with alternative typologies that have different constraint formulations (mainly other combinations of: 1) categorical constraints, 2) gradual constraints and 3) stringency hierarchies).

8.2.0.1 Number of PoAs and sibilant markedness

GEN: **A** B C D n as inputs and outputs.
CON: M.SIB: Assign a violation for each *sibilant* in the output (**A**, B, C or D).

As the perception and categorisation experiment in 2.3 has shown (see also 4.2), four is the largest attested number of contrastable places of articulation and in all likelihood the largest learnable for an average human population. This means that both bSibInv and eSibInv allow for 4 PoAs as a maximum, formulated as the abstract categories **A** - B - C - D. As said in (73), the A category includes dental and alveolar sibilants, B includes retracted alveolar and alveopalatal sibilants, C includes palatoalveolars and plain retroflexes, while D includes subapical retroflexes. Reasons behind this categorisation are discussed in 2.2.1. This categorisation means that the attested 1s languages are represented as B, 2s as A - B (for the /s̪ - s̠/ and /s̪ - ɕ/ types), B - C (for the /s̪ - ʃ/ type) or A - C (for the /s - ʃ/ type), 3s as A - B - C (for both /s̪ - ɕ - ṣ/ and /s̪ - s̠ - ʃ̻/) and 4s as A - B - C - D (for both /s̪ - ɕ - ṣ - s̮̠/ and /s̪ - s̠ - ʃ̻ - s̮̠/). Besides this, a non-sibilant segment is represented as n in the typologies. Most of the time, this underspecified non-sibilant corresponds more precisely to a (denti-)alveolar /t, d/, although in languages without /t, d/ like Hawaiian, it corresponds to /k/ (recall the example with /s/ → /k/ in 3.2.2.1). One shall observe that it must be a *less* marked segment than sibilants: /θ/ could not be the non-sibilant segment because if M.SIB dominates F.IDENT(SIB), there will be no sibilant in the language and no segment such as the more marked /θ/ (Jekiel 2012: 24) neither, because the presence of /θ/ implies that of sibilants in any natural language.[1] A language that does not allow sibilants does not allow segments that are even more marked, and will map foreign sibilants to less marked segments (typically stops).

To sum up, the attested inventories reported in 3.2.2 would be categorised as n, B, AB, BC, AC, ABC and ABCD, as an emergent consequence of the average human categorisation ability reaching 4 PoAs at most. As evidenced in 4.2.2.1, the tendentially lower number of contrasted PoAs with respect to the logically predicted average is a consequence of sibilant markedness, which is a teleological principle grounded in articulatory difficulty. It is expressed as the constraint M.SIB, which penalises any sibilant because of its articulatory complexity (cfr. e.g. Adam 2013; Baroni 2014*b*: 34). As Sebregts (2014: 134) mentions (citing McGowan 1992: 218f. and Vihman 1996), sibilants and trills are among the speech sounds mastered last in language acquisition. Further evidence for the

[1] Hall (1997*b*: 91), citing Maddieson (1984: 334), mentions the language Lakkia as having no sibilants and /θ/, but this turns out to be incorrect (Theraphan 1992: 64f.).

constraint m.Sib comes from sibilantless languages, which in fact do not simply lack sibilants (something which could still be an accidental gap in the phoneme distribution) but actively forbid them, as can be seen in loanwords (recall e.g. *mele Kalikimaka* in 3.2.2.1). A typology without a sibilant markedness constraint would thus fail at predicting the existing sibilantless languages.

8.2.0.2 Centralisation

m.Periph(Sib): Assign a violation for each *non-central sibilant* in the output (**A** or **D**).

Centralisation is formulated as the constraint m.Periph(Sib), which penalises sibilants proportionally to how peripheral (more front or more back) they are. I have not found other mentions of centralisation as a constraint in the literature than that of Boersma and Hamann (2008: 243). In subsequent work like Seinhorst (2012: 5) and Boersma et al. (2020), it corresponds to the effect of the articulatory node in Neural Networks. The scarcity of centralisation as an OT constraint might be a consequence of the lack of awareness in much of the literature that retracted sibilants exist as a category of its own (see 3.1.5.1).

The constraint m.Periph(Sib) is intended as penalising *proportionally* to peripherality rather than in a binary way, even if with 4 PoAs, there is no formal difference between a binary or proportional definition. For example, if one distinguishes 5 PoAs rather than 4, adding a distinction between dental and alveolar sibilants, dental [s̪] and subapical [s̺̠] are worse than alveolar [s] and palatoalveolar [ʃ], which are worse than [s̠] (i.e. the hierarchy [s̪, s̺̠] ≫ [s, ʃ] ≫ [s̠]). This is reflected in the use of several articulatory constraints that penalise each a particular sibilant realisation in Boersma and Hamann (2008); their universal(ly learnt) ranking provides for the effect that 'more peripheral = worse' because the articulatory constraints are ranked in such a way that constraints against more peripheral realisations are ranked higher. This universal graduality of violations is expressed differently in the typology I propose, with no different outcome: instead of representing centralisation as a universally fixed ranking, they are formulated as a set of constraints that constitute a stringency hierarchy (Prince 1999; de Lacy 2002, 2004; Merchant and Krämer 2018). In a stringency hierarchy, a constraint penalises e.g. [s̪, s̺̠] + [s, ʃ] + [s̠], another penalises [s̪, s̺̠] + [s, ʃ], and another constraint penalises [s̪, s̺̠] only (note that the first constraint is equal in formulation to the sibilant markedness constraint and therefore redundant). Interacting with faithfulness, this ensures that [s̪, s̺̠] can be allowed to surface only if [s, ʃ] also does, and both can be allowed only if [s̠] is allowed, as illustrated in the lower half of table(au) 29, which illustrates the similarity between Boersma and Hamann (2008)'s centralisation constraints and the stringency hierarchy used here. To illustrate the graduality of centralisation constraints, this case showed what the typology would look like if there were phonemic contrasts between dental and plain alveolar sibilants (5 rather than 4 PoAs), but since they do not contrast and form one category (namely A) instead (Maddieson 1984: 31f.), we will need a constraint less in the typology I propose. In an alternative typology with an upper limit of 5 PoAs, m.Periph(Sib) would be formulated as the constraints *A, B, D, E and *A, E (having removed *A, B, C, D, E because it is identical with m.Sib). However, in the typology with

B. & H. (2008)	*[36.0 Erb]	*[20.0 Erb]	*[24.0 Erb]	*[32.0 Erb]	*[28.0 Erb]
dental s̪	*!				
alveolar s				*!	
retracted alveolar s̠					*
palatoalveolar ʃ			*!		
subapical ʂ		*!			
Stringency hierarchy	NoAE	NoABDE	NoABCDE		
dental s̪	*!	*	*		
alveolar s		*!	*		
retracted alveolar s̠			*		
palatoalveolar ʃ		*!	*		
subapical ʂ	*!	*	*		

Table 29: *Comparison between articulatory constraints with a universally fixed ranking in Boersma and Hamann (2008) and the freely rankable constraints of a stringency hierarchy to represent centralisation. NoAE and NoABDE are equivalent to merging *[36.0 Erb] with *[20.0 Erb] and *[24.0 Erb] with *[32.0 Erb], respectively.*

4 places of articulation, there is only one constraint in the stringency hierarchy (*A, D), which comes down to the same as a binary constraint ('either central or peripheral') in this case. Formally, m.Periph(Sib) would thus be a stringency hierarchy consisting of several constraints if there were more than 4 PoAs, but is logically formulated as only one constraint in the SibInv sibilant typologies.

As has been argued from a typological perspective in 4.3.1, the centre of the articulatory continuum from the most anterior [s̪] to the most posterior [ʂ], which correspond to 0% and 100% on the sibilant PoA continuum, should be located at 50%. In a sibilant typology with 5 (or any odd number of) PoAs, this centre at 50% will perfectly match the centre of the set of sibilants (e.g. C for 5 PoAs, M for 25 PoAs). However, if the number of PoAs is 4 (like for other even numbers), the centre lies in between B and C. m.Periph(Sib) is thus defined as *A, D accordingly. This contradicts the typological observation that languages with only one sibilant have a single B rather than C, and thus that B seems to be the central value of the continuum. Nevertheless, in 8.2.1.1, as a part of 'Alternative typologies that must be discarded', I will show that *A, D rather than other formulations is the correct constraint definition regardless of typological intuitions, and I will show that B surfacing in 1s is in fact not just a consequence of centralisation, but the result of the conjuncture of centralisation and posterior markedness.

8.2.0.3 Posterior markedness

m.PostCD(Sib): Assign a violation for each *posterior sibilant* in the output (C or D).
m.PostD(Sib): Assign a violation for each *subapical posterior sibilant* in the output (D).

Posterior markedness occurs in the OT literature e.g. as *[coron, -anter] in Rubach (2007: 99), *[-anterior] in de Lacy (2002: 426) and Itō and Mester (2003: 675) or *[posterior, fricative] and *[posterior, stop] in Flynn (2012: 106). In all these cases, it is formulated as a single markedness constraint against posterior sibilants (also including alveopalatals). In this sibilant typology, I include posterior markedness as subset inclusion constraints forming a stringency hierarchy: the more posterior a sibilant is, the greater the vertical tongue displacement it requires and therefore the more it is disfavoured. This is reflected in the observation that subapical retroflexes such as [ʂ̻] are only found in languages that have three other PoAs for sibilants; if [ʃ] and [ʂ̻] were on equal footing wrt. posterior markedness, sibilant inventories with less than 4 PoAs and subapical /ʂ̻/ would be predicted not only to exist, but also to be much more frequent than inventories with a plain retroflex /ṣ/, since the former is distinguished from the latter by a better auditory dispersion wrt. the anterior PoAs. However, one observes quite the contrary: the more posterior a sibilant is, the less frequent it is (in natural languages in general, cfr. fig. XLI, as well as inside one and the same language, cfr. 4.4.1.2). Furthermore, a sibilant typology that does not have a stringency hierarchy in the form of m.PostCD(Sib) and m.PostD(Sib) does not predict 3s languages to exist (because either a language does not have posterior sibilants and is a 2s with AB or a 1s with B, or a language does have posterior sibilants and is a 4s with ABCD or a 2s with BC). Considering the numerous well-attested examples of 3s, a typology that fails at predicting them is very unlikely to be right.

8.2.0.4 Voicing markedness

m.Voice(Sib): Assign a violation for each *voiced sibilant* in the output.

Voicing markedness as a constraint against voiced obstruents (or sibilants in particular) is widespread and uncontroversial in the literature. The constraint against voiced sibilants, named m.Voice(Sib) in the SibInv typology, corresponds to e.g. *[+strident, +voice] in Bradley and Delforge (2006: 45), or *Lar in Lombardi (1996) as well as *VoicedObstruents in Baroni (2014*b*: 33) and Alderete (2004: 396) applied to sibilants. It must be formulated as a distinct constraint from *VoicedObstruents, targeting sibilants in particular, because voiced non-sibilant obstruents can exist in a language that forbids voiced sibilants (e.g. Icelandic, Latin or Norwegian). Importantly, it must be formulated as a single constraint m.Voice(Sib) rather than as several constraints that each target a specific voiced sibilant (e.g. *Av, *Bv etc.) for reasons explained in detail in 8.2.1.2, where this alternative typology is discarded.

8.2.0.5 Affricate markedness

m.Affr(Sib): Assign a violation for each *sibilant affricate* in the output.

The constraint m.Affr(Sib) has been used rather extensively in the OT literature on language acquisition, reflecting the later and generally less successful acquisition of (sibilant) affricates (cfr. e.g. Kinney 2004; Dinnsen and Farris-Trimble 2009). Parallelly to the observation in adult language that most frequently, sibilant affricates have corresponding fricatives and that the existence of sibilant affricates presupposes that of sibilant fricatives in general (cfr. 3.2.1.4), the markedness of affricates is expressed as the constraint m.Affr(Sib) in the SibInv OT typology. Since I am not aware of languages that ban sibilant affricates but have non-sibilant affricates,[2] it cannot be established whether the constraint m.Affr(Sib) is distinct from the more general constraint *Affricate.

When used as a single constraint as in this dissertation, m.Affr(Sib) interacts with the faithfulness constraint f.Ident(Sib); if the former is ranked higher than the latter, sibilant affricates will not be allowed in the language and surface as sibilant fricatives. Importantly, many languages do not have sibilant affricates as phonemes and do have sibilant fricatives, but do not have m.Affr(Sib) dominating f.Ident(Sib), although the apparent absence of affricates in the language suggests so. A test to verify whether m.Affr(Sib) dominates f.Ident(Sib) is to check whether foreign words with affricates keep their affricates when loaned into the relevant language. For example, Norwegian does not have sibilant affricates in the native lexicon but in loanwords, they are taken over faithfully rather than being deaffricated or replaced by a stop + fricative combination (e.g. *sandwich* pronounced [sɛnvit͡ʃ]; Worren 1986). As I show in 5.4.2, it must be concluded in such cases that affricates are not forbidden by the OT grammar but rather constitute an accidental gap that is easily filled. Now, considering that most such languages do allow sibilant affricates even when they are rare or absent in native vocabulary (e.g. French, Dutch or Finnish), this raises the question whether the existence of the constraint m.Affr(Sib) is legitimate. In fact, it is, because some phenomena need the constraint m.Affr(Sib) to be explained: besides the markedness of affricates observed in child speech, diachronic deaffrications such as the one which took place from Old French to Middle French to deaffricate all affricates are the result of m.Affr(Sib) dominating f.Ident(Sib). After such an event takes place, affricates have become absent from the native input, but combinations of stops + fricatives are still phonetically present and potentially undistinguished from affricates (if the language allows for codas and/or CC clusters). For example, Middle French had words like *chats* 'cats' (*chat* + *s*), irrespectively of whether still pronounced as [t͡ʃat͡s̪] or already [ʃat͡s̪]. Affricates manage to maintain their presence in such a language despite deaffrication, because children are exposed to what is phonetically [t͡s̪] and start to rank f.Ident(Sib) above m.Affr(Sib) again accordingly.

As in the case of voicing, formulating affricate markedness as several constraints that target specific PoAs individually leads to improbable results. This is illustrated in detail in 8.2.1.3.

8.2.0.6 Fricative - affricate and voicing symmetry

As shown in 4.1.1.2 and 4.7.2.1, the symmetry between sibilant fricatives and affricates is not due to teleological constraints on the well-formedness of the entire inventory, but automatically results from the typology I propose without needing such constraints. Indeed, if one considers the fact that the typology predicts a certain number of possible phonemes (cfr. 5.4.2), there is no need to have symmetry constraints on the inventory because symmetry is a direct consequence of the typology: if the underlying grammar

[2] Wolof could be an example of this by having palatal affricates and no sibilant affricates, but sibilant affricates in English loanwoards are preserved (cfr. Ngom 2006: 110f.), so it is not a clear-cut case.

allows for e.g. /ʂ̻/, it also allows for /t͡ʂ̻/ (if f.Ident(Sib) dominates m.Affr(Sib)) and native speakers will be able to produce the affricate with the same acquired tongue position as that for /ʂ̻/, since the articulation of the affricate predominantly adopts the gesture of the sibilant part (recall the quote from Recasens and Espinosa 2007: 146 in 4.7.1.2). A sibilant typology is therefore successful without any principle active in grammar against asymmetries.

8.2.0.7 Auditory dispersion

As demonstrated in 4.8.2.1 and in detail in 5.1.1.1, auditory dispersion is not part of grammar but emerges in the interaction of speakers and listeners. There is thus no need for auditory constraints in OT. It is included in the model I propose in 5.4, when less-dispersed 2s inventories can become the well-dispersed 2s after several generations through the interaction of OT grammars and Neural Networks.

8.2.0.8 Conservatism and contact

f.Ident(Sib): Assign a violation, if the input and the output are sibilants, for each *sibilant PoA situated in between that of the input sibilant and the output sibilant* on the scale **A** > **B** > **C** > **D**, and if one of input and output is a sibilant and the other a non-sibilant, for each *sibilant PoA in the universal inventory of the typology* (e.g. 1 for **C** → **B**, 2 for **A** → **C**, 4 for **A** → n etc.).

Faithfulness constraints in OT are subdivided into: Max against the deletion of segments from input to output, Ident against the modification of segments or features at corresponding places in the input and output strings, and Dep against the insertion of segments from input to output (McCarthy and Prince 1995). For example, Ident(Strident) prohibits the mapping of an underlying non-sibilant (with [-strident]) to a sibilant and the other way around ([+strident] → [-strident]). Gnanadesikan (2004: 76) explains how Ident constraints operate as follows:

> Ident[F] is a family of constraints which demand identity between a segment's value for some feature [F] and the value of that feature in the segment's correspondent. Thus Ident[F]-IO requires that if an input segment possesses the feature value [αF] the output correspondent of that segment must also possess the value [αF]. (...) Ident[F] can only be violated by a segment which has an output correspondent. If the segment is deleted, Max is violated, but Ident[F] is not.

In the typologies proposed in this dissertation, I have included only one faithfulness constraint (f.Ident(Sib)), with a definition assigning violation marks proportionally to distance in the set of sibilant PoAs. I will address in 8.2.1 the alternative possibilities in which 1) the constraint Max is added to the typology (8.2.1.4); 2) faithfulness is formulated as several Ident constraints, according to the type of feature which is (not) preserved (8.2.1.5); 3) faithfulness is formulated in three different ways: categorically, gradually and as a stringency hierarchy (8.2.1.6 and 8.2.1.7).

Language contact will be modelled in the interface between OT and NN, in 5.4.1. There, it will be expressed outside of OT as the fact that the 'child' gets to perceive acoustic input from two different types of sibilant inventories at the same time.

8.2.1 Alternative typologies that must be discarded

This subsection reports alternative typologies of the sibilant typology, and explains why they are to be discarded.

8.2.1.1 Centralisation: the centre at B instead

Since the 1s inventory has its single sibilant realised as B, one could hypothesise that the most central value of the sibilant continuum is B (similarly to Boersma and Hamann 2008: 229, see fig. LX in this dissertation). This would mean a stringency hierarchy with one centralisation constraint penalising A, C and D, i.e. all sibilants that are not at the centre (B), and a second one penalising D for being even further away from the centre than A and C. This translates to the constraints CentralACD(Sib) (formulated as *A, C, D) and CentralD(Sib) (formulated as *D, i.e. identical to m.PostD(Sib) and therefore superfluous). Such a sibilant typology then yields the same languages as the main typology, except for the BC language, which is absent. As table(au) 30 shows, /C/ cannot surface as [C] under the ranking for the BC language anymore, because centralisation now also penalises C. Considering the absence of the BC language here (while being

	/C/	*D	*A, C, D	[custom]	*C, D	*A, B, C, D
	output	m.PostD(Sib)/ CentralD(Sib)	CentralACD(Sib)	f.Ident(Sib)	m.PostCD(Sib)	m.Sib
	A	0	1!	2	0	1
☞	B	0	0	1	0	1
	C	0	1!	0	1	1
	D	1!	1	1	1	1
	m.Sib	0	0	4!	0	0

Table 30: *Violation Tableau illustrating the impossibility of obtaining a BC language if the centre of the sibilant continuum is considered to be at B rather than in between B and C (i.e. when centralisation also penalises C).*

attested in natural languages like Faroese or Catalan), the main typology presented in this dissertation should be considered more accurate. Besides this typological argument, assuming that the centre of the continuum falls precisely in its centre (i.e. right between B and C) is more logical than assuming it has a slightly forward tongue position. It is thus a positive outcome of the SibInv typology that the centre of the continuum assumed is the most logical one, and simultaneously that the 1s language is correctly predicted as being a language with B.

