UNIVERSITÁ DEGLI STUDI DI VERONA

DEPARTMENT OF FOREIGN LANGUAGES AND LITERATURES

GRADUATE SCHOOL OF HUMANITIES

DOCTORAL PROGRAM IN FOREIGN LITERATURES, LANGUAGES AND LINGUISTICS

XXXIV Cycle / 2018

How children acquire adjectives: Evidence from three eye-tracking studies on Italian

---

IN *CO-TUTELLE DE THÉSE* WITH
UNIVERSITÄT KONSTANZ

S.S.D. L-LIN/01

Coordinator: Prof. Dr. Stefan RABANUS (University of Verona)

Tutor: Prof. Dr. Chiara MELLONI (University of Verona)

Tutor: Prof. Dr. Tanja KUPISCH (Universität Konstanz, UiT The Arctic University of Norway)

Doctoral Student: Michela REDOLFI

How children acquire adjectives: Evidence from three eye-tracking studies on Italian – Michela Redolfi
Doctoral Thesis
Verona, 2022

Table of Contents

Acknowledgements

Throughout the writing of this dissertation I have received a great deal of support and assistance.

First and foremost, I would like to thank my supervisor Chiara Melloni for sparking my interest in psycholinguistics and language acquisition and for her support and feedback before and during my Doctoral program. Your expertise was invaluable in figuring out my objectives and in guiding me through the writing process of this thesis. I am also very grateful to my second supervisor, Tanja Kupisch, for believing in this project from our very first meeting, and for encouraging the cooperation with the Linguistics department at the University of Konstanz. Thank you for the availability and for offering me the chance to meet inspiring researchers and colleagues with whom I have enjoyed working. This thesis benefitted greatly from both your insightful comments and suggestions.

I am profoundly thankful to Bettina Braun for her interesting and informative classes in statistics and for the huge help she gave me with data analysis. You provided me with the tools that I needed and made the most challenging part of this work the most stimulating.

I wish to extend my special thanks to Denis Delfitto, Maria Vender, Marta Tagliani and all the colleagues of the LaTeC research group for their precious advices and for the helpful discussions about the empirical and the theoretical aspects of my project.

I would