8.2.1.2 Voicing markedness: individual constraints

In ??, it has been argued against constraints that penalise single sibilants, and in 8.2.0.4, voicing markedness has been formulated accordingly as one constraint against voiced sibilants. In what follows, I argue against voicing markedness as several constraints targeting voiced sibilants individually (i.e. *Av, *Bv, *Cv, *Dv).

A first observation is that such a typology generates unattested inventories. For example, since it allows certain voiced sibilants (i.e. their constraint being dominated by faithfulness) and forbids other voiced sibilants (i.e. their constraint dominating faithfulness), a 4s can have e.g. *Cv against voiced C dominating f.Ident(Sib) which dominates the rest. This yields a language with /s̪, z̪, s̠, z̠, ʃ, _, ʂ, ʐ/, i.e. without /ʒ/, in which a /ʒ/ is mapped to /z̠/ instead (when m.PostCD(Sib) dominates *Bv), as illustrated in fig. CXXIII. This means that a language could consider faithfulness as more important than posterior markedness for a voiceless sibilant but do the opposite for its voiced counterpart (while voicing is possible for all other sibilants), *and* not even map this voiced sibilant to its voiceless counterpart but rather to a sibilant at another PoA. In a typology with only one constraint m.Voice(Sib), such an unattested language would be impossible because m.Voice(Sib) either dominates f.Ident(Sib)or is dominated by it, either allowing voiced sibilants or mapping them all to their voiceless counterpart.

Secondly, a typology with individual voicing markedness constraints (without even counting the fricative - affricate distinctions) predicts 57 different languages, among which the vast majority is as unrealistic as the example in fig. CXXIII (for example, a 3s that maps all voiced sibilants to B).

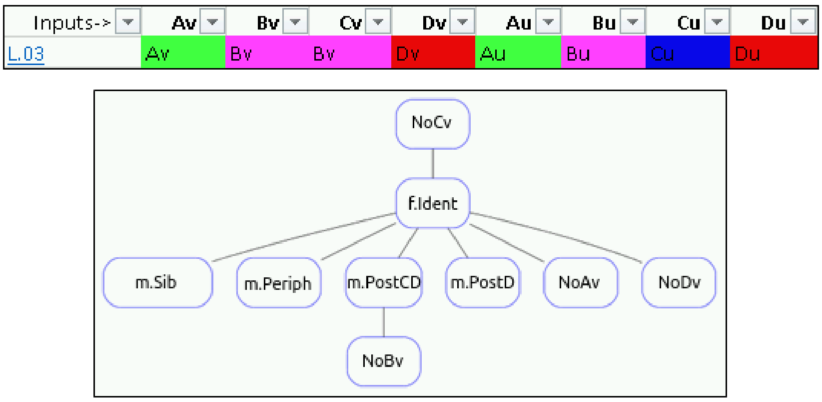


Figure CXXIII: *Grammar of an improbable language in which /ʒ/ is mapped to /z̠/ due to independent voicing markedness constraints (in OTWorkplace; Prince et al. 2007-2021).*

8.2.1.3 Affricate markedness: individual constraints

What was said above about individual voicing markedness constraints applies identically to affricate markedness. With the constraints *tA, *tB, *tC, *tD (*t* standing for an affricate, *s* standing for a fricative), one obtains precisely the same typology with 57 languages as that detailed in the previous subsubsection. This means that there is a 4s inventory in which all sibilants are mapped faithfully except affricates at C, which are mapped to B. In this case as well, a language that maps an underlying /t͡ʃ/ to [t͡s̺] but /ʃ/ to [ʃ] is unrealistic, like many of the other resulting languages.

When allowing both individual voicing markedness constraints and individual affricate markedness constraints, one obtains a total number of 2306 (!) languages. Out of these 2306 languages, there is e.g. one AB language that allows voiced A but forbids voiced B and all affricates, so that [d͡z̠] is mapped to /s̠/ although [d͡z̺] is mapped to /z̺/. Such languages do not reflect the actual behaviour of attested languages: if [d͡z̠] is mapped to /s̠/, [d͡z̺] must be mapped to /s̺/ accordingly, since voicing and affricates are either allowed or forbidden.

8.2.1.4 Faithfulness: a typology with Dep or Max

Faithfulness is mainly subdivided into Max, Dep and Ident constraints. Having constraints that ban epenthesis (i.e. Dep constraints) is of no use in the typology I propose, because the markedness of (certain types of) sibilants is not going to be circumvented by epenthesis. Since nothing is 'solved' with respect to markedness by epenthesising (i.e. by adding segments to the already marked sibilants), Dep is not relevant at all for the sibilant typology.

Relevant faithfulness constraints are of the Ident type, which prevents sibilants from being mapped to other sibilants or non-sibilants, and of the Max type, which prevents sibilants from being deleted to circumvent their markedness. When adding a Max constraint to the basic typology, the 6 original languages surface together with 5 *Doppelgänger* – versions of these languages in which an unfaithfully mapped sibilant is not mapped to the phonetically closest sibilant but is deleted instead (see fig. CXXIV). Pre-

Inputs->	A	B	C	D
L.04	A	B	C	-
L.05	A	B	-	-
L.08	-	B	C	-
L.09	-	B	-	-
L.10	-	-	-	-

Figure CXXIV: *Five additional languages with deletion predicted by a typology with Max (in OTWorkplace; Prince et al. 2007-2021).*

cisely the same additional languages are obtained whether faithfulness is formulated as a single categorical constraint, assigning 0 or 1 violation, or as a single gradual constraint that assigns violations proportionally to the distance on the continuum from A to D. The reason for which there are only 5 rather than 6 additional languages with respect to the basic sibilant typology is that the fully faithful language (ABCD) allows no unfaithful mapping, and thus no deletion. Whether Max crucially dominates or is crucially dominated by certain constraints determines whether a sibilant will be deleted, and if deletion is possible, it affects all sibilants that are not faithfully mapped (for example, A and D in a BC grammar). This implies that if the speaker of e.g. a BC language hears a foreign

[s], (s)he will map it to nihil in production, i.e. delete it (for example, French *clair-obscur* 'chiaroscuro' would be realised as [klɛ.ʀɔp.kyːʀ̥] in Hollandic Dutch and *farce* as [faɹ]). Such a systematic sibilant deletion phenomenon is unattested in natural languages and made unlikely by the fact that there is a strong faithfulness pressure to keep underlying sibilants in the surface form, even when fast speech lowers faithfulness and raises markedness constraints (Baroni 2014*b*: 34). In Baroni (2014*b*)'s "study on Italian casual speech (...) sibilants were practically never reduced or deleted, regardless of their position, whereas all the other consonants and the vowels could potentially disappear". A typology with Max is thus to be avoided. In accordance with the quote by Gnanadesikan (2004: 76) in 8.2.0.8, there are no Ident violations here for deletional candidates; alternatively, if there is one Ident violation also for every deletion, three additional languages wrt. the 11 languages of the typology just described appear: the AB, B and BC languages obtain a new version in which only some unfaithfully mapped sibilants are deleted, others not. This is even less realistic and must also be discarded.

8.2.1.5 Faithfulness: several Ident constraints

input	output	Ident(strident)	Ident(PoA)	Ident(anterior)	Ident(distributed)
A	A	0	0	0	0
	B	0	1	0	1
	C	0	1	1	0
	D	0	1	1	1
	m.Sib	1	0	0	0

Table 31: *Possible Ident constraints applied to A as input in the VT of the sibilant typology.*

Following the definition of Ident constraints above by Gnanadesikan (2004: 76), f.Ident(Sib) could be formulated in the narrow sense as "no non-sibilant becomes a sibilant and no sibilant becomes a non-sibilant", i.e. as a constraint Ident(strident) that only evaluates the preservation of the feature [±strident]. Parallelly, Itō and Mester (2003: 676) use the constraint Ident(anterior) for anterior - posterior contrasts. One can in fact either use Ident(strident) for faithfulness violations between sibilants and non-sibilants and Ident(PoA) for faithfulness violations from one sibilant PoA to another; or, Ident(strident), Ident(anterior) and Ident(distributed), i.e. splitting PoA into [±anterior] and [±distributed], which is closer to the DFT analysis because it uses precisely the features used to characterise the sibilant typology in DFT (see 5.5.2). An example of the violations that such Ident constraints attribute to an A is provided in table(au) 31. A typology with several Ident constraints, as I will show now, leaves excessive space for distinctions.

In the case of Ident(strident) and Ident(PoA), one obtains the same languages as those of the basic typology, with the addition that all languages in which some sibilants are mapped faithfully and some are not (i.e. all but the 0s and the 4s) come in one version in which unfaithful mappings lead to another sibilant (with

Ident(Strident) ≫ Ident(PoA)) and one version in which they lead to a non-sibilant (with Ident(PoA) ≫ Ident(Strident)). The Property Analysis (see fig. CXXV) is then the same as the PA of the basic typology, except that there is an additional property Ident(PoA) ≪≫ Ident(Strident) and that faithfulness needs to be formulated as {Ident(Strident), Ident(PoA)}.sub when interacting against all markedness constraints except m.Sib (where it is m.Sib ≪≫ Ident(Strident)). It thus becomes clear that adding

SibInv	FacTyp				
Inputs->	A	B	C	D	n
L.01	A	B	B1	B1	n
L.02	A	B	C	B1	n
L.03	A	B	C	D	n
L.04	A	B	C	ni	n
L.05	A	B	ni	ni	n
L.06	B1	B	B1	B1	n
L.07	B1	B	C	B1	n
L.08	ni	B	C	ni	n
L.09	ni	B	ni	ni	n
L.10	ni	ni	ni	ni	n

i = violation of Ident(strident)
1 = violation of Ident(PoA)

Props	a	b
FaithCD	NoCD	{Ident(PoA), Ident(strident)}.sub
FaithD	NoD	{Ident(PoA), Ident(strident)}.sub
FaithCentral	Central	{Ident(PoA), Ident(strident)}.sub
SibOrNo	NoSibilant	Ident(strident)
IdentWar	Ident(strident)	Ident(PoA)

PA table	FaithCD	FaithD	FaithCentral	SibOrNo	IdentW	ok?
L.01	a	moot	b	b	a	::
L.02	b	a	b	b	a	::
L.03	b	b	b	b	moot	::
L.04	b	a	b	b	b	::
L.05	a	moot	b	b	b	::
L.06	a	moot	a	b	a	::
L.07	b	moot	a	b	a	::
L.08	b	moot	a	b	b	::
L.09	a	moot	a	b	b	::
L.10	moot	moot	moot	a	moot	::

Figure CXXV: *Factorial Typology and Property Analysis of a typology with Ident(Strident) and Ident(PoA) (in OTWorkplace; Prince et al. 2007-2021). Languages only distinguished by the last property are linked to each other (in red) in the PA Table.*

the strident vs. PoA distinction within the faithfulness constraints does not change much of the analysis: the two constraints bring one additional distinction (namely 'will an unfaithful sibilant be mapped to another sibilant or to a non-sibilant?') and interact for the rest with the markedness constraints exactly as the one-constraint faithfulness does. When it comes to the question which version of faithfulness is more realistic wrt. attested languages, one-constraint faithfulness must be preferred, because natural languages map non-native sibilants to their closest sibilant rather than to a stop (Adams 1975; Vijūnas 2010).

With three Ident constraints (Ident(strident), Ident(anterior) and Ident(distributed)), even more distinctions appear: for example, the B language comes in one version where Ident(strident) ≫ Ident(anterior), Ident(distributed), so that all sibilants are mapped to B (as in the basic sibilant typology); one version in which inversely, Ident(anterior), Ident(distributed) ≫ Ident(strident), so that all sibilants except B are mapped to non-sibilants (since non-sibilants satisfy both Ident(anterior) and Ident(distributed)). Besides these two, there are two other languages that either only allow A to surface as B since they are [+anterior] (Ident(anterior) ≫ Ident(strident) ≫ Ident(distributed)) or only allow D to surface as B since they are [-distributed] (Ident(distributed) ≫ Ident(strident) ≫ Ident(anterior)). However, recall from 3.2.2.2 that only the type in which any sibilant is mapped to B is attested. This typology thus introduces distinctions that are not made by natural languages and is to be avoided.

8.2.1.6 Faithfulness: one categorical f.Ident(Sib) constraint

The most widely used form of Ident constraints is one in which precisely one violation is assigned each time a segment has a different output correspondent, irrespectively of the phonetic distance corresponding to the violation (cfr. e.g. van Oostendorp 2005*b*; de Lacy 2011: 1f., 12). Ident(F) requires the identity of the value of feature *F* between input and output, but is not traditionally used to penalise e.g. the mapping /a/ > [ɕ] more than /a/ > [ɑ]. It is rather usually assumed that since the first mapping violates many Ident constraints (Ident(consonantal), Ident(high), Ident(strident)...) and the second much fewer (Ident(back)), these many Ident constraints will make it much less likely that /a/ ever surfaces as [ɕ]. As just shown in 8.2.1.5, using several Ident constraints with different features in the sibilant typology in such a way introduces unattested distinctions. As will be shown in 8.2.1.7, using a stringency hierarchy of Ident constraints correlating with phonetic distance does not turn out to be a better choice. In what follows, I will show that a typology with the 'classical' categorically formulated single Ident constraint predicts unattested patterns.

If f.Ident(Sib) assigns precisely one violation for every input - output pair that is not identical, the winner of unfaithful mappings is a non-sibilant, as illustrated by the FacTyp in CXXVI. This is because if the faithful option is not available, the best candi-

SibInv	FacTyp				
Inputs->	A	B	C	D	n
AB	A	B	n	n	n
ABC	A	B	C	n	n
ABCD	A	B	C	D	n
B	n	B	n	n	n
BC	n	B	C	n	n
n	n	n	n	n	n

Figure CXXVI: *Factorial Typology of a typology with a single categorically formulated f.Ident(Sib) constraint (in OTWorkplace; Prince et al. 2007-2021).*

date is the least marked one. The non-sibilant is thus selected because it does not have any violation for anything else than f.Ident(Sib), since the other constraints only apply to sibilants. If there were no non-sibilant candidates, the winner of unfaithful mappings would always be B, because it only violates m.Sib besides f.Ident(Sib). That is the only difference between this typology and the basic sibilant typology: in this typology, unfaithful mappings are in favour of the least marked segment, because f.Ident(Sib) does not prefer any among the unfaithful candidates, and the markedness constraints do. Such a typology is unrealistic in the sense that natural languages with e.g. B - C do not map a foreign D to a non-sibilant or to B, but rather to C, since it is the phonetically closest sibilant (as in the basic sibilant typology of 5.2.2.2).

8.2.1.7 Faithfulness: a stringency hierarchy of f.Ident(Sib) constraints

Faithfulness as a stringency hierarchy implies that in order to model the fact that mapping e.g. an A to a D is worse than mapping it to a B, there is one constraint that penalises unfaithful mappings, one that penalises unfaithful mappings with a distance larger than 1, one that penalises unfaithful mappings with a distance larger than 2 etc., as illustrated in fig. CXXVII. This typology yields 18 languages (see fig. CXXVIII), which are either

SibInv	VT.ini		ffunc:ALT1	ffunc:ALT2	ffunc:ALT3	ffunc:ALT4	*A, B, C, D	*A, D	*C, D	*D
input	output	opt	f.ident(Sib)1	f.Ident(Sib)2	f.Ident(Sib)3	f.Ident(Sib)4	m.Sib	m.Periph(Sib)	m.PostCD(Sib)	m.PostD(Sib)
A	A		0	0	0	0	1	1	0	0
	B		1	0	0	0	1	0	0	0
	C		1	1	0	0	1	0	1	0
	D		1	1	1	0	1	1	1	1
	n		1	1	1	1	0	0	0	0
B	A		1	0	0	0	1	1	0	0
	B		0	0	0	0	1	0	0	0
	C		1	0	0	0	1	0	1	0
	D		1	1	0	0	1	1	1	1
	n		1	1	1	1	0	0	0	0

Figure CXXVII: *Extract of a sibilant typology with a stringency hierarchy of* Ident *constraints (in OTWorkplace; Prince et al. 2007-2021).*

variants of one the 6 basic typology languages with differing unfaithful mappings, or a new language: ABD (in which Ident2 $\gg$ m.PostCD(Sib) $\gg$ Ident1 $\gg$ m.PostD(Sib)). Many of these languages are improbable, such as e.g. a language that maps C to B but D to a non-sibilant (because m.PostCD(Sib) $\gg$ Ident2 $\gg$ m.Sib). Such distinctions brought by this alternative typology do not match attested patterns and are *a priori* unlikely to exist.

To summarise, we want to express the fact that some mappings are more unfaithful than others, which generates the need for a typology in which this hierarchy is encoded. However, a stringency hierarchy of faithfulness is not a solution because it leaves too much space for unrealistic possibilities as the ones described here.

Inputs->	A	B	C	D	n		A	B	C	D	n
L.01	A	B	B	B	n	L.10	B	B	B	B	n
L.02	A	B	B	C	n	L.11	B	B	B	C	n
L.03	A	B	B	D	n	L.12	B	B	B	n	n
L.04	A	B	B	n	n	L.13	B	B	C	B	n
L.05	A	B	C	B	n	L.14	B	B	C	C	n
L.06	A	B	C	C	n	L.15	B	B	C	n	n
L.07	A	B	C	D	n	L.16	n	B	C	n	n
L.08	A	B	C	n	n	L.17	n	B	n	n	n
L.09	A	B	n	n	n	L.18	n	n	n	n	n

Figure CXXVIII: *Extract of a sibilant typology with a stringency hierarchy of* Ident *constraints (in OTWorkplace; Prince et al. 2007-2021).*

8.3 Comparison of possible/predicted/attested sibilant inventories

Before one focuses on the *predicted* inventories, let us calculate the number of *logically possible* inventories, i.e. inventories that would be possible if 'anything goes'. A script that generates all possible and predicted inventories for 1, 2, 3 and 4 PoAs can be found in 8.4.1 below in this appendix.

For each number of sibilant PoAs N_{sib} from 1 to 4 (1s, 2s, 3s and 4s), there will be 4 possibilities of voicing and manner (XX, VX, XV and VV); there will be each time 4 total numbers of segments that are possible if there is no gap (in table 24 in 5.4.2.2, it was 1, 2, 2 and 4, respectively, from XX to VV), numbers which we will call N_{seg} in general and e.g. $N_{seg}[2]$ in the case of VV, with the number in between square brackets representing the number of V's (written N_V); and there will be maximal numbers of gaps ranging from 0 to N_{seg} (e.g. in table 24: $N_{seg}[2] = 4$ gaps at most). To give an example, if $N_{sib} = 1$ and the voicing and manner are set to XV, there will be $N_{seg}[1]$ segments ([1] because there is precisely one V in XV), and the maximum number of gaps for that row will be $N_{seg}[1]$. To determine N_{seg}, one simply calculates $N_{seg} = 2^{N_V} \cdot N_{sib}$. With e.g. $N_{sib} = 1$ and $N_V = 1$, it will be $2^1 \cdot 1 = 2$, as can be seen in table 24 (/ʂ t͡ʂ/, i.e. 2 segments). The different variables used here are summarised in (82):

(82) N_{sib} = number of contrasting PoAs

N_V = number of contrasting voicing and manner features, i.e. number of V's (always 0, 1 or 2 if there is only voicing and manner)

N_{gap} = number of gaps

N_{seg} = number of segments, given the number of V's N_V, the number of PoAs N_{sib} and the number of gaps N_{gap} (if unspecified, $N_{gap} = 0$ by default)

$N_{seg}[N_V] = 2^{N_V} \cdot N_{sib}$

$N_{seg}[N_V][N_{gap}] = 2^{N_V} \cdot N_{sib} - N_{gap}$

This means that $N_{seg}[0]$ will be 1, 2, 3 and 4 for N_{sib} = 1, 2, 3 and 4 respectively; $N_{seg}[1]$ will be 2, 4, 6 and 8; and $N_{seg}[2]$ will be 4, 8, 12 and 16. It can be written with a second argument, N_{gap}, which then does not anymore refer to the number of segments given that there are no gaps, but rather given that there are $[gap]$ gaps. For example, $N_{seg}[2][3]$ refers to the number of segments if both voicing and affricates are allowed as well as if there are 3 gaps. The number of segments then becomes $N_{seg}[N_V][N_{gap}] = N_{seg}[N_V] - N_{gap}$, so e.g. $N_{seg}[2][3] = 4 - 3 = 1$ in a 1s, $8 - 3 = 5$ in a 2s, $12 - 3 = 9$ in a 3s and $16 - 3 = 13$ in a 4s.

The number of logically possible inventories for each row, i.e. for each value of N_V, is then written N_{poss}, taking N_{seg} as an argument (written $N_{poss}[N_{seg}]$). It is calculated with the formula in (83):

(83) $$N_{poss} = 2^{N_{seg}}$$

so that for example, if there are 4 contrasting sibilant PoAs ($N_{sib} = 4$) as well as both voicing and affricates ($N_V = 2$), N_{seg} is then $2^2 \cdot 4 = 16$, and N_{poss} will thus be $2^{16} = 65536$, i.e. 65536 logically possible 4s with voicing and affricates. To illustrate the formulas in (82) and (83) in a tangible way, table 32 summarises the calculations in the 1s, which was detailed in 24. With these formulas, one can calculate the number of logically possible

$N_{sib} = 1$	$N_{gap} = 0$	$N_{gap} = 1$	$N_{gap} = 2$	$N_{gap} = 3$	$N_{gap} = 4$	N_{poss}
$N_V = 0$	$N_{seg}[N_V][N_{gap}]$ = $2^0 \cdot 1 - 0$ = 1	$N_{seg}[N_V][N_{gap}]$ = $2^0 \cdot 1 - 1$ = 0	/	/	/	$2^{N_{seg}}$ = 2
$N_V = 1$	$N_{seg}[N_V][N_{gap}]$ = $2^1 \cdot 1 - 0$ = 2	$N_{seg}[N_V][N_{gap}]$ = $2^1 \cdot 1 - 1$ = 1	$N_{seg}[N_V][N_{gap}]$ = $2^1 \cdot 1 - 2$ = 0	/	/	$2^{N_{seg}}$ = 4
$N_V = 1$	$N_{seg}[N_V][N_{gap}]$ = $2^1 \cdot 1 - 0$ = 2	$N_{seg}[N_V][N_{gap}]$ = $2^1 \cdot 1 - 1$ = 1	$N_{seg}[N_V][N_{gap}]$ = $2^1 \cdot 1 - 2$ = 0	/	/	$2^{N_{seg}}$ = 4
$N_V = 2$	$N_{seg}[N_V][N_{gap}]$ = $2^2 \cdot 1 - 0$ = 4	$N_{seg}[N_V][N_{gap}]$ = $2^2 \cdot 1 - 1$ = 3	$N_{seg}[N_V][N_{gap}]$ = $2^2 \cdot 1 - 2$ = 2	$N_{seg}[N_V][N_{gap}]$ = $2^2 \cdot 1 - 3$ = 1	$N_{seg}[N_V][N_{gap}]$ = $2^2 \cdot 1 - 4$ = 0	$2^{N_{seg}}$ = 16

Table 32: *Logically possible 1s inventories as a general mathematical representation.*

inventories both if 1) any PoA pattern is allowed (e.g. C, AD, CD, ABC...) and any gap is allowed, and if 2) only the 7 predicted inventories are allowed (sibilantless, B, AB, AC, BC, ABC, ABCD) but any gap is allowed. In the first case, there are 16 logically possible basic patterns; in the second case, 7 are allowed. For each 0s, there will be one possible inventory; for each 1s (4 in the first case vs. 1 in the second), there will be 2 + 4 + 4 + 16 = 26 possible inventories. For each 2s (6 in the first case vs. 3 in the second), there will be 4 + 16 + 16 + 256 = 292 possibilities; for each 3s (4 in the first case vs. 1 in the second), there will be 8 + 64 + 64 + 4096 = 4232 possibilities, and for each 4s (1 in the first case vs. 1 in the second), there will be 16 + 256 + 256 + 65536 = 66064 possibilities. This means 1 + 104 + 1752 + 16928 + 66064 = 84849 inventories in the first case, if any PoA pattern and any gap are allowed. It means 1 + 26 + 876 + 4232 + 66064 = 71199 in the second case, if only the 7 PoA patterns but any gap are allowed (see footnote 10 in 5.4.2.2).

Turning now to the number of logically possible inventories for each value of N_V and N_{gap}, one observes that the numbers of logically possible inventories in the rows of table 24 are 1 1, 1 2 1 and 1 4 6 4 1, from top to bottom. These correspond to binomial coefficients (i.e. the number of possibilities when one chooses *k* elements out of *n* possible elements): for example, the number of logically possible inventories given 1s, 1 gap and VV will be $k = N_{seg}[2] - N_{gap} = 4 - 1 = 3$, chosen among $n = N_{seg} = 4$ segments. The binomial coefficient being the number of ways in which one can choose, in this case, 3 segments out of 4, its result, written $\binom{n}{k}$, will be $\binom{4}{3}$, which is 4. The formula to calculate binomial coefficients is as follows:

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

One shall also observe that the rows 1 1, 1 2 1 and 1 4 6 4 1 are simply the 2nd, 3rd and 5th of Pascal's triangle, which is a visual representation of binomial coefficients. Given that the first row in Pascal's triangle is the row if $N_{seg} = 0$, one can thus generate the row for any N_{seg} based on N_{sib} and N_V. For example, in a 3s with VV, N_{seg} will be $2^2 \cdot 3 = 12$, and the numbers of logically possible inventories in that row will thus be the 13th row of Pascal's triangle (also called row 12), namely 1 12 66 220 495 792 924 792 495 220 66 12 1. This is illustrated in table 33, which takes up the pattern for the 1s of tables 24 and 32 to represent this time the number of logically possible inventories for each cell. Note that N_{poss} is used in the table as a shorter form to refer to $N_{poss}[N_{seg}[N_V][N_{gap}]] = \binom{N_{seg}}{N_{seg}[N_V][N_{gap}]}$ (e.g. $\binom{4}{2}$ for the number of logically possible inventories if the number of segments without gaps is 4 and the number of segments with gaps is 2). This allows to calculate e.g. the

$N_{sib} = 1$	N_{seg}	**0 gap**	**1 gap**	**2 gaps**	**3 gaps**	**4 gaps**
$N_V = 0$	$N_{seg}[N_V] = 2^0 \cdot 1$ = 1	$N_{poss} = \binom{1}{0}$ = 1	$\binom{1}{1}$ = 1	/	/	/
$N_V = 1$	$N_{seg}[N_V] = 2^1 \cdot 1$ = 2	$N_{poss} = \binom{2}{0}$ = 1	$\binom{2}{1}$ = 2	$\binom{2}{2}$ = 1	/	/
$N_V = 1$	$N_{seg}[N_V] = 2^1 \cdot 1$ = 2	$N_{poss} = \binom{2}{0}$ = 1	$\binom{2}{1}$ = 2	$\binom{2}{2}$ = 1	/	/
$N_V = 2$	$N_{seg}[N_V] = 2^2 \cdot 1$ = 4	$N_{poss} = \binom{4}{0}$ = 1	$\binom{4}{1}$ = 4	$\binom{4}{2}$ = 6	$\binom{4}{3}$ = 4	$\binom{4}{4}$ = 1

Table 33: *Logically possible 1s inventories as a general mathematical representation.*

number of logically possible inventories if $N_{sib} = 4$, $N_V = 2$ and $N_{gap} = 3$, as follows: $N_{seg}[2][3]$ will be $2^2 \cdot 4 - 3 = 13$, and $N_{poss}[N_{seg}[2][3]]$ will be $\binom{2^2 \cdot 4}{13} = \frac{16!}{13! \cdot (16-13)!} = 560$. With the same parameters but $N_{gap} = 8$, one reaches the largest number of logically possible inventories for a single cell in this typological model of sibilant inventories (i.e. given max 4 PoAs, voicing and affricates), which is 12870.

In table 24, the number of *predicted* surface inventories is 1 + 2 + 2 + 6 = 11 out of 26: 1/2 if XX, 2/4 if VX or XV, and 6/16 if VV. To understand how one can calculate this number of 11 out of 26, one will now focus on the number of inventories per cell that conform to the two conditions named above (or inversely, on the number of logically possible inventories that are to be removed as 'implausible', which amounts to the same thing).

Starting with the gap at the voiceless anterior sibilant fricatives, it can obviously not occur if $N_{gap} = 0$. The number of inventories in that column will thus never be reduced, and always equal $N_{seg}[N_V]$. With 1 gap, there will be only 1 case in which /s/ is lacking, also irrespectively of the number of sibilants; the number of conforming inventories is thus $N_{seg}[N_V][1] - 1$ in that column. With 2 gaps and more, the number of inventories without a gap at /s/ is expressible as the number of possibilities when choosing a number corresponding to $N_{seg}[N_V][N_{gap}]-1$ segments out of the sets of sibilants that are not /s/, i.e. out of $N_{seg}-1$. One observes that this formula can be applied to all cells, and that it again implies binomial coefficients (i.e. the number of possibilities when one chooses *k* elements out of *n* possible elements). The formula to calculate the number of inventories that conform to the voiceless anterior sibilant fricative universal is thus:

$$\binom{N_{seg}-1}{N_{seg}[N_V][N_{gap}]-1} = \frac{(N_{seg}-1)!}{(N_{seg}[N_V][N_{gap}]-1)! \cdot ((N_{seg}-1)-(N_{seg}[N_V][N_{gap}]-1))!}$$

To give an example, for the row with VV ($N_V = 2$) in a 1s ($N_{sib} = 1$), as detailed in table 24, the number of inventories that conform to this universal will be: $\binom{3}{3}$ ($N_{gap} = 0$), $\binom{3}{2}$ ($N_{gap} = 1$), $\binom{3}{1}$ ($N_{gap} = 2$), $\binom{3}{0}$ ($N_{gap} = 3$) and $\binom{3}{-1}$ ($N_{gap} = 4$). The results will be 1, 3, 3, 1 and undefined (since one cannot choose -1 object out of a set; in this case, of course, this 'undefined' will mean that if the number of gaps equals the number of segments, the result will be 0 inventories conforming to the universal). Turning now to the voicing gaps, one shall observe that contrarily to the /s/-gap, it does not apply everywhere: while every inventory is susceptible to contain or not contain an anterior voiceless sibilant fricative, some exclude by definition the possibility of voicing (XX and XV). One calculates the number of inventories that conform to the voicing gap condition as follows. Every time a voiced sibilant is present, it must have its voiceless counterpart: if none is present, any voiceless segment can be present or absent in the inventory; if one is present, its counterpart must be present; if two are present, their two counterparts must be present, and so forth. Choosing a voiced sibilant thus means also choosing its voiceless counterpart, if one wants to stay within the bounds of predicted inventories. There are two important implications because of this: if one chooses too few voiced sibilants given the number of gaps (e.g. 2 voiced sibilants out of 16 segments in total, in a 4s VV inventory with 2 gaps), there will not be enough voiceless sibilants to fill the remaining slots (in our example, 2 voiced sibilants imply 2 corresponding voiceless sibilants i.e. 4 taken, thus 6 voiceless sibilants left to fill 12 remaining slots). On the opposite end, selecting too many voiced sibilants given the number of gaps (e.g. 8 voiced sibilants out of 16 segments in total, in a 4s VV inventory with 2 gaps) will make it impossible to fit all necessary voiceless counterparts into the remaining slots (in our example, 8 voiced sibilants imply 8 voiceless sibilants, but there are only 14 available slots). The first condition can be formulated as:[3]

$$N_{vd} \geq N_{seg}[N_V][N_{gap}] - \frac{N_{seg}}{2}$$

where *vd* means 'voiced sibilants *chosen*'. For example, there cannot be 0 voiced segment in a 1s with VV and 1 gap, because $0 < 3 - 2$, but there can be 1 voiced segment in that inventory, because $1 = 3 - 2$. There can be 0 voiced segment in a 1s with VV and 3 gaps, because $0 > 1 - 2$. The second condition can be formulated as follows:

[3] In certain cases (when the number of gaps is larger than half the number of segments), the condition states that e.g. $N_{vd} \geq -1$. In these cases, the actual outcome should be understood as 0, because one cannot select a negative number of voiced sibilants.

$$N_{vd} \leq \frac{N_{seg}[N_V][N_{gap}]}{2}$$

so that e.g. in a 1s with VV and 1 gap, there cannot be 2 voiced segments because $2 > 1.5$. On the opposite, 0 voiced segments would be fine because $0 < 1.5$; this shows that in order to be relevant, the number of voiced segments chosen (N_{vd}) must satisfy *both* conditions. The two conditions can be merged as follows:

$$N_{seg}[N_V][N_{gap}] - \frac{N_{seg}}{2} \leq N_{vd} \leq \frac{N_{seg}[N_V][N_{gap}]}{2}$$

Once a given possible number of voiced sibilants is chosen, there will be a number of possibilities based on how many voiceless sibilants are yet to be chosen. For example, in a 2s with VV and 1 gap, there are $N_{seg}[N_V][N_{gap}] = 7$ empty slots in general; if there are 3 voiced segments (which is the only number satisfying the condition $3 \leq 3 \leq 3.5$), The possibilities of choosing 3 voiced segments out of 4 available will be $\binom{\frac{N_{seg}}{2}}{N_{vd}} = \binom{4}{3} = 4$. However, when one has chosen 3 voiced segments, one must still choose among the segments that 1) are not voiced (because the voiced segments have already been selected) and 2) are not the voiceless counterparts (because they must automatically have been selected). One will thus choose among the remaining voiceless segments. In our example, this means $N_{seg}[N_V][N_{gap}] - 2 \cdot N_{vd} = 7 - 2 \cdot 3 = 1$: if 3 voiced segments have been chosen, their 3 voiceless counterparts have also and there is thus only 1 segment left to choose. If there were 2 left to choose, the number of possibilities in total would be twice as large, because there are 2 options for each selection of 3 voiced segments out of 4. The formula to calculate how many inventories do not have a voicing gap is thus:

$$\binom{\frac{N_{seg}}{2}}{N_{vd}} \cdot \binom{\frac{N_{seg}}{2} - N_{vd}}{N_{seg}[N_V][N_{gap}] - 2 \cdot N_{vd}}$$

which is only of application, of course, if the condition about the number of voiced segments formulated above holds. One can illustrate how both formulas for /s/- and voicing gaps work e.g. with the VV cell of a 2s with 2 gaps (visible in table 25 in 5.4.2.2). The number of logically possible inventories is $\binom{N_{seg}}{N_{seg}[N_V][N_{gap}]} = \binom{8}{6} = 28$; the number of inventories that conform to the /s/-gap condition is $\binom{N_{seg}-1}{N_{seg}[N_V][N_{gap}]-1} = \binom{7}{5} = 21$. The number of inventories that conform to the voicing gap condition is calculated first by determining $N_{seg}[N_V][N_{gap}] - \frac{N_{seg}}{2} \leq N_{vd} \leq \frac{N_{seg}[N_V][N_{gap}]}{2}$ i.e. $6 - 4 \leq N_{vd} \leq 3$, which admits 2 and 3 as values for N_{vd}. If $N_{vd} = 2$, one applies $\binom{\frac{N_{seg}}{2}}{N_{vd}} \cdot \binom{\frac{N_{seg}}{2} - N_{vd}}{N_{seg}[N_V][N_{gap}] - 2 \cdot N_{vd}} = \binom{4}{2} \cdot \binom{4-2}{6-2 \cdot 2} = 6 \cdot 1 = 6$. If $N_{vd} = 3$, it becomes $\binom{4}{3} \cdot \binom{4-3}{6-2 \cdot 3} = 4 \cdot 1 = 4$. One only needs to sum up the results for $N_{vd} = 2$ and $N_{vd} = 3$ to obtain 6 + 4 = 10, which is the number of inventories that conform to the voicing gap condition. A quick look at the outcome of the script in 8.4.1, which generates all logically possible inventories and goes *a posteriori* to determine which ones conform to the two conditions,[4] confirms the numbers of 21 and 10 conforming inventories for the two conditions. However, one observes thanks to the script that the number of inventories that conform to *both* conditions at the same time is certainly not 21 + 10 = 31, but only 9. One can deduce that among the 10 inventories that conform to the second condition, all but one also conform to the first. The non-conforming one has a gap at both /s/ and /z/, which allows it to conform to the second condition but not to the first one. To combine both conditions, one first applies the first one and bears in mind that one voiceless segment has necessarily been used. This means that one applies the 'usual' $\binom{N_{seg}-1}{N_{seg}[N_V][N_{gap}]-1}$ for the /s/-gap, followed by a modified version of the condition for voicing gaps:

$$N_{seg}[N_V][N_{gap}] - 1 - \frac{N_{seg}}{2} \leq N_{vd} \leq \frac{N_{seg}[N_V][N_{gap}] - 1}{2}$$

This is because if /s/ is already selected, the number of available slots and the number of voiceless segments out of which one can choose, respectively, will be lowered by 1 (hence the '-1' on both sides). The formula for voicing gaps must then be modified as follows:

$$\binom{\frac{N_{seg}}{2}-1}{N_{vd}} \cdot \binom{\frac{N_{seg}}{2}-N_{vd}}{N_{seg}[N_V][N_{gap}]-1-2\cdot N_{vd}}$$

which is the same as the 'usual' formula for voicing gaps, but with a '-1' where relevant, to account for the /s/ already selected. To give an example, let us consider a 4s with VX and 4 gaps: it will have 4 free slots for 8 segments, which means $\binom{8-1}{4-1} = 35$ inventories that satisfy the /s/-gap condition. The modified condition for the voicing gap condition will be $4 - 1 - 4 = -1 \leq N_{vd} \leq 1.5$, which allows for 0 and 1 as numbers of voiced segments. The modified formula for the voicing gap condition is then $\binom{\frac{8}{2}-1}{0} \cdot \binom{\frac{8}{2}-0}{4-1-2\cdot 0} = \binom{3}{0} \cdot \binom{4}{3} = 1 \cdot 4 = 4$ for 0 and $\binom{\frac{8}{2}-1}{1} \cdot \binom{\frac{8}{2}-1}{4-1-2\cdot 1} = \binom{3}{1} \cdot \binom{3}{1} = 3 \cdot 3 = 9$ for 1. The sum of both results is 13, which is confirmed as correct by the data generated by the script in 8.4.1. Table 34 below lists the number of predicted sibilant inventories out of the number of logically possible sibilant inventories, per number of sibilants, voicing/affricate distinctions and gaps.

[4] Not by using the formulas detailed here, but by counting how many of the inventories formulated as strings e.g. do start with "1" (meaning /s/), i.e. counting the inventories without knowing what the formulas are. This provides 'independent evidence' to guarantee that the formulas used here are sound.

Nsib	NV	Nseg	0 gap	1 gap	2 gaps	3 gaps	4 gaps	5 gaps	6 gaps	7 gaps	8 gaps	9 gaps	10 gaps	11 gaps	12 gaps	13 gaps	14 gaps	15 gaps	16 gaps
1	XX	1	1/1	0/1	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
	VX	2	1/1	1/2	0/1	/	/	/	/	/	/	/	/	/	/	/	/	/	/
	XV	2	1/1	1/2	0/1	/	/	/	/	/	/	/	/	/	/	/	/	/	/
	VV	4	1/1	2/4	2/6	1/4	0/1	/	/	/	/	/	/	/	/	/	/	/	/
2	XX	2	1/1	1/2	0/1	/	/	/	/	/	/	/	/	/	/	/	/	/	/
	VX	4	1/1	2/4	2/6	1/4	0/1	/	/	/	/	/	/	/	/	/	/	/	/
	XV	4	1/1	3/4	3/6	1/4	0/1	/	/	/	/	/	/	/	/	/	/	/	/
	VV	8	1/1	4/8	9/28	13/56	13/70	9/56	4/28	1/8	0/1	/	/	/	/	/	/	/	/
3	XX	3	1/1	2/3	1/3	0/1	/	/	/	/	/	/	/	/	/	/	/	/	/
	VX	6	1/1	3/6	5/15	5/20	3/15	1/6	0/1	/	/	/	/	/	/	/	/	/	/
	XV	6	1/1	5/6	10/15	10/20	5/15	1/6	0/1	/	/	/	/	/	/	/	/	/	/
	VV	12	1/1	6/12	20/66	45/220	75/495	96/792	96/924	75/792	45/495	20/220	6/66	1/12	0/1	/	/	/	/
4	XX	4	1/1	3/4	3/6	1/4	0/1	/	/	/	/	/	/	/	/	/	/	/	/
	VX	8	1/1	4/8	9/28	13/56	13/70	9/56	4/28	1/8	0/1	/	/	/	/	/	/	/	/
	XV	8	1/1	7/8	21/28	35/56	35/70	21/56	7/28	1/8	0/1	/	/	/	/	/	/	/	/
	VV	16	1/1	8/16	35/120	105/560	238/1820	427/4368	623/8008	750/11440	750/12870	623/11440	427/8008	238/4368	105/1820	35/560	8/120	1/16	0/1

Table 34: *All logically possible and predicted inventories for each number of sibilants, voicing/affricate combinations and number of gaps.*

8.4 Language contact and functional load in NN

This table represents the combinations tested of the AB language with AC, BC and ABC, with contact intensities of 20, 40, 60 and 80% and functional loads of 20-80, 40-60, 60-40, 80-20 for 2s or 20-20-60, 20-60-20, 60-20-20, 40-40-20, 40-20-40, 20-40-40 for the 3s.

	AC		BC		ABC	
	CI	FL	CI	FL	CI	FL
AB	20-80	20-80 vs. 20-80 [AC]	20-80	20-80 vs. 20-80 [BC]	20-80	20-80 vs. 20-20-60 [ABC]
		20-80 vs. 40-60 [AC]		20-80 vs. 40-60 [ABC]		20-80 vs. 20-60-20 [ABC]
		20-80 vs. 60-40 [AC]		20-80 vs. 60-40 [ABC]		20-80 vs. 60-20-20 [ABC]
		20-80 vs. 80-20 [AB]*		20-80 vs. 80-20 [ABC]		20-80 vs. 40-40-20 [ABC]
		40-60 vs. 20-80 [AC]		40-60 vs. 20-80 [ABC]		20-80 vs. 40-20-40 [ABC]
		40-60 vs. 40-60 [AC]		40-60 vs. 40-60 [ABC]		20-80 vs. 20-40-40 [ABC]
		40-60 vs. 60-40 [AC]		40-60 vs. 60-40 [ABC]		40-60 vs. 20-20-60 [ABC]
		40-60 vs. 80-20 [AC]		40-60 vs. 80-20 [BC]		40-60 vs. 20-60-20 [ABC]
		60-40 vs. 20-80 [AC]		60-40 vs. 20-80 [ABC]		40-60 vs. 60-20-20 [ABC]
		60-40 vs. 40-60 [AC]		60-40 vs. 40-60 [ABC]		40-60 vs. 40-40-20 [ABC]
		60-40 vs. 60-40 [ABC]		60-40 vs. 60-40 [ABC]		40-60 vs. 40-20-40 [ABC]
		60-40 vs. 80-20 [AC]*		60-40 vs. 80-20 [ABC]		40-60 vs. 20-40-40 [ABC]
		80-20 vs. 20-80 [AC]		80-20 vs. 20-80 [ABC]		60-40 vs. 20-20-60 [ABC]
		80-20 vs. 40-60 [AC]		80-20 vs. 40-60 [ABC]		60-40 vs. 20-60-20 [ABC]
		80-20 vs. 60-40 [AC]		80-20 vs. 60-40 [ABC]		60-40 vs. 60-20-20 [ABC]
		80-20 vs. 80-20 [AC]		80-20 vs. 80-20 [ABC]		60-40 vs. 40-40-20 [ABC]
	/	/	/	/		60-40 vs. 40-20-40 [ABC]
	/	/	/	/		60-40 vs. 20-40-40 [ABC]
	/	/	/	/		80-20 vs. 20-20-60 [ABC]
	/	/	/	/		80-20 vs. 20-60-20 [ABC]
	/	/	/	/		80-20 vs. 60-20-20 [AC]*
	/	/	/	/		80-20 vs. 40-40-20 [ABC]
	/	/	/	/		80-20 vs. 40-20-40 [ABC]
	/	/	/	/		80-20 vs. 20-40-40 [ABC]
	40-60	20-80 vs. 20-80 [ABC]	40-60	20-80 vs. 20-80 [ABC]	40-60	20-80 vs. 20-20-60 [ABC]
		20-80 vs. 40-60 [ABC]		20-80 vs. 40-60 [ABC]		20-80 vs. 20-60-20 [ABC]
		20-80 vs. 60-40 [ABC]		20-80 vs. 60-40 [ABC]		20-80 vs. 60-20-20 [ABC]
		20-80 vs. 80-20 [ABC]		20-80 vs. 80-20 [ABC]		20-80 vs. 40-40-20 [ABC]
		40-60 vs. 20-80 [ABC]		40-60 vs. 20-80 [ABC]		20-80 vs. 40-20-40 [ABC]
		40-60 vs. 40-60 [AC]		40-60 vs. 40-60 [ABC]		20-80 vs. 20-40-40 [ABC]
		40-60 vs. 60-40 [AC]*		40-60 vs. 60-40 [ABC]		40-60 vs. 20-20-60 [ABC]
		40-60 vs. 80-20 [ABC]		40-60 vs. 80-20 [ABC]		40-60 vs. 20-60-20 [ABC]
		60-40 vs. 20-80 [AC]		60-40 vs. 20-80 [ABC]		40-60 vs. 60-20-20 [ABC]
		60-40 vs. 40-60 [ABC]		60-40 vs. 40-60 [ABC]		40-60 vs. 40-40-20 [ABC]
		60-40 vs. 60-40 [AC]		60-40 vs. 60-40 [ABC]		40-60 vs. 40-20-40 [ABC]
		60-40 vs. 80-20 [ABC]		60-40 vs. 80-20 [ABC]		40-60 vs. 20-40-40 [ABC]
		80-20 vs. 20-80 [AC]		80-20 vs. 20-80 [AC]		60-40 vs. 20-20-60 [ABC]
		80-20 vs. 40-60 [ABC]		80-20 vs. 40-60 [ABC]		60-40 vs. 20-60-20 [ABC]
		80-20 vs. 60-40 [ABC]		80-20 vs. 60-40 [ABC]		60-40 vs. 60-20-20 [ABC]
		80-20 vs. 80-20 [AC]		80-20 vs. 80-20 [ABC]		60-40 vs. 40-40-20 [ABC]
	/	/	/	/		60-40 vs. 40-20-40 [ABC]
	/	/	/	/		60-40 vs. 20-40-40 [ABC]
	/	/	/	/		80-20 vs. 20-20-60 [AC]*
	/	/	/	/		80-20 vs. 20-60-20 [ABC]
	/	/	/	/		80-20 vs. 60-20-20 [ABC]
	/	/	/	/		80-20 vs. 40-40-20 [ABC]
	/	/	/	/		80-20 vs. 40-20-40 [AC]
	/	/	/	/		80-20 vs. 20-40-40 [ABC]
	60-40	20-80 vs. 20-80 [ABC]	60-40	20-80 vs. 20-80 [ABC]	60-40	20-80 vs. 20-20-60 [ABC]
		20-80 vs. 40-60 [ABC]		20-80 vs. 40-60 [ABC]		20-80 vs. 20-60-20 [ABC]
		20-80 vs. 60-40 [ABC]		20-80 vs. 60-40 [ABC]		20-80 vs. 60-20-20 [ABC]
		20-80 vs. 80-20 [ABC]		20-80 vs. 80-20 [ABC]		20-80 vs. 40-40-20 [ABC]
		40-60 vs. 20-80 [ABC]		40-60 vs. 20-80 [ABC]		20-80 vs. 40-20-40 [AC]
		40-60 vs. 40-60 [ABC]		40-60 vs. 40-60 [ABC]		20-80 vs. 20-40-40 [ABC]
		40-60 vs. 60-40 [ABC]		40-60 vs. 60-40 [ABC]		40-60 vs. 20-20-60 [ABC]
		40-60 vs. 80-20 [ABC]		40-60 vs. 80-20 [ABC]		40-60 vs. 20-60-20 [ABC]
		60-40 vs. 20-80 [ABC]		60-40 vs. 20-80 [AC]		40-60 vs. 60-20-20 [ABC]
		60-40 vs. 40-60 [AB]*		60-40 vs. 40-60 [ABC]		40-60 vs. 40-40-20 [ABC]
		60-40 vs. 60-40 [ABC]		60-40 vs. 60-40 [ABC]		40-60 vs. 40-20-40 [AB]*
		60-40 vs. 80-20 [ABC]		60-40 vs. 80-20 [ABC]		40-60 vs. 20-40-40 [ABC]
		80-20 vs. 20-80 [AC]		80-20 vs. 20-80 [AC]		60-40 vs. 20-20-60 [AB]*
		80-20 vs. 40-60 [AC]		80-20 vs. 40-60 [ABC]		60-40 vs. 20-60-20 [ABC]
		80-20 vs. 60-40 [AC]		80-20 vs. 60-40 [ABC]		60-40 vs. 60-20-20 [ABC]
		80-20 vs. 80-20 [ABC]		80-20 vs. 80-20 [ABC]		60-40 vs. 40-40-20 [ABC]
	/	/	/	/		60-40 vs. 40-20-40 [ABC]
	/	/	/	/		60-40 vs. 20-40-40 [ABC]
	/	/	/	/		80-20 vs. 20-20-60 [AC]
	/	/	/	/		80-20 vs. 20-60-20 [ABC]
	/	/	/	/		80-20 vs. 60-20-20 [ABC]
	/	/	/	/		80-20 vs. 40-40-20 [ABC]
	/	/	/	/		80-20 vs. 40-20-40 [AC]*
	/	/	/	/		80-20 vs. 20-40-40 [ABC]
	80-20	20-80 vs. 20-80 [ABC]	80-20	20-80 vs. 20-80 [ABC]	80-20	20-80 vs. 20-20-60 [ABC]
		20-80 vs. 40-60 [ABC]		20-80 vs. 40-60 [ABC]		20-80 vs. 20-60-20 [ABC]
		20-80 vs. 60-40 [ABC]		20-80 vs. 60-40 [ABC]		20-80 vs. 60-20-20 [ABC]
		20-80 vs. 80-20 [ABC]		20-80 vs. 80-20 [ABC]		20-80 vs. 40-40-20 [ABC]
		40-60 vs. 20-80 [ABC]		40-60 vs. 20-80 [ABC]		20-80 vs. 40-20-40 [AC]
		40-60 vs. 40-60 [ABC]		40-60 vs. 40-60 [ABC]		20-80 vs. 20-40-40 [ABC]
		40-60 vs. 60-40 [ABC]		40-60 vs. 60-40 [ABC]		40-60 vs. 20-20-60 [ABC]
		40-60 vs. 80-20 [ABC]		40-60 vs. 80-20 [AB]*		40-60 vs. 20-60-20 [ABC]
		60-40 vs. 20-80 [ABC]		60-40 vs. 20-80 [ABC]		40-60 vs. 60-20-20 [ABC]
		60-40 vs. 40-60 [ABC]		60-40 vs. 40-60 [ABC]		40-60 vs. 40-40-20 [ABC]
		60-40 vs. 60-40 [ABC]		60-40 vs. 60-40 [ABC]		40-60 vs. 40-20-40 [ABC]
		60-40 vs. 80-20 [ABC]		60-40 vs. 80-20 [ABC]		40-60 vs. 20-40-40 [ABC]
		80-20 vs. 20-80 [AC]*		80-20 vs. 20-80 [AC]*		60-40 vs. 20-20-60 [AC]*
		80-20 vs. 40-60 [ABC]		80-20 vs. 40-60 [ABC]		60-40 vs. 20-60-20 [ABC]
		80-20 vs. 60-40 [ABC]		80-20 vs. 60-40 [ABC]		60-40 vs. 60-20-20 [AB]*
		80-20 vs. 80-20 [ABC]		80-20 vs. 80-20 [ABC]		60-40 vs. 40-40-20 [ABC]
	/	/	/	/		60-40 vs. 40-20-40 [ABC]
	/	/	/	/		60-40 vs. 20-40-40 [ABC]
	/	/	/	/		80-20 vs. 20-20-60 [AC]
	/	/	/	/		80-20 vs. 20-60-20 [ABC]
	/	/	/	/		80-20 vs. 60-20-20 [ABC]
	/	/	/	/		80-20 vs. 40-40-20 [ABC]
	/	/	/	/		80-20 vs. 40-20-40 [AC]
	/	/	/	/		80-20 vs. 20-40-40 [ABC]

Table 35: *Possible combinations of contact intensity and functional load for 2s AB. [AB]* is an AB language that is very close to being an AC language, and [AC]* the opposite.*

8.4.1 Praat script generating all possible and predicted inventories

```
# Creating the tables.
nV$[1] = "XX"
nV$[2] = "VX"
nV$[3] = "XV"
nV$[4] = "VV"
nV[1] = 0
nV[2] = 1
nV[3] = 1
nV[4] = 2
for nSib to 4
   for v to 4
      nSeg = 2^nV[v] * nSib
      nPoss = 2^nSeg
      Create Table with column names: "Inventories-"+string$(nSib)+"-"+
         nV$[v], 2, "Inventory Gaps Universal1 Universal2 Predicted"
      Set string value: 1, "Inventory", "1"
      Set string value: 2, "Inventory", "0"
      for i from 2 to nSeg
         numrows = Get number of rows
         for j to numrows
            thisstring$ = Get value: j, "Inventory"
            Set string value: j, "Inventory", thisstring$+"1"
            Append row
            Set string value: numrows+j, "Inventory", thisstring$+"0"
         endfor
      endfor
      for i to nPoss
         thisstring$ = Get value: i, "Inventory"
         segsum = 0
         universal1 = 0
         universal2 = 1
         thisseg$ = ""
         for j to nSeg
            prevseg$ = thisseg$
            thisseg$ = mid$(thisstring$, j, 1)
            if j = 1 && thisseg$ = "1"
               Set numeric value: i, "Universal1", 1
               universal1 = 1
            endif
            if j/2 = round(j/2) && prevseg$ = "0" && thisseg$ = "1" &&
               v/2 = round(v/2)
               universal2 = 0
            endif
            segsum = segsum+number(thisseg$)
         endfor
         if universal2 = 1
            Set numeric value: i, "Universal2", universal2
            if universal1 = 1
               Set numeric value: i, "Predicted", 1
            endif
         endif
```

```
            Set numeric value: i, "Gaps", nSeg-segsum
         endfor
         Sort rows: "Gaps"
      endfor
endfor
# Calculating the stats.
Create Table with column names: "InventoryStats", 0, "Nsib VorX NV Nseg
     Ngap Nposs Universal1 Universal2 Predicted"
for nSib to 4
   for v to 4
      nSeg = 2^nV[v] * nSib
      selectObject: "Table Inventories-"+string$(nSib)+"-"+nV$[v]
      maxgaps = Get maximum: "Gaps"
      for i from 0 to maxgaps
         selectObject: "Table Inventories-"+string$(nSib)+"-"+nV$[v]
         nowarn Extract rows where: "self$[""Gaps""]="""+string$(i)
            +""""
         numposs = Get number of rows
         Remove
         selectObject: "Table Inventories-"+string$(nSib)+"-"+nV$[v]
         nowarn Extract rows where: "self$[""Gaps""]="""+string$(i)+"""
             && self$[""Universal1""]=""1"""
         numvalid1 = Get number of rows
         Remove
         selectObject: "Table Inventories-"+string$(nSib)+"-"+nV$[v]
         nowarn Extract rows where: "self$[""Gaps""]="""+string$(i)+"""
             && self$[""Universal2""]=""1"""
         numvalid2 = Get number of rows
         Remove
         selectObject: "Table Inventories-"+string$(nSib)+"-"+nV$[v]
         nowarn Extract rows where: "self$[""Gaps""]="""+string$(i)+"""
             && self$[""Predicted""]=""1"""
         numvalid3 = Get number of rows
         Remove
         selectObject: "Table InventoryStats"
         Append row
         numrows = Get number of rows
         Set numeric value: numrows, "Nsib", nSib
         Set string value: numrows, "VorX", nV$[v]
         Set numeric value: numrows, "NV", nV[v]
         Set numeric value: numrows, "Nseg", nSeg
         Set numeric value: numrows, "Ngap", i
         Set numeric value: numrows, "Nposs", numposs
         Set numeric value: numrows, "Universal1", numvalid1
         Set numeric value: numrows, "Universal2", numvalid2
         Set numeric value: numrows, "Predicted", numvalid3
      endfor
   endfor
endfor
```

8.5 Scripts for the OT and NN production - perception interaction

8.5.1 Python script running the Praat scripts in alternation

```
#!/usr/bin/env python3

import os

os.system ('/usr/bin/praat --run "AdultOTandNN.praat" "30000"')
# The number specified as argument is the number of tokens per generation.

for x in range(1, 21):
    # range(1, 11) for 10 generations, (1, 37) for 36 generations etc.
    gen = "{}"
    print("The current generation is generation " + gen.format(x))
    os.system ('/usr/bin/praat --run "ChildOTandNN.praat" "' + gen.format(x) + '"')
    # The number specified as argument refers to the generation number.
```

8.5.2 Adult production

```
# Joachim Kokkelmans 03.03.2020. Based on scripts by Paul Boersma and Klaas Seinhorst.
# Requires: A folder InitialObjects at the same level as this script containing the OTGrammar InitialChildGrammar.txt and a folder OTGrammars containing 1s.OTGrammar, 2sAB.OTGrammar etc.; a folder AdultOTandNN containing a folder Distributions (containing 1s.csv, 2sAB.csv etc.) and a folder OTGrammars containing the OTGrammars of the five languages of the 'adult speaker'; a folder ChildOTandNN containing a folder Distributions, a folder DistributionStages, a folder GrammarStages and a folder NetworkStages; the uppermost folder level contains InitialObjects, AdultOTandNN, ChildOTandNN, Run.py, AdultOTandNN.praat and ChildOTandNN.praat.

# Parameters of the simulation in general. The functional loads are set further down.

    form Frequencies
        # Number of times the adult speaker of a grammar pronounces a sibilant to the child.
        natural nbOfSteps 30000
    endform

# Loading all OTGrammar files in the directory (adapted from a script by Kevin Ryan).

    sourcedirectory$ = "InitialObjects/OTGrammars"
    filenameorinitialsubstring$ = ""
    fileextension$ = ".txt"
```

```

# Here below, a list of all sound files in the 'InitialObjects/OTGrammars' folder is created. Based on a script by Kevin Ryan.

   Create Strings as file list... list 'sourcedirectory$'/'filenameorinitialsubstring$'*'fileextension$'
   head_words = selected("Strings")
   file_count = Get number of strings

   for current_file to file_count
      select Strings list
      filename$ = Get string: current_file
      Read from file: sourcedirectory$+"/"+filename$
      grammar$ = filename$ - ".txt"

### Here, the 'real stuff' happens, for the respective OTGrammar has just been loaded. The non-sibilant is not a possible output because we are not interested in making the 'child' perceive that, since it is unmappable on the sibilant continuum.

      # Creating a table to record every input - output ever produced by the adult.
      Create Table with column names: "pronouncedUFSFtokens", 0, "UF SF PV NbMismatches"

      # Getting to know which are the categories present in the native language of the adult. This is unproblematic because the adult already has perfect lexical knowledge; the child won't have this lexical knowledge before or during learning.
      selectObject: "OTGrammar "+grammar$
      winnerstring$ = ""
      numtableaux = Get number of tableaus
      for i to numtableaux
         winner = Get winner: i
         alreadywon = index(winnerstring$, "_"+string$(winner))
         if alreadywon = 0
            winnerstring$ = winnerstring$+"_"+string$(winner)
         endif
      endfor
      # winnerstring$ now contains: _2 (i.e. B), _1_2 (i.e. AB), _2_3 (i.e. BC), _1_2_3 (i.e. ABC) and _1_2_3_4 (i.e. ABCD), depending on the language.
      nbOfPhonemes = length(winnerstring$)/2
      for i to nbOfPhonemes
         originallength = length(winnerstring$)/2
         placeoflastbar = rindex(winnerstring$, "_")
         phoneme'i' = number(right$(winnerstring$, (originallength*2)-placeoflastbar))
         winnerstring$ = left$(winnerstring$, placeoflastbar-1)
      endfor
      # winnerstring$ has been split into parts (phoneme'1', phoneme'2' etc., depending on the number of phonemes), from right to left.
```

```

    # Setting the functional loads.
    if nbOfPhonemes = 1
       funcload[1] = 1
    elsif nbOfPhonemes = 2
       funcload[1] = 1/2
       funcload[2] = 1/2
    elsif nbOfPhonemes = 3
       funcload[1] = 1/3
       funcload[2] = 1/3
       funcload[3] = 1/3
    elsif nbOfPhonemes = 4
       funcload[1] = 1/4
       funcload[2] = 1/4
       funcload[3] = 1/4
       funcload[4] = 1/4
    endif

    ### Checking the preset ranking values of all constraints in the OTGrammar object to see whether some have 0 and must thus get random ranking values between min - 10 and max + 10.
    # First, getting the minimum and the maximum of the constraint ranking values.
    selectObject: "OTGrammar "+grammar$
    numconstr = Get number of constraints
    minrv = 100
    maxrv = 100
    for j to numconstr
       thisrv = Get ranking value: j
       if thisrv < minrv && thisrv <> 0
          minrv = thisrv
       elsif thisrv > maxrv ;NB: && thisrv <> 0 is logically entailed.
          maxrv = thisrv
       endif
    endfor
    # Then, remembering for later which constraints should have random values at every new step by putting them in the noncrucialconstr array.
    nbofnoncrucialconstr = 0
    for j to numconstr
       thisrv = Get ranking value: j
       if thisrv = 0
          nbofnoncrucialconstr = nbofnoncrucialconstr + 1
          noncrucialconstr'nbofnoncrucialconstr' = j
       endif
    endfor

    ### Now, our 'perfect' adult speaker is about to speak.
    for i to nbOfSteps
       selectObject: "OTGrammar "+grammar$
       # Setting a random ranking value for uncrucially ranked constraints
```

```
        for j to nbofnoncrucialconstr
            thisrv = randomUniform(minrv-10, maxrv+10)
            Set ranking: noncrucialconstr'j', thisrv, thisrv
        endfor
        # I hesitate between:
        # Selecting randomly an A, B, C or D input with equal (1/4th)
            probabilities, irrespectively of what the language is.
            This is full Richness of the Base, but implies that the
            lexicon of the adult speaker can contain non-phonemes.
        # Or:
        # Selecting randomly an A, B, C or D input with equal
            probabilities but only if it exists as a phoneme (e.g. if
            the language is BC, only B or C can be an input).
        # This sounds as going against ROTB, and implies that the
            adult draws tokens only from the native lexicon, and does
            thus not attempt e.g. to pronounce underlying foreign non-
            native sibilants.
        # What follows is a switch between the ROTB-version and a
            version in which only existing phonemes are taken from the
             lexicon. Default is 1 (phoneme-only version).
        # A problem of the ROTB version is that in the ABC language, A
             will frequently be mapped to A, B to B, C to C, but D
            will also be frequently mapped to C, which does that C
            ends up being +/- twice as frequent as A or B.
        switch = 1
        if switch = 1
            repeat
                rand = randomInteger(1, nbOfPhonemes)
                productionphoneme = phoneme'rand'
                randzeroone = randomUniform(0, 1)
            until funcload[rand] > randzeroone
        else
            repeat
                rand = randomInteger(1, 4)
                productionphoneme = rand
                randzeroone = randomUniform(0, 1)
            until funcload[rand] > randzeroone
        endif
        if productionphoneme = 1
            randomUFtoken$ = "A"
        elsif productionphoneme = 2
            randomUFtoken$ = "B"
        elsif productionphoneme = 3
            randomUFtoken$ = "C"
        elsif productionphoneme = 4
            randomUFtoken$ = "D"
        endif
        # Getting once a winning output for the input with a
            stochastic noise in the constraint ranking.
        pronouncedSFtoken$ = Input to output: randomUFtoken$, 4
        # Updating the table with the input - output pairs pronounced.
        selectObject: "Table pronouncedUFSFtokens"
        Append row
```

```
      numrows = Get number of rows
      Set string value: numrows, "UF", randomUFtoken$
      Set string value: numrows, "SF", pronouncedSFtoken$
      ### Converting the resulting SF into a value on an abstract continuum from 0 (dental sibilant) to 100 (subapical sibilant). A = 0-25 with prototype 12.5, B = 25-50 with prototype 37.5 etc., with a SD of 'standev'.
      # The noise is inversely proportional to the number of contrasting sibilants, since the permissible variation in the realisation of a sibilant is larger in smaller inventories (cfr. Boersma and Hamann 2008: 222; Sieber 2020).
      standev = 10/sqrt(nbOfPhonemes)
      if pronouncedSFtoken$ = "A"
        repeat
          pronouncedvalue = randomGauss(12.5, standev)
        until pronouncedvalue >= 0 && pronouncedvalue <= 100
      elsif pronouncedSFtoken$ = "B"
        repeat
          pronouncedvalue = randomGauss(37.5, standev)
        until pronouncedvalue >= 0 && pronouncedvalue <= 100
      elsif pronouncedSFtoken$ = "C"
        repeat
          pronouncedvalue = randomGauss(62.5, standev)
        until pronouncedvalue >= 0 && pronouncedvalue <= 100
      elsif pronouncedSFtoken$ = "D"
        repeat
          pronouncedvalue = randomGauss(87.5, standev)
        until pronouncedvalue >= 0 && pronouncedvalue <= 100
      endif
      Set numeric value: numrows, "PV", pronouncedvalue
    endfor

    ### Saving all the tokens produced in the 'AdultOTandNN/Distributions' folder.
    selectObject: "Table pronouncedUFSFtokens"
    Save as comma-separated file: "AdultOTandNN/Distributions/"+grammar$+".csv"
    Save as comma-separated file: "ChildOTandNN/Distributions/"+grammar$+".csv"
    Remove
    selectObject: "OTGrammar "+grammar$
    Save as text file: "AdultOTandNN/OTGrammars/"+grammar$+".OTGrammar"
    Remove

  endfor
  selectObject: "Strings list"
  Remove
```

8.5.3 Child perception and production

```
# Joachim Kokkelmans 03.03.2020. Based on scripts by Paul Boersma and
```

```
Klaas Seinhorst.
# Requires: A folder InitialObjects at the same level as this script containing the OTGrammar InitialChildGrammar.txt and a folder OTGrammars containing 1s.OTGrammar, 2sAB.OTGrammar etc.; a folder AdultOTandNN containing a folder Distributions (containing 1s.csv, 2sAB.csv etc.) and a folder OTGrammars containing the OTGrammars of the five languages of the 'adult speaker'; a folder ChildOTandNN containing a folder Distributions, a folder DistributionStages, a folder GrammarStages and a folder NetworkStages; the uppermost folder level contains InitialObjects, AdultOTandNN, ChildOTandNN, Run.py, AdultOTandNN.praat and ChildOTandNN.praat.

# Parameters of the simulation in general (those of the neural network are below).

   form Generation
      integer gennumber 1
   endform

# Loading the AdultDistributions files in the directory (adapted from a script by Kevin Ryan).

   # The learning child has access to the distributions of the adult speech i.e. what (s)he has heard from the adult, but obviously not to underlying information (SF or UF) nor to the grammar itself (OTGrammar objects).
   sourcedirectory$ = "ChildOTandNN/Distributions"
   filenameorinitialsubstring$ = ""
   fileextension$ = ".csv"

# Here below, a list of all sound files in the 'ChildOTandNN/Distributions' folder is created.

   Create Strings as file list... list 'sourcedirectory$'/'filenameorinitialsubstring$'*'fileextension$'
   file_count = Get number of strings

   for current_file to file_count
      select Strings list
      filename$ = Get string... current_file
      Read Table from comma-separated file: sourcedirectory$+"/"+filename$
      grammar$ = filename$ - ".csv"

# Here, the 'real stuff' happens, for the respective initial Distribution has just been loaded.

### PRELEXICAL PERCEPTION: the child does not speak yet. (S)he hears the adult distribution as auditory input, updates the connection weights within the network accordingly and will start to posit what the phoneme categories are. The UF nodes are created later, when all the input tokens have been heard once.

```

```
	# A Neural Network is created, which will receive all the
	   perceived values from the adult's distribution (thereby
	   tuning the perception of the child learner).
	# It has an UF layer with a number of nodes corresponding to the
	   number of categories (0 at first, then created after learning
	   ), an SF layer and an AudF layer.

	# Parameters of the network.

	sf.NbOfNodes = 12
	audf.NbOfNodes = 32 ; shall be divisible by 4, in order not to
	   complicate calculations with A, B, C and D as equally wide
	   regions.
	if (audf.NbOfNodes/4) <> round(audf.NbOfNodes/4)
	   exitScript: "The number of AudF nodes must be a multiple of
	      four."
	endif
	spreadpercentage = 0.05

	### Creating the network.
	Select outer viewport: 0, 12, 0, 7
	Create empty Network: grammar$, 0.01, "linear", 0, 1, 1, 0.01,
	   -1, 1, 0, 0, 10, 0, 5
	for i to audf.NbOfNodes
	   Add node: (10/audf.NbOfNodes)*(i-1)+(5/audf.NbOfNodes), (5/3)
	      +(5/6), 0, "yes" ; the AudF nodes are clamped from the
	      start.
	endfor
	firstAudF = 1
	lastAudF = audf.NbOfNodes
	for i to sf.NbOfNodes
	   Add node: (10/sf.NbOfNodes)*(i-1)+(5/sf.NbOfNodes), (5/3)
	      +(5/6)+(10/6), 0, "no" ; the SF nodes are unclamped from
	      the start.
	endfor
	firstSF = audf.NbOfNodes+1
	lastSF = audf.NbOfNodes+sf.NbOfNodes
	Set instar: 0.5
	Set outstar: 0.5
	Set weight leak: 0
	Set activity leak: 1
	Set shunting: 0

	### Creating the connections between SF and SF nodes with fixed
	   -0.1 weights (as in Boersma et al. 2018), then between SF and
	    AudF with flexible random (0-0.1) weights.
	for i from firstSF to lastSF
	   for j from firstSF to lastSF
	         Add connection: i, j, -0.1, 0
	   endfor
	endfor
	for i from firstSF to lastSF
	   for j from firstAudF to lastAudF
```

```
            randomweight = randomUniform(0, 0.1)
            Add connection: i, j, randomweight, 1
         endfor
      endfor

      ### Now, the child hears the tokens pronounced by the adult. They are taken one by one from the Distribution file. The mispronunciations of small children are not heard, because with the small number of tokens (less than 100000) used in these simulations that exerts an influence to increase the proportion of B pronunciations at every generation. In real life, we hear billions of sibilants and are thus not significantly influenced by the few initial mispronunciations.
      selectObject: "Table "+grammar$
      Extract rows where column (text): "NbMismatches", "is equal to", ""
      Rename: "temp"
      selectObject: "Table "+grammar$
      Remove
      selectObject: "Table temp"
      Rename: grammar$
      numrows = Get number of rows
      Randomize rows
      for i to numrows
         selectObject: "Table "+grammar$
         heardvalue = Get value: i, "PV"
         # This translates the percentual formulation of the perceived value (from 0 to 100) onto the Audf node continuum (from 1 to audf.NbOfNodes).
         heardvalue = (heardvalue/100*(audf.NbOfNodes-1))+1
         # Feeding the network with the perceived token.
         selectObject: "Network "+grammar$
         Zero activities: 0, 0
         for j from firstAudF to lastAudF
            activity = exp(-0.5 * (j-heardvalue) ^ 2 / (spreadpercentage*(audf.NbOfNodes-1)) ^2)
            Set activity: j, activity
         endfor
         Spread activities: 100
         Update weights
      endfor

### POSTLEXICAL PERCEPTION: the child does not speak yet. (S)he posits what the existing phoneme categories are, then hears the adult distribution as auditory input again, updating the connection weights between SF and UF.

      # Pacing through AudF to see which SF nodes get activated simultaneously, and grouping their numbers in one string for each category.
      currentpattern$ = ""
      nbOfPatterns = 0
```

```
      for perceivedNode to audf.NbOfNodes
         Zero activities: 0, 0
         for i from firstAudF to lastAudF
            activity = exp(-0.5 * (i-perceivedNode) ^ 2 / (
                spreadpercentage*(audf.NbOfNodes-1)) ^2)
            Set activity: i, activity
         endfor
         Spread activities: 100
         nodepatternstring$ = ""
         for i from firstSF to lastSF
            thisactivity = Get activity: i
            if thisactivity > 0.5
               nodepatternstring$ = nodepatternstring$+string$(i)+"_"
            endif
         endfor
         if nodepatternstring$ <> "" && nodepatternstring$ <>
             currentpattern$
            currentpattern$ = nodepatternstring$
            nbOfPatterns = nbOfPatterns + 1
            patterns$[nbOfPatterns] = nodepatternstring$
         endif
      endfor

      ### Now the patterns$ array contains each different activation
          pattern, but also their combination (e.g. in the transition
          from one category to another, both patterns are activated and
           this is thus counted as a separate pattern). I eliminate
          these in what follows.
      currentpattern$ = ""
      nbOfFilteredPatterns = 0
      for i to nbOfPatterns
         for j to nbOfPatterns
            if i <> j
               foundadifferentone = 0
               firststring$ = patterns$[j]
               secondstring$ = patterns$[i]
               underscore = index(firststring$, "_")
               while underscore <> 0
                  thisnodenumber$ = left$(firststring$, underscore)
                  foundinthesecondstring = index(secondstring$,
                      thisnodenumber$)
                  if foundinthesecondstring = 0
                     foundadifferentone = 1
                  endif
                  firststring$ = mid$(firststring$, (underscore+1), (
                      length(firststring$)-underscore))
                  underscore = index(firststring$, "_")
               endwhile
               if foundadifferentone = 1 && patterns$[i] <>
                   currentpattern$
                  currentpattern$ = patterns$[i]
                  nbOfFilteredPatterns = nbOfFilteredPatterns + 1
                  Create Table with column names: "Category"+string$(
```

```
                    nbOfFilteredPatterns), 0, "Node"
                underscore = index(secondstring$, "_")
                while underscore <> 0
                   thisnodenumber$ = left$(secondstring$, (underscore
                      -1))
                   Append row
                   numnewrow = Get number of rows
                   Set numeric value: numnewrow, "Node", number(
                      thisnodenumber$)
                   secondstring$ = mid$(secondstring$, (underscore+1)
                      , (length(secondstring$)-underscore))
                   underscore = index(secondstring$, "_")
                endwhile
             endif
          endif
       endfor
    endfor
    ### The if i <> j condition above excludes the single category in
         a 1s, because of course the only pattern is not different
        from itself. In that case, I add the category here.
    if nbOfPatterns = 1
       Create Table with column names: "Category1", 0, "Node"
       thisstring$ = patterns$[1]
       underscore = index(thisstring$, "_")
       while underscore <> 0
          thisnodenumber$ = left$(thisstring$, (underscore-1))
          Append row
          numnewrow = Get number of rows
          Set numeric value: numnewrow, "Node", number(
             thisnodenumber$)
          thisstring$ = mid$(thisstring$, (underscore+1), (length(
             thisstring$)-underscore))
          underscore = index(thisstring$, "_")
       endwhile
       nbOfFilteredPatterns = 1
    endif

    ### Looping through the SF categories and seeing which SF
        category corresponds to which out of A, B, C or D.
    thereisA = 0
    thereisB = 0
    thereisC = 0
    thereisD = 0
    selectObject: "Network "+grammar$
    for i from firstAudF to lastAudF
       Set clamping: i, "no"
    endfor
    for i from firstSF to lastSF
       Set clamping: i, "yes"
    endfor
    minnodespercategory = audf.NbOfNodes
    for i to nbOfFilteredPatterns
       thisisA = 0
```

```
        thisisB = 0
        thisisC = 0
        thisisD = 0
        selectObject: "Network "+grammar$
        Zero activities: 0, 0
        selectObject: "Table Category"+string$(i)
        numberofrows = Get number of rows
        minnodespercategory[i] = numberofrows
        if numberofrows < minnodespercategory
           minnodespercategory = numberofrows
        endif
        for j to numberofrows
           thisnodenumber = Get value: j, "Node"
           selectObject: "Network "+grammar$
           Set activity: thisnodenumber, 1
           selectObject: "Table Category"+string$(i)
        endfor
        selectObject: "Network "+grammar$
        Spread activities: 100
        for j from firstAudF to lastAudF
           thisactivity = Get activity: j
           if j <= (audf.NbOfNodes/4)
              thisisA = thisisA + thisactivity
           elsif j > (audf.NbOfNodes/4) && j <= (2*(audf.NbOfNodes/4))
              thisisB = thisisB + thisactivity
           elsif j > (2*(audf.NbOfNodes/4)) && j <= (3*(audf.NbOfNodes
              /4))
              thisisC = thisisC + thisactivity
           elsif j > (3*(audf.NbOfNodes/4))
              thisisD = thisisD + thisactivity
           endif
        endfor
        # It still is possible that two different categories have been
             assigned the same letter (e.g. a rather retracted A as B
            and a rather fronted C as B also, so potential conflict in
             sight). To solve this, the number of the category is also
             kept.
        selectObject: "Table Category"+string$(i)
        maxcategory = imax(thisisA, thisisB, thisisC, thisisD)
        if maxcategory = 1
           Rename: "Category"+string$(i)+"A"
           category$[i] = "Category"+string$(i)+"A"
           thereisA = 1
        elsif maxcategory = 2
           Rename: "Category"+string$(i)+"B"
           category$[i] = "Category"+string$(i)+"B"
           thereisB = 1
        elsif maxcategory = 3
           Rename: "Category"+string$(i)+"C"
           category$[i] = "Category"+string$(i)+"C"
           thereisC = 1
        elsif maxcategory = 4
           Rename: "Category"+string$(i)+"D"
```

```
            category$[i] = "Category"+string$(i)+"D"
            thereisD = 1
         endif
      endfor

      ### Now, the child hears 10000 more (randomly selected) tokens,
         updates the weights and stores the information about the
         frequency of occurrence of each category.
      selectObject: "Network "+grammar$
      for i from firstAudF to lastAudF
         Set clamping: i, "yes"
      endfor
      for i from firstSF to lastSF
         Set clamping: i, "no"
      endfor
      selectObject: "Table "+grammar$
      Randomize rows
      catstring$ = ""
      for i to nbOfFilteredPatterns
         category[i] = 0
         catstring$ = catstring$ + category$[i]
      endfor
      for i to 10000
         selectObject: "Table "+grammar$
         heardvalue = Get value: i, "PV"
         # This translates the percentual formulation of the perceived
            value (from 0 to 100) onto the Audf node continuum (from 1
             to audf.NbOfNodes).
         heardvalue = (heardvalue/100*(audf.NbOfNodes-1))+1
         # Feeding the network with the perceived token.
         selectObject: "Network "+grammar$
         Zero activities: 0, 0
         for j from firstAudF to lastAudF
            activity = exp(-0.5 * (j-heardvalue) ^ 2 / (
               spreadpercentage*(audf.NbOfNodes-1)) ^2)
            Set activity: j, activity
         endfor
         Spread activities: 100
         maxaverageactivityofthiscategory = 0
         for j to nbOfFilteredPatterns
            averageactivityofthiscategory = 0
            selectObject: "Table "+category$[j]
            numberofrows = Get number of rows
            for k to numberofrows
               thisnodenumber = Get value: k, "Node"
               selectObject: "Network "+grammar$
               thisactivity = Get activity: thisnodenumber
               averageactivityofthiscategory =
                  averageactivityofthiscategory + thisactivity
               selectObject: "Table "+category$[j]
            endfor
            averageactivityofthiscategory =
               averageactivityofthiscategory / k
```

```
            if averageactivityofthiscategory > maxaverageactivityofthiscategory
               maxaverageactivityofthiscategory = averageactivityofthiscategory
               itisthiscategory = j
            endif
         endfor
         category[itisthiscategory] = category[itisthiscategory] + 1
         selectObject: "Network "+grammar$
         Update weights
      endfor
      for i to nbOfFilteredPatterns
         category[i] = category[i] / 10000
      endfor
      # A sample calculation of the accuracy of the network in identifying the intended category (SF) based on the perceived token (PV) gave a difference of 0.0005 (0.05%) as margin error (for 2 sibilants, B and C).

### CHILD PRODUCTION: the child tries to speak and, having heard the adult say X times one sibilant or Y times another, tries to pronounce the same sibilants (not necessarily in the same order, but with the same lexical frequencies). To this end, the relevant SF nodes are activated.

      # The ArtF nodes are created. They penalise posterior and peripheral sibilants gradually by means of connections to AudF with inhibition. They have the same strength (i.e. the same acceleration from least marked to most marked, namely -x/124).
      selectObject: "Network "+grammar$
      Add node: 1, (5/3)-(5/6), 1, "yes"
      frontnode = audf.NbOfNodes+sf.NbOfNodes+1
      for i from ((audf.NbOfNodes/4)+1) to lastAudF
         Add connection: i, frontnode, 0.25*(-(i-((audf.NbOfNodes-1)/4)-1)/(audf.NbOfNodes-1))*((i-((audf.NbOfNodes-1)/4)-1)/(audf.NbOfNodes-1)), 0
      endfor
      Add node: 5, (5/3)-(5/6), 1, "yes"
      centralnode = frontnode+1
      for i from firstAudF to lastAudF
         Add connection: i, centralnode, 0.25*(-abs(i-((audf.NbOfNodes-1)/2)-1)/(audf.NbOfNodes-1))*(abs(i-((audf.NbOfNodes-1)/2)-1)/(audf.NbOfNodes-1))-0.8, 0
      endfor
      selectObject: "Network "+grammar$
      for i from firstAudF to lastAudF
         Set clamping: i, "no"
      endfor
      for i from firstSF to lastSF
         Set clamping: i, "yes"
      endfor

```

```
      # Creating a table to record every input - output ever produced by the child. It also records the number of mismatches (i.e. of non-faithful mappings), the ranking value of the five constraints (Central(Sib), Ident(Sib), PostCD(Sib), PostD(Sib), *Sib) and the category tags.
      Create Table with column names: "pronouncedUFSFtokens", 0, "UF SF PV NbA NbB NbC NbD Nb. NbMismatches Constr1 Constr2 Constr3 Constr4 Constr5 CatString"
      # The current child grammar must start as tabula rasa with all markedness constraints ranked highest (120 for markedness vs. 100 for faithfulness), copied from the model 'InitialChildGrammar' OTGrammar object.
      Read from file: "InitialObjects/InitialChildGrammar.txt"
      Rename: "ChildGrammar"
      numconstr = Get number of constraints
      ### For the same number of tokens as the number of tokens heard, with approximatively the same functional load per category as in the 'mother language', the child starts trying to articulate with all markedness constraints ranked above faithfulness. Only non-sibilants come out at first...
      num["A"] = 0
      num["B"] = 0
      num["C"] = 0
      num["D"] = 0
      num["."] = 0
      nummismatches = 0
      # This below makes sure that the proportion of functional load of the A, B, C and D categories will stay the same (otherwise there is a statistical bias towards B later in the script).
      for i to nbOfFilteredPatterns
         produced[category$[i]] = 0
         maxnum[category$[i]] = ceiling(category[i] * numrows)
      endfor
      i = 1
      exitthisscript = 0
      # Setting the phonetic deviation from the intended token (noise in the signal).
      phondev = (audf.NbOfNodes-1)/10
      repeat
         # The child starts to speak and wants to pronounce the sibilants heard in her/his initial distribution (with the same token frequency, already calculated as category[i] above).
         repeat
            randcat = randomInteger(1, nbOfFilteredPatterns)
            randzeroone = randomUniform(0, 1)
         until category[randcat] > randzeroone && produced[category$[randcat]] <= maxnum[category$[randcat]]
         selectObject: "OTGrammar ChildGrammar"
         pronouncedUF$ = right$(category$[randcat], 1)
         pronouncedSF$ = Input to output: pronouncedUF$, 2
         ### Updating the table with the input - output pairs pronounced.
```

```
         selectObject: "Table pronouncedUFSFtokens"
         Append row
         numrows2 = Get number of rows
         Set string value: numrows2, "UF", pronouncedUF$
         Set string value: numrows2, "SF", pronouncedSF$
         num[pronouncedSF$] = num[pronouncedSF$] + 1
         Set numeric value: numrows2, "NbA", num["A"]
         Set numeric value: numrows2, "NbB", num["B"]
         Set numeric value: numrows2, "NbC", num["C"]
         Set numeric value: numrows2, "NbD", num["D"]
         Set numeric value: numrows2, "Nb.", num["."]
         ### Checking that the pronouncedSF$ token is not a category which does not exist phonemically (e.g. B in an AC language). If it is, a pronounced value is chosen around the centre of the non-phonemic category.
         mispronounced = 0
         if pronouncedSF$ = "A" && thereisA = 0
            mispronounced = 1
            thistoken = randomGauss(((audf.NbOfNodes-1)/8)/(audf.NbOfNodes-1)*100, phondev)
         elsif pronouncedSF$ = "B" && thereisB = 0
            mispronounced = 1
            thistoken = randomGauss((((audf.NbOfNodes-1)/4 + (audf.NbOfNodes-1)/8))/(audf.NbOfNodes-1)*100, phondev)
         elsif pronouncedSF$ = "C" && thereisC = 0
            mispronounced = 1
            thistoken = randomGauss((((audf.NbOfNodes-1)/2 + (audf.NbOfNodes-1)/8))/(audf.NbOfNodes-1)*100, phondev)
         elsif pronouncedSF$ = "D" && thereisD = 0
            mispronounced = 1
            thistoken = randomGauss((((audf.NbOfNodes-1)/2 + 3*(audf.NbOfNodes-1)/8))/(audf.NbOfNodes-1)*100, phondev)
         endif
         if mispronounced = 1
            selectObject: "Table pronouncedUFSFtokens"
            Set numeric value: numrows2, "PV", thistoken
         endif
         if pronouncedSF$ <> "." && mispronounced = 0
            selectObject: "Network "+grammar$
            Zero activities: 0, 0
            for j from frontnode to centralnode
               Set activity: j, 1
            endfor
            selectObject: "Table "+category$[randcat]
            numberofrows = Get number of rows
            for j to numberofrows
               thisnodenumber = Get value: j, "Node"
               selectObject: "Network "+grammar$
               Set activity: thisnodenumber, (minnodespercategory/minnodespercategory[randcat])
               selectObject: "Table "+category$[randcat]
            endfor
            selectObject: "Network "+grammar$
```

```
            Spread activities: 500
            # Choosing a random AudF node and if its activation is
               larger than a random randomUniform(0, 1) number, it is
               the token produced.
            repeat
               security = 0
               repeat
                  if security > 5000
                     selectObject: "Network "+grammar$
                     Save as text file: "error"+grammar$+string$(
                        gennumber)+".Network"
                     exitthisscript = 1
                  endif
                  security = security + 1
                  randnode = randomInteger(firstAudF, lastAudF)
                  thisactivity = Get activity: randnode
                  randzeroone = randomUniform(0, 1)
               until thisactivity > randzeroone  exitthisscript = 1
               thistoken = randomGauss(((randnode-1)/(audf.NbOfNodes-1)
                  *100), phondev)
            until (thistoken > 0 && thistoken < 100)  exitthisscript =
               1
            selectObject: "Table pronouncedUFSFtokens"
            Set numeric value: numrows2, "PV", thistoken
         endif
         ### Life is hard, and sibilants hard to pronounce. If I did
            not manage to produce a sibilant now, I will improve by
            lowering the constraints which were penalising the correct
             output and raising the constraints which were favouring
            it.
         if pronouncedUF$ <> pronouncedSF$
            nummismatches = nummismatches + 1
            selectObject: "Table pronouncedUFSFtokens"
            Set numeric value: numrows2, "NbMismatches", nummismatches
            selectObject: "OTGrammar ChildGrammar"
            Learn one: pronouncedUF$, pronouncedUF$, 2, "Symmetric all
               ", 0.1, 0, "yes"
            for j to numconstr
               constranking = Get ranking value: j
               selectObject: "Table pronouncedUFSFtokens"
               Set numeric value: numrows2, "Constr"+string$(j),
                  constranking
               selectObject: "OTGrammar ChildGrammar"
            endfor
         else
            i = i + 1
            produced[category$[randcat]] = produced[category$[randcat]]
                + 1
         endif
      until i = (numrows + 1)  exitthisscript = 1

      ### Saving and removing the objects.
      if exitthisscript = 0
```

```
            selectObject: "Table pronouncedUFSFtokens"
            Set string value: numrows2, "CatString", catstring$
            Save as comma-separated file: "ChildOTandNN/DistributionStages
                /"+grammar$+string$(gennumber)+".csv"
            Save as comma-separated file: "ChildOTandNN/Distributions/"+
                grammar$+".csv"
            Remove
            selectObject: "Network "+grammar$
            Save as text file: "ChildOTandNN/NetworkStages/"+grammar$+
                string$(gennumber)+".Network"
            Remove
            selectObject: "OTGrammar ChildGrammar"
            Save as text file: "ChildOTandNN/GrammarStages/ChildGrammar"+
                grammar$+string$(gennumber)+".txt"
            Remove
            selectObject: "Table "+grammar$
            for i to nbOfFilteredPatterns
                plusObject: "Table "+category$[i]
            endfor
            Remove
        endif

    endfor
    selectObject: "Strings list"
    Remove
```

8.5.4 Script for a button that allows to pace through AudF

Note: this script modifies itself (and shall thus be modified with caution to avoid data loss). It works whenever a network object is selected in Praat.

```
               form Pace through AudF
positive Standard_Deviation 0.04
integer Nb._of_first_AudF_node 1
integer Nb._of_last_AudF_node 32
boolean There_are_only_AudF_and_SF_nodes 1
integer If_not_Nb._of_first_SF_node 33
integer If_not_Nb._of_last_SF_node 44
boolean Clamp_all_other_nodes 1
endform
# Above, the info is asked to the user. Every time this script is run,
    the form above is replaced by the same form with the values used by
     the user.

# Calculating the total number of nodes in the network (by listing the
    nodes, then extracting the last number in the list).
   nodelist$ = List nodes: 1, 1000, "yes", "no", "no", 6, "no", "no", "
       no", 6
   nodelist$ = left$(nodelist$, (length(nodelist$)-1))
   endline$ = mid$(nodelist$, 5, 1)
   lastendline = rindex(nodelist$, endline$)
   nodelist$ = right$(nodelist$, (length(nodelist$)-lastendline))
   numberofnodes = number(nodelist$)
```

```

# Calculating the number of AudF nodes and SF nodes with shorter variable names.
   firstAudF = nb._of_first_AudF_node
   lastAudF = nb._of_last_AudF_node
   firstSF = if_not_Nb._of_first_SF_node
   lastSF = if_not_Nb._of_last_SF_node
   standev = standard_Deviation

# Calculating, if There_are_only_AudF_and_SF_nodes = 1, what the range of the SF nodes is.
   if there_are_only_AudF_and_SF_nodes = 1
      if firstAudF = 1
         firstSF = lastAudF + 1
         lastSF = numberofnodes
      else
         firstSF = 1
         lastSF = firstAudF - 1
      endif
   endif

# Calculating the distance between the first and last nodes.
   distanceAudF = lastAudF - firstAudF
   distanceSF = lastSF - firstSF

#Checking that the standard deviation is located between 0 (exclusive) and 1 (exclusive)
   if standev <= 0 || standev >= 1
      exitScript: "The standard deviation must be located between 0 (exclusive) and 1 (exclusive)."
   endif

# Checking that the node numbers specified are not 0 or less, because it would make no sense.
   if firstAudF < 1 || lastAudF < 1
      exitScript: "Negative or nihil node numbers are not allowed."
   endif
   if there_are_only_AudF_and_SF_nodes = 0
      if firstSF < 1 || lastSF < 1
         exitScript: "Negative or nihil node numbers are not allowed."
      endif
   endif

# Checking that the user did not specify a first node with a higher number than the corresponding last node.
   if firstAudF > lastAudF
      exitScript: "The number of the first node(s) must be lower than the number of the last node(s)."
   endif
   if there_are_only_AudF_and_SF_nodes = 0
      if firstSF > lastSF
         exitScript: "The number of the first node(s) must be lower than the number of the last node(s)."
```

```
      endif
   endif

# Checking that the user did not specify node ranges which are larger (
    and smaller, if There_are_only_AudF_and_SF_nodes = 1) than the
    total number of nodes. Then, if the node ranges are contiguous.
   if (distanceAudF + distanceSF) > numberofnodes
      exitScript: "The number of AudF and SF nodes cannot exceed the
          total number of nodes in the network."
   endif
   if there_are_only_AudF_and_SF_nodes = 1
      if (distanceAudF + distanceSF) <> (numberofnodes - 2)
         exitScript: "The node ranges must be contiguous. The AudF node
              numbers you gave are probably wrong."
      endif
   endif

# The network is prepared: all nodes are clamped except the SF nodes.
   if clamp_all_other_nodes = 1
      for i to numberofnodes
         Set clamping: i, "yes"
      endfor
      for i from firstSF to lastSF
         Set clamping: i, "no"
      endfor
   else
      for i to numberofnodes
         Set clamping: i, "no"
      endfor
      for i from firstAudF to lastAudF
         Set clamping: i, "yes"
      endfor
   endif

# One loops from firstAudF to lastAudF, erasing all previous activities
    , then setting the new activities and spreading them.
   for perceivedNode from firstAudF to lastAudF
      Zero activities: 1, 0
      for i from firstAudF to lastAudF
         activity = exp(-0.5 * (i-perceivedNode) ^ 2 / (standev*
             distanceAudF) ^2)
         Set activity: i, activity
      endfor
      Spread activities: 100
      thisobject$ = selected$("Network")
      Copy: thisobject$+"-"+string$(perceivedNode)
      selectObject: "Network "+thisobject$
   endfor

# Drawing the demo window.
   demoWindowTitle: "Pacing through the AudF nodes of network "+
       thisobject$
   demo Erase all
```

```
demo Black
demo Times
demo 24
demo Select inner viewport: 0, 100, 0, 100
demo Axes: 0, 100, 0, 100
selectObject: "Network "+thisobject$
demo Draw: "yes"
@displayinfo
currentAudF = lastAudF
while demoWaitForInput ( )
   if demoKey$() = →""
      demo Erase all
      if currentAudF = lastAudF
         currentAudF = firstAudF
      else
         currentAudF = currentAudF + 1
      endif
      selectObject: "Network "+thisobject$+"-"+string$(currentAudF)
      demo Draw: "yes"
      @displayinfo
   endif
   if demoKey$() = ←""
      demo Erase all
      if currentAudF = firstAudF
         currentAudF = lastAudF
      else
         currentAudF = currentAudF - 1
      endif
      selectObject: "Network "+thisobject$+"-"+string$(currentAudF)
      demo Draw: "yes"
      @displayinfo
   endif
   if demoClickedIn (90, 100, 0, 10)
      goto END
   endif
endwhile
procedure displayinfo
   demo Select inner viewport: 0, 100, 0, 100
   demo Axes: 0, 100, 0, 100
   demo Text: 50, "centre", 92, "half", "Use the left or right
       arrows on your keyboard to move leftwards or rightwards."
   demo Paint rectangle: "grey", 90, 100, 0, 10
   demo Draw line: 90, 10, 100, 10
   demo Draw line: 90, 0, 100, 0
   demo Draw line: 90, 10, 90, 0
   demo Draw line: 100, 10, 100, 0
   demo Text: 95, "centre", 5, "half", "Exit"
endproc
label END
demo Select inner viewport: 0, 100, 0, 100
demo Axes: 0, 100, 0, 100
demo Paint rectangle: "white", 0, 100, 0, 100
demo Text: 50, "centre", 50, "half", "Please close this window."
```

```

# Deleting the networks created.
   for i from firstAudF to lastAudF
      selectObject: "Network "+thisobject$+"-"+string$(i)
      Remove
   endfor
   selectObject: "Network "+thisobject$

# Replacing the values of the form at the beginning of this script.
   thisverysamefile$ = readFile$("PaceThroughAudF.praat")
   formstop = index(thisverysamefile$, "endform")
   totallength = length(thisverysamefile$)
   rightform$ = mid$(thisverysamefile$, (formstop+8), (totallength-(
      formstop+7)))
   writeFileLine: "PaceThroughAudF.praat", "form Pace through AudF"
   appendFileLine: "PaceThroughAudF.praat", "positive
      Standard_Deviation "+string$(standev)
   appendFileLine: "PaceThroughAudF.praat", "integer Nb.
      _of_first_AudF_node "+string$(firstAudF)
   appendFileLine: "PaceThroughAudF.praat", "integer Nb.
      _of_last_AudF_node "+string$(lastAudF)
   appendFileLine: "PaceThroughAudF.praat", "boolean
      There_are_only_AudF_and_SF_nodes "+string$(
      there_are_only_AudF_and_SF_nodes)
   appendFileLine: "PaceThroughAudF.praat", "integer If_not_Nb.
      _of_first_SF_node "+string$(firstSF)
   appendFileLine: "PaceThroughAudF.praat", "integer If_not_Nb.
      _of_last_SF_node "+string$(lastSF)
   appendFileLine: "PaceThroughAudF.praat", "boolean
      Clamp_all_other_nodes "+string$(clamp_all_other_nodes)
   appendFileLine: "PaceThroughAudF.praat", "endform"
   appendFile: "PaceThroughAudF.praat", rightform$
```

8.6 VBA Functions

8.6.1 Gradual Faithfulness

```
Function AGradualFaith(sInput As String, sOutput As String) As Long

sInput = LCase(sInput)
sOutput = LCase(sOutput)
Dim InputLen As Long
Dim OutputLen As Long
InputLen = Len(sInput)
OutputLen = Len(sOutput)

Dim nbofviolations As Long
nbofviolations = 0
Dim inc As Long
Dim thischar As String
```

```
Dim thatchar As String

If InputLen = OutputLen Then

    For inc = 1 To Len(sOutput)

        Dim this As Long
        this = CLng(inc)
        thischar = Mid(sInput, this, 1)
        thatchar = Mid(sOutput, this, 1)
        If Asc(thischar) > 96 And Asc(thatchar) > 96 And Asc(thischar)
            < 123 And Asc(thatchar) < 123 Then
            nbofviolations = nbofviolations + (Abs(Asc(thischar) - Asc(
                thatchar)))
        End If
        If (thischar = "n" Or thatchar = "n") And thischar <> thatchar
            Then
            nbofviolations = 4
        End If

    Next

Else

    nbofviolations = 666
    Dim alert As Integer
    alert = MsgBox("Input and Output should have the same length: No go
        !", vbOKOnly, "Error")

End If

AGradualFaith = nbofviolations

End Function
```

8.7 Sibilant inventory combinations for Shibboleth

Gileadites	**Ephraimites**	**V if possible, explanation if impossible**	**Combined probability based on statistical frequency in the SibInv**
B	B	No mismatch i.e. no reason for mispronunciations	
	AB	The Ephraimites would have been able to pronounce B	
	AC	The Gileadites hear a sibilant as another, so could not have only one	
	BC	The Ephraimites would have been able to pronounce B	
	ABC	The Ephraimites would have been able to pronounce B	
AB	B	V	$\frac{19}{258} \cdot \frac{60}{258} = 1.71\%$
	AB	No mismatch i.e. no reason for mispronunciations	
	AC	V	$\frac{19}{258} \cdot \frac{129}{258} = 3.68\%$
	BC	V	$\frac{19}{258} \cdot \frac{9}{258} = 0.26\%$
	ABC	The Ephraimites would have been able to pronounce A and B	
AC	B	V	$\frac{129}{258} \cdot \frac{60}{258} = 11.63\%$
	AB	V	$\frac{129}{258} \cdot \frac{19}{258} = 3.68\%$
	AC	No mismatch i.e. no reason for mispronunciations	
	BC	V	$\frac{129}{258} \cdot \frac{9}{258} = 1.74\%$
	ABC	The Ephraimites would have been able to pronounce A and C	
BC	B	V	$\frac{9}{258} \cdot \frac{60}{258} = 0.81\%$
	AB	V	$\frac{9}{258} \cdot \frac{19}{258} = 0.26\%$
	AC	V	$\frac{9}{258} \cdot \frac{129}{258} = 1.74\%$
	BC	No mismatch i.e. no reason for mispronunciations	
	ABC	The Ephraimites would have been able to pronounce B and C	
ABC	B	V	$\frac{25}{258} \cdot \frac{60}{258} = 2.25\%$
	AB	V	$\frac{25}{258} \cdot \frac{19}{258} = 0.71\%$
	AC	V	$\frac{25}{258} \cdot \frac{129}{258} = 4.84\%$
	BC	V	$\frac{25}{258} \cdot \frac{9}{258} = 0.34\%$
	ABC	No mismatch i.e. no reason for mispronunciations	

Table 36: *Possible combinations of 1s, 2s and 3s to account for the* Shibboleth *story, with an explanation if the combination would not have been able to yield the* Shibboleth-*confusion and the statistical probability based on the SibInv database if possible.*

8.8 All inventories in the database

8.8.1 All inventories

The number in front of the inventories presented here corresponds to the ID of the inventory in the typological database.

(1) All Australian languages (except Torres strait languages) (Dixon 2002: 63f., 606)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.								
Sib. fri.								
N.-sib. fri.		θ̠ ð̠						

(2) Hawaiian (Gussenhoven and Jacobs 2017; Schütz 1994)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.								
N.-sib. fri.								

(3) Proto-Dravidian (Steever 2019)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				
Sib. aff.								
Sib. fri.								
N.-sib. fri.								

(4) Old Tamil (Steever 2019)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				
Sib. aff.								
Sib. fri.								
N.-sib. fri.						ɻ̝̊		

(5) Older Malagasy (O'Neill 2015: 37)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.					t͡ɹ̝̊ d͡ɹ̝			
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(6) Older Mauritian Creole (Baker and Kriegel 2013: 253; Dufour et al. 2014)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(7) Proto-Indo-European (Vijūnas 2010)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̺					
N.-sib. fri.								

(8) Old Latin (Roberts 2012)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̺ z̺					
N.-sib. fri.								ɧ/x

(9) Latin (Galmés De Fuentes 1962: 117ff.; citealtVijunas2010)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̺					
N.-sib. fri.								ɧ/x

(10) Veronese regional standard Italian (Avesani et al. 2017)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪		t͡s̺ d͡z̺					
Sib. fri.			s̺ z̺					
N.-sib. fri.								

(11) Northern Italian dialects with dentalised /ts/ (northern Veneto) (Belloni 2006: 25)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̺ d͡z̺					
Sib. fri.			s̺ z̺					
N.-sib. fri.	θ ð							

(12) Castilian Spanish (Adams 1975; Colantoni and Kochetov 2011)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.			s̺					
N.-sib. fri.		θ̠ ð̠						ɧ/x

(13) Mexican Spanish (Lope 2004; Vijūnas 2010)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.			t͡ɬ	c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.				ç ʝ				

(14) Proto-Germanic (Vijūnas 2010)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̠ z̠					
N.-sib. fri.		θ̠ ð̠						ɧ/x ɣ

(15) Gothic (Wright 1910)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̠ z̠					
N.-sib. fri.		θ̠ ð̠						ɧ/x

(16) Old Norse (Vijūnas 2010)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.								
Sib. fri.			s̠					
N.-sib. fri.		θ̠ ð̠						

(17) Old Icelandic (Pétursson 1971)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.								
Sib. fri.			s̠					
N.-sib. fri.		θ̠ ð̠						

(18) Icelandic (Vijūnas 2010; Pétursson 1971)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				
Sib. aff.								
Sib. fri.			s̠					
N.-sib. fri.		θ̠ ð̠		ç				ɧ/x

(19) Danish (Basbøll 2005; Grønnum 2005: 305f.)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠					
Sib. fri.			s̠	ɕ				
N.-sib. fri.								

(20) Southern Scanian dialects (Lucazin 2010)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.	s̪			ɕ				
N.-sib. fri.								

(21) Older Low German (Adams 1975)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
Sib. fri.			s̠ z̠					
N.-sib. fri.				ç				

(22) West Frisian

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠		t͡ʃ			
Sib. fri.			s̠ z̠		ʃ			
N.-sib. fri.								ɧ/x

(23) Standard Hollandic Dutch (Collins and Mees 1999: 160; Kwakkel 2008; Ooijevaar 2011; Seinhorst and Ooijevaar 2011)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				
Sib. aff.			t͡s̠					
Sib. fri.			s̠ z̠	ɕ ʑ				
N.-sib. fri.								

(24) Flemish (Kokkelmans 2017)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				
Sib. aff.								
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.				ç				

(25) Younger Afrikaans (van Wyk et al. 1989; Wissing et al. 2015)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				
Sib. aff.					t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(26) Finnish (Adams 1975; Suomi et al. 2008: 27; Vijūnas 2010: 49)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠					
Sib. fri.			s̠					
N.-sib. fri.				ç				ɧ/x

(27) Estonian (Asu and Teras 2009: 368)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠					
Sib. fri.			s̠					
N.-sib. fri.								

(28) Ancient Greek

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(29) Athens (standard) Greek (Adams 1975)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠ z̠					
N.-sib. fri.		θ̠ ð̠		ç ʝ				

(30) Peloponnesian Late Byzantine Greek (Adams 1975)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠ z̠					
N.-sib. fri.		θ̠ ð̠						

(31) Proto-Celtic

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̠					
N.-sib. fri.								

(32) Modern Cornish (Adams 1975)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.	θ ð							ɧ/x

(33) Calcutta Bengali

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠					
N.-sib. fri.								

(34) Albanian (Orel 2000)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.	θ ð							

(35) Eastern Armenian

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			

N.-sib. fri.								

(36) Proto-Slavic

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɧ/x

(37) Proto-Baltic

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.	s̪ z̪			ɕ ʑ		ṣ ẓ		
N.-sib. fri.								

(38) Sanskrit (Hall 1997*a*)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.	s̪			ɕ		ṣ		
N.-sib. fri.								

(39) Farsi (Persian)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɧ/x ɣ

(40) Hindi (Koul 2008: 17)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɧ/x ɣ

(41) Bangladeshi Bengali (Khan 2010)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(42) Sinte Romani (Matras 2002: 40)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɧ/x

(43) Romanian (Chitoran 2002: 10)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɦ/x

(44) Cypriot Greek (Adams 1975; Arvaniti 2010)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.	θ ð							ɦ/x ɣ

(45) Kozani Greek (Baltazani et al. 2016)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.	θ ð			ç ʝ				

(46) Asia Minor Greek (Janse 2002)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ			
N.-sib. fri.	θ ð			ç				

(47) Turkish (Adams 1975)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(48) Maltese (Lucas and Čéplö 2020)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(49) Moroccan Arabic (Harris 1942; Lahrouchi 2018)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(50) Serbo-Croatian (Morén 2006)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ d͡ʑ		t͡ṣ d͡ẓ		
Sib. fri.	s̪ z̪					ṣ ẓ		
N.-sib. fri.								ɦ/x

(51) Ukrainian (Buk et al. 2008)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɦ/x

(52) Slovak (Pavlík 2004: 106; Hanulíková and Hamann 2010)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɦ/x

(53) Czech (Dankovičová 1997)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɦ/x

(54) Lithuanian (Ambrazas 2006)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(55) Latvian

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(56) Hungarian (Adams 1975; Siptár and Torkenczy 2000)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.				ç				

(57) Transylvanian Saxon (Scheiner 1895)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ			
N.-sib. fri.				ç				ɧ/x

(58) Luxemburgish (Gilles and Trouvain 2013)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(59) Standard German (Wiese 1996)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ			
N.-sib. fri.				ç				

(60) West Middle German regional varieties (Hall 2013; Jannedy and Weirich 2016, 2017; Conrad 2018)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(61) Irish (Hickey 2003: 268; citealt[214]Hickey2011)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.				ç ʝ				

(62) Standard Scottish Gaelic

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.				ç ʝ				ɧ/x ɣ

(63) Welsh (Hannahs 2013)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.			ɬ	ç				

(64) English (Cercignani 1983; Adams 1975)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.					t͡ɹ̝̊ d͡ɹ̝			
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.	θ ð							

(65) Breton

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.	θ ð							ɦ/x ɣ

(66) Standard French (Jakobson and Lotz 1949; Picard 1987; Goelzer 2005: 54; Capliez 2016: 83)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(67) Standard Italian (Vietti 2019)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(68) Romansh

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(69) Coastal Peruvian Spanish

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.	θ ð			ç ʝ				ɦ/x

(70) Portuguese

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(71) Younger Urban Norwegian (e.g. Bergen) (Kristoffersen 2000: 23; citealtSjevik2015)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.								

(72) Northern Sami

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ			
N.-sib. fri.	θ ð			ç ʝ				

(73) Faroese (Vijūnas 2010)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.			s̠		ʃ			
N.-sib. fri.			ɬ			ɰ̥˖		

(74) Archaic inland West Norwegian (e.g. Sogn og Fjordane or Setesdal)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				
Sib. aff.								
Sib. fri.			s̠					
N.-sib. fri.								

(75) Veneto regional standard Italian

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠					
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(76) Catalan (Vijūnas 2010; Recasens and Espinosa 2007; Benet et al. 2012; Bonet and Lloret 2018)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠		t͡ʃ d͡ʒ			
Sib. fri.			s̠ z̠		ʃ ʒ			
N.-sib. fri.								

(77) Old French (Adams 1975)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪				t͡ʃ d͡ʒ			
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(78) Middle High German (Adams 1975)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.

N.-sib. aff.								
Sib. aff.	t͡s̪				t͡ʃ			
Sib. fri.	s̪		s̠ z̠		ʃ			
N.-sib. fri.								ɧ/x

(79) Old Spanish (Adams 1975)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪				t͡ʃ			
Sib. fri.			s̠ z̠		ʃ ʒ			
N.-sib. fri.								

(80) Mirandese

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.	s̪ z̪		s̠ z̠		ʃ ʒ			
N.-sib. fri.								

(81) Eastern Galician

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.			s̠		ʃ			
N.-sib. fri.		θ̠ ð̠						

(82) Standard Basque (Euskara Batua) (Alvarez Enparantza 1982; de Rijk 2008: 8ff.)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪		t͡s̠		t͡ʃ			
Sib. fri.	s̪		s̠		ʃ			
N.-sib. fri.								

(83) Swedish (Alvarez Enparantza 1982; Hall 1997*a*)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.	s̪			ɕ		ṣ		
N.-sib. fri.								ɧ/x

(84) Archi (Berns 2013: 62)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.			t͡ɬ					
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.			ɬ ɮ					

(85) Polish (Hall 1997*b*; Ladefoged and Maddieson 1996; Żygis and Padgett 2010)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
Sib. aff.	t͡s̪ d͡z̪			t͡ɕ d͡ʑ		t͡ṣ d͡ẓ		
Sib. fri.	s̪ z̪			ɕ ʑ		ṣ ẓ		
N.-sib. fri.								

(86) Russian (Padgett and Żygis 2007: 296; Kochetov 2017: 322)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ				
Sib. fri.	s̪ z̪			ɕ		ṣ ẓ		
N.-sib. fri.								ɧ/x

(87) Alutiiq

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s						
N.-sib. fri.			ɬ					ɧ/x ɣ

(88) Ubykh (Vogt 1963; Ladefoged and Maddieson 1996; Colarusso 1999; Fenwick 2011)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠		t͡ʃ d͡ʒ		t͡ʂ d͡ʐ	
Sib. fri.	s̪ z̪		s̠ z̠		ʃ ʒ		ʂ ʐ	
N.-sib. fri.			ɬ					ɧ/x ɣ

(89) Toda (Emeneau 1984; Ladefoged and Maddieson 1996; Boersma and Hamann 2008)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.	s̪		s̠		ʃ ʒ		ʂ ʐ	
N.-sib. fri.	θ		ɬ					ɧ/x ɣ

(90) Waorani (Huaorani or Auca) (Peeke 1973)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.								
Sib. fri.								
N.-sib. fri.								

(91) Chukchee (Dunn 1999)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.			s̠					
N.-sib. fri.			ɬ					

(92) Ket

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
Sib. fri.			s̠					
N.-sib. fri.			ɬ ɮ	ç ʝ				ɦ/x ɣ

(93) Proto-Bantu (Creissels 1999: 314)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.								
Sib. fri.								
N.-sib. fri.								

(94) Tswana (Creissels 1999: 314)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.			t͡ɬ					
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ			
N.-sib. fri.			ɬ					

(95) Western Basque (Muxika-Loitzate 2017)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠		t͡ʃ			
Sib. fri.			s̠		ʃ			
N.-sib. fri.								

(96) Quebec French (Picard 1987: 46)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z						
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(97) Acadian French (Emmitte 2013)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(98) Trentino dialects (Alber and Kokkelmans in prep.)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.	s̪ z̪		s̠ z̠					
N.-sib. fri.								

(99) Karuk (Bright 1978)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠					
Sib. fri.	s̪		s̠					

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. fri.								

(100) Luiseno

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.	s̪		s̠		ʃ			
N.-sib. fri.								

(101) Canary Islands Spanish (Hualde 2005)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.		s						
N.-sib. fri.								

(102) Evolene dialect (Elmiger et al. 2013)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠					
Sib. fri.	s̪ z̪		s̠ z̠					
N.-sib. fri.								

(103) Zulu

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.			ɬ ɮ					

(104) Shona

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠		t͡ʃ d͡ʒ			
Sib. fri.	s̪ z̪		s̠ z̠		ʃ ʒ			
N.-sib. fri.								

(105) Xhosa (Bennett and Braver 2020)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.						ɬ̊ ɮ		

(106) Sotho

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ ʒ			
N.-sib. fri.						ɬ̊		

(107) Telugu

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.	s̪			ɕ		ṣ		
N.-sib. fri.								

(108) Southern German

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.			ɬ					

(109) Fijian

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.			ʂ					
N.-sib. fri.		θ̠ ð̠						

(110) Maori (Harlow 1996: 2)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.								
N.-sib. fri.								

(111) Tuamotuan (Kuki 1969: 48f.)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.								
N.-sib. fri.								

(112) Standard Mandarin Chinese (Duanmu 2000)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ		t͡ṣ		
Sib. fri.	s̪			ɕ		ṣ		
N.-sib. fri.						ɻ̝̊		

(113) Malay

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(114) Middle French (Adams 1975)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								

Sib. fri.	s̪ z̪		s̠ z̠		ʃ ʒ			
N.-sib. fri.								

(115) Finland Swedish

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.	s̪			ɕ				
N.-sib. fri.								

(116) Japanese (Toda 2009)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪			t͡ɕ d͡ʑ				
Sib. fri.	s̪ z̪			ɕ ʑ				
N.-sib. fri.				ç				

(117) Khalkha Mongolian (Svantesson 2006)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(118) Uzbek

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(119) Etruscan (Agostiniani 2014)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s						
Sib. fri.		s			ʃ			
N.-sib. fri.								

(120) Standard Greenlandic (Kalaallisut) (Vijūnas 2010: 49)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠					
Sib. fri.			s̠					
N.-sib. fri.								

(121) Walloon

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.				ç				

(122) Brazilian Portuguese

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(123) Kazakh

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.	s̪ z̪			ɕ ʑ				
N.-sib. fri.								

(124) Upper Sorbian (Jocz 2015: 174)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɦ/x ɣ

(125) Lower Sorbian

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ		t͡ṣ		
Sib. fri.	s̪ z̪			ɕ ʑ		ṣ ẓ		
N.-sib. fri.								

(126) Somali (Gabbard 2010)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.	θ ð							

(127) Tigrinya

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(128) Oromo

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(129) Standard Igbo

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.

N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(130) Bambara

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(131) Tashelhiyt Berber (Kossmann and Stroomer 1997)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(132) Oowekyala

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠					
N.-sib. fri.								

(133) Sandawe (Hunziker et al. 2008; Steeman 2012)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.			t͡ɬ d͡ɮ					
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s						
N.-sib. fri.			ɬ			ɻ̟̊		

(134) Komnzo (Döhler 2018)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠					
N.-sib. fri.		θ̠ ð̠						

(135) Yiddish (Kleine 2003)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.						ɻ̟̊		

(136) Modern Hebrew (Laufer 1990)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			

Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.						ɻ̊˖		

(137) Picard

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(138) Nganasan

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡ʂ					
Sib. fri.			ʂ					
N.-sib. fri.								

(139) Older (South) Korean (Ko 2013; Cho 2016)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.	s̪			ɕ				
N.-sib. fri.								

(140) Younger (South) Korean (Kim 1999, 2004)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪							
Sib. fri.	s̪			ɕ				
N.-sib. fri.								

(141) Pyongyang Korean

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡ʂ					
Sib. fri.			ʂ					
N.-sib. fri.								

(142) Kwakiutl

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡ʂ					
Sib. fri.			ʂ					
N.-sib. fri.						ɻ̊˖		

(143) Kala Lagaw Ya (Round et al. 2011: 113)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			ʂ ʐ					
N.-sib. fri.								

(144) Betsimisaraka Malagasy (O'Neill 2015)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠		t͡ʃ d͡ʒ			
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(145) Mono

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(146) Eton (Van de Velde 2008)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z						
N.-sib. fri.								

(147) Mbuko

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠ z̠					
N.-sib. fri.			ɬ ɮ					ɧ/x

(148) Dar Daju Daju (Aviles 2008; Ibrahim 2019: 46)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.								
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(149) Eys Limburgish (Kokkelmans 2020*b*)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɧ/x ɣ

(150) Estonian Swedish (Asu et al. 2015)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.	s̪			ɕ		ṣ		
N.-sib. fri.				ç				

(151) Early Romani (Matras 2002)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.

N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɧ/x

(152) Abun

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.								

(153) Amarasi

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̺					
N.-sib. fri.								

(154) Guernesiais

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.				ç				ɧ/x ɣ

(155) Argentinian Spanish

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.	θ ð							

(156) Assamese

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̺ z̺					
N.-sib. fri.								ɧ/x

(157) Asturian (Hermo del Teso 2017)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.			s̺		ʃ			
N.-sib. fri.		θ̠						

(158) Southern Azerbaidjani

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɧ/x ɣ

(159) Bearnais (Gascon)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.								
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(160) Cagliari Sardinian

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(161) Saterland Frisian

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.								ɧ/x ɣ

(162) Scots

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɧ/x

(163) Bzyp Abkhaz (Colarusso 1999; Chirikba 2003: 12; Hewitt 2013: 168)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠		t͡ʃ d͡ʒ		t͡s̢ d͡z̢	
Sib. fri.	s̪ z̪		s̠ z̠		ʃ ʒ		s̢ z̢	
N.-sib. fri.								

(164) Literary (Abzhywa) Abkhaz (Colarusso 1999; Hewitt 2013)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠		t͡ʃ d͡ʒ		t͡s̢ d͡z̢	
Sib. fri.	s̪ z̪				ʃ ʒ		s̢ z̢	
N.-sib. fri.								

(165) Younger Mauritian Creole (Baker and Kriegel 2013: 253; citealtDufour2014)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			

N.-sib. fri.								

(166) Beijing Mandarin Chinese (Duanmu 2000)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̺					t͡ʂ		
Sib. fri.	s̺			ɕ		ʂ		
N.-sib. fri.						ɻ̊˔		

(167) South Uist Scottish Gaelic

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.				ç ʝ				ɦ/x ɣ

(168) Middle Welsh (Liu 2018)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.	θ ð		ɬ					ɦ/x

(169) Quechua (Cusco) (Rataj 2005)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s						
N.-sib. fri.								

(170) Quechua (Ecuador) (Rataj 2005)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(171) Maastricht Limburgish (Gussenhoven and Aarts 1999)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				
Sib. aff.								
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɦ/x ɣ

(172) Petalcingo Tzeltal (Shklovsky 2005: 9)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(173) Amuzgo (Berns 2013: 65)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(174) Old Javanese (Teselkin and Echols 1972)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̠					
N.-sib. fri.								

(175) Antwerp dialect (De Schutter 1999)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(176) Aymara (Hardman 2001; Briggs 2007)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s						
N.-sib. fri.								ɧ/x

(177) Medieval Picard

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(178) Late medieval Eys Limburgish (Kokkelmans 2020*b*)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								ɧ/x

(179) Early medieval Eys Limburgish (Kokkelmans 2020*b*)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠					
Sib. fri.			s̠ z̠					
N.-sib. fri.								ɧ/x

(180) Swiss German (Berger 1913)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.

N.-sib. aff.								k͡x
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(181) Romansh

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(182) Ladino (Judaeo-Spanish)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɦ/x ɣ

(183) Kabyle Berber

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.				ç ʝ				

(184) Egyptian Arabic

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(185) Classical Arabic (Ferguson 1969)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.	s̪ z̪			ɕ				
N.-sib. fri.	θ ð		ɬ ɮ					

(186) Ainu

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ				
Sib. fri.	s̪			ɕ				
N.-sib. fri.								

(187) Pompeii graffitti Latin (Bateman 2007: 115f.)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠					
N.-sib. fri.								

(188) Tebul Ure (Heath 2015)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z						
N.-sib. fri.								

(189) Tigre

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(190) Middle Egyptian

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.				ç				

(191) Late medieval southern Spanish

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(192) Aragonese (Ledgeway and Maiden 2016)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.			s̠		ʃ			
N.-sib. fri.		θ̠						

(193) Central Alaskan Yupik

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s z						
N.-sib. fri.			ɬ					ɦ/x ɣ

(194) Tajik

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			

N.-sib. fri.								

(195) Indonesian

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(196) Esperanto

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(197) Central Venetian dialects (Belloni 2006: 25, Avesani et al. 2017)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̺ d͡z̺					
Sib. fri.			s̺ z̺					
N.-sib. fri.								

(198) Hone Aostian French dialect (Gruppo 'Amis du patois' 2007)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(199) St. Veran French dialect (Mathieu 2001)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(200) Samoan (Arakin 1973; Alderete and Bradshaw 2012)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̺					
N.-sib. fri.								

(201) Proto-Polynesian (Elbert 1964; Rutter 2001)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̺					
N.-sib. fri.								

(202) Younger Malagasy

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.					t͡ɹ̝̊ d͡ɹ̝			
Sib. aff.		t͡s d͡z						
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(203) Urban Eastern Norwegian (Solhaug 2010)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.				ç				

(204) Mundabli (Voll 2017)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪			t͡ɕ d͡ʑ				
Sib. fri.	s̪			ɕ				
N.-sib. fri.								

(205) Younger Low German

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s z			ʃ			
N.-sib. fri.				ç				ɦ/x ɣ

(206) Late Vulgar Latin (Pope 1934; Bateman 2007)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪			t͡ɕ d͡ʑ				
Sib. fri.			s̠					
N.-sib. fri.								

(207) Early Gallo-Romance

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪				t͡ʃ d͡ʒ			
Sib. fri.			s̠					
N.-sib. fri.								

(208) Middle Dutch

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̠ z̠					
N.-sib. fri.								ɦ/x ɣ

(209) Late West Germanic (van der Hoek 2010)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
Sib. aff.								
Sib. fri.			s̠					
N.-sib. fri.		θ̠ ð̠						ɦ/x ɣ

(210) Older Afrikaans (Wissing et al. 2015)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̠					
N.-sib. fri.								

(211) Middle Cornish (Adams 1975)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̠					
N.-sib. fri.		θ̠						

(212) Western Armenian

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(213) Proto Anglo-Frisian (van der Hoek 2010: 47f.)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.			s̠					
N.-sib. fri.		θ̠ ð̠						ɦ/x ɣ

(214) Old High German (Adams 1975)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								k͡x
Sib. aff.			t͡s̠					
Sib. fri.			s̠					
N.-sib. fri.		θ̠ ð̠						ɦ/x

(215) Late Proto-Finnic (Pajusalu 2012; Koivulehto 1986: 293)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.			s̠					
N.-sib. fri.								ɦ/x

(216) Early Proto-Finnic (Pajusalu 2012; Koivulehto 1986: 293)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.								

(217) Rotokas (Robinson 2006)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.								
N.-sib. fri.								

(218) Mocheno from Eichleit and Florutz (Rowley 2017)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								k͡x
Sib. aff.	t͡s̪				t͡ʃ d͡ʒ			
Sib. fri.	s̪ z̪		s̠ z̠		ʃ ʒ			
N.-sib. fri.								ɧ/x

(219) Wari (Everett and Kern 2002)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(220) Abipon (Najlis 1966; Lev et al. 2012)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				
Sib. aff.								
Sib. fri.								
N.-sib. fri.								ɧ/x ɣ

(221) Old Japanese (Pintér 2015)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̠					
N.-sib. fri.								

(222) Older Elfdalian (Sapir 2005; Garbacz 2008)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.			s̠					
N.-sib. fri.		θ̠ ð̠		ç ʝ				

(223) Younger Elfdalian (Sapir 2005; Garbacz 2008)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠					
N.-sib. fri.		θ̠ ð̠		ç ʝ				

(224) Northern Scanian dialects (Lucazin 2010)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.	s̪			ɕ				
N.-sib. fri.								

(225) White Hmong (Golston and Yang 2001; Jarkey 2015)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ		t̪͡ɭ̊d̪͡ɭ		
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ ʒ			
N.-sib. fri.			ɬ	ç ʝ				

(226) Jalapa Mazatec (Silverman et al. 2015: 83)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(227) Chalcotongo Mixtec (Macaulay 1996: 19, Botma and van 't Veer 2013: 47)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s			ʃ ʒ			
N.-sib. fri.	θ ð			ç				ɦ/x

(228) Great Andamanese (Abbi 2013)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.								ɦ/x

(229) Val Camonica dialects (Cresci 2014)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠ z̠					
N.-sib. fri.		θ̠ ð̠						

(230) Forest Nenets (Salminen 2007)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.	s̪			ɕ				
N.-sib. fri.			ɬ	ç				ɦ/x

(231) Green Hmong (Bruhn 2006)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				

Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s			ʃ ʒ			
N.-sib. fri.			ɬ	ç				

(232) Middle Spanish (Adams 1975)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.	s̪		s̠		ʃ			
N.-sib. fri.				ç ʝ				

(233) Proto-Greek

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.			s̠					
N.-sib. fri.								

(234) Late Proto-Greek

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠					
N.-sib. fri.								

(235) Kiribati (Gilbertese) (Blevins and Harrison 1999: 206)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.								
N.-sib. fri.								

(236) Cazet Fassanese Ladin (Chiochetti 2016)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ		t͡ṣ		
Sib. fri.	s̪ z̪					ṣ ẓ		
N.-sib. fri.								

(237) Margi (Hall 1997*b*: 18)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.			ɬ ɮ	ç ʝ				ɦ/x ɣ

(238) Zoque (Hall 2000: 713)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			

Sib. fri.		s			ʃ			
N.-sib. fri.								

(239) Amsterdam Hollandic Dutch (Collins and Mees 1999; Ooijevaar 2011)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				
Sib. aff.			t͡s̠					
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(240) Burushaski (Hunza and Nager) (Berger 1998: 13, 21f.)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ		t͡ṣ		
Sib. fri.	s̪ z̪			ɕ ʑ		ṣ ẓ		
N.-sib. fri.						ɻ̝̊		

(241) Burushaski (Jammu and Kashmir) (Munshi 2006: 58)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ d͡ʑ		t͡ṣ		
Sib. fri.	s̪ z̪			ɕ		ṣ ẓ		
N.-sib. fri.						ɻ̝̊		

(242) Wolof (Ngom 2006)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.								
Sib. fri.			s̠					
N.-sib. fri.								ɧ/x

(243) Lardil (Prince and Smolensky 1993/2004: 109)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ				
Sib. aff.								
Sib. fri.								
N.-sib. fri.						ɻ̝̊		

(244) Navajo (Oberly 2008: 40)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.			t͡ɬ					
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.			ɬ					ɧ/x ɣ

(245) Pomattertitsch Walser German (Dal Negro 2004: 77)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			

Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɧ/x

(246) Tohono O'odham (Dart 1991, 1993; Gafos 1999)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.	s̪		s̠					
N.-sib. fri.								

(247) Iskonawa (Zariquiey 2015: 56)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(248) Piro (Lin 1997: 404)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç				
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.				ç				

(249) Korubo (de Oliveira 2009)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.			ɬ					

(250) Epena (Harms 1984: 160)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s						
N.-sib. fri.								

(251) Sicilian (Privitera 1998; Rausch and Baiamonte 2017)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.				c/c͡ç ɟ/ɟ͡ʝ		t̠͡ɹ̊d̠͡ɹ		
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ			
N.-sib. fri.				ç				

(252) Georgian (Aronson 1990)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			

N.-sib. fri.								

(253) Kyrgyz (Abylkasymova and Jumabaeva 1997: 14)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								ɦ/x ɣ

(254) Seneca (Chafe 2015: 10)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(255) Slovenian (Šuštaršič et al. 1995)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								ɦ/x

(256) Kven (Söderholm 2017)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.	θ ð							

(257) Oklahoma Cherokee (Montgomery-Anderson 2008)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠					
Sib. fri.			s̠					
N.-sib. fri.			ɬ					

(258) Qualla Cherokee (King 1975)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠					
Sib. fri.			s̠					
N.-sib. fri.								

8.8.2 Every different kind of inventory attested

(1)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.								
N.-sib. fri.								

(2)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(3)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(4)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠					
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(5)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(6)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠		t͡ʃ d͡ʒ			
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(7)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪				t͡ʃ d͡ʒ			
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(8)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠					
Sib. fri.			s̠ z̠					
N.-sib. fri.								

(9)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠					
Sib. fri.			s̠ z̠		ʃ ʒ			
N.-sib. fri.								

(10)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠		t͡ʃ d͡ʒ			
Sib. fri.			s̠ z̠		ʃ ʒ			
N.-sib. fri.								

(11)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪				t͡ʃ			
Sib. fri.			s̠ z̠		ʃ ʒ			
N.-sib. fri.								

(12)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠		t͡ʃ			
Sib. fri.			s̠ z̠		ʃ			
N.-sib. fri.								

(13)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.			s̠					
N.-sib. fri.								

(14)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.			s̠					
N.-sib. fri.								

(15)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.

N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.			s̠					
N.-sib. fri.								

(16)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠					
Sib. fri.			s̠					
N.-sib. fri.								

(17)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.			s̠					
N.-sib. fri.								

(18)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪				t͡ʃ d͡ʒ			
Sib. fri.			s̠					
N.-sib. fri.								

(19)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪			t͡ɕ d͡ʑ				
Sib. fri.			s̠					
N.-sib. fri.								

(20)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.			s̠		ʃ			
N.-sib. fri.								

(21)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠		t͡ʃ			
Sib. fri.			s̠		ʃ			
N.-sib. fri.								

(22)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			

Sib. fri.		s z						
N.-sib. fri.								

(23)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z						
N.-sib. fri.								

(24)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(25)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(26)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(27)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(28)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(29)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z						
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(30)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(31)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ ʒ			
N.-sib. fri.								

(32)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(33)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(34)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(35)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(36)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(37)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s z			ʃ			
N.-sib. fri.								

(38)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s						
N.-sib. fri.								

(39)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.		s						
N.-sib. fri.								

(40)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s						
N.-sib. fri.								

(41)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s			ʃ ʒ			
N.-sib. fri.								

(42)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s			ʃ ʒ			
N.-sib. fri.								

(43)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ ʒ			
N.-sib. fri.								

(44)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								

Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ ʒ			
N.-sib. fri.								

(45)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.		s			ʃ			
N.-sib. fri.								

(46)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(47)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(48)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s						
Sib. fri.		s			ʃ			
N.-sib. fri.								

(49)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(50)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s			t͡ʃ d͡ʒ			
Sib. fri.		s			ʃ			
N.-sib. fri.								

(51)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.		t͡s d͡z			t͡ʃ d͡ʒ			

Sib. fri.		s			ʃ			
N.-sib. fri.								

(52)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ		t͡ṣ		
Sib. fri.	s̪ z̪					ṣ ẓ		
N.-sib. fri.								

(53)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ d͡ʑ		t͡ṣ d͡ẓ		
Sib. fri.	s̪ z̪					ṣ ẓ		
N.-sib. fri.								

(54)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠		t͡ʃ d͡ʒ		t͡ʂ̠ d͡ʐ̠	
Sib. fri.	s̪ z̪				ʃ ʒ		ʂ̠ ʐ̠	
N.-sib. fri.								

(55)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.	s̪ z̪			ɕ ʑ				
N.-sib. fri.								

(56)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪			t͡ɕ d͡ʑ				
Sib. fri.	s̪ z̪			ɕ ʑ				
N.-sib. fri.								

(57)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.	s̪ z̪			ɕ ʑ		ṣ ẓ		
N.-sib. fri.								

(58)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ		t͡ṣ		
Sib. fri.	s̪ z̪			ɕ ʑ		ṣ ẓ		

N.-sib. fri.

(59)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪			t͡ɕ d͡ʑ		t͡ṣ d͡ẓ		
Sib. fri.	s̪ z̪			ɕ ʑ		ṣ ẓ		
N.-sib. fri.								

(60)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.	s̪ z̪			ɕ				
N.-sib. fri.								

(61)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ				
Sib. fri.	s̪ z̪			ɕ		ṣ ẓ		
N.-sib. fri.								

(62)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ d͡ʑ		t͡ṣ		
Sib. fri.	s̪ z̪			ɕ		ṣ ẓ		
N.-sib. fri.								

(63)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠ d͡z̠					
Sib. fri.	s̪ z̪		s̠ z̠					
N.-sib. fri.								

(64)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠					
Sib. fri.	s̪ z̪		s̠ z̠					
N.-sib. fri.								

(65)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.	s̪ z̪		s̠ z̠		ʃ ʒ			
N.-sib. fri.								

(66)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.	s̪ z̪		s̠ z̠		ʃ ʒ			
N.-sib. fri.								

(67)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪				t͡ʃ d͡ʒ			
Sib. fri.	s̪ z̪		s̠ z̠		ʃ ʒ			
N.-sib. fri.								

(68)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠		t͡ʃ d͡ʒ			
Sib. fri.	s̪ z̪		s̠ z̠		ʃ ʒ			
N.-sib. fri.								

(69)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪		t͡s̠ d͡z̠		t͡ʃ d͡ʒ		t͡ʂ̠ d͡ʐ̠	
Sib. fri.	s̪ z̪		s̠ z̠		ʃ ʒ		ʂ̠ ʐ̠	
N.-sib. fri.								

(70)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ				
Sib. fri.	s̪			ɕ				
N.-sib. fri.								

(71)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.	s̪			ɕ				
N.-sib. fri.								

(72)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪							
Sib. fri.	s̪			ɕ				
N.-sib. fri.								

(73)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ				
Sib. fri.	s̪			ɕ				
N.-sib. fri.								

(74)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪ d͡z̪			t͡ɕ d͡ʑ				
Sib. fri.	s̪			ɕ				
N.-sib. fri.								

(75)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.								
Sib. fri.	s̪			ɕ		ṣ		
N.-sib. fri.								

(76)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.	s̪			ɕ		ṣ		
N.-sib. fri.								

(77)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪					t͡ṣ		
Sib. fri.	s̪			ɕ		ṣ		
N.-sib. fri.								

(78)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪			t͡ɕ		t͡ṣ		
Sib. fri.	s̪			ɕ		ṣ		
N.-sib. fri.								

(79)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪				t͡ʃ			
Sib. fri.	s̪		s̠ z̠		ʃ			
N.-sib. fri.								

(80)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.

N.-sib. aff.								
Sib. aff.				t͡ɕ d͡ʑ				
Sib. fri.	s̪		s̠					
N.-sib. fri.								

(81)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.			t͡s̠					
Sib. fri.	s̪		s̠					
N.-sib. fri.								

(82)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ d͡ʒ			
Sib. fri.	s̪		s̠		ʃ ʒ		ʂ̠ ʐ̠	
N.-sib. fri.								

(83)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.					t͡ʃ			
Sib. fri.	s̪		s̠		ʃ			
N.-sib. fri.								

(84)

	Dent.	Alve.	Retr.	Alpa.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪		t͡s̠		t͡ʃ			
Sib. fri.	s̪		s̠		ʃ			
N.-sib. fri.								

(85)

	Dent.	Alve.	Retr.	Alve.	Pala.	Plai.	Suba.	Vela.
N.-sib. aff.								
Sib. aff.	t͡s̪		t͡s̠		t͡ʃ			
Sib. fri.	s̪		s̠		ʃ			
N.-sib. fri.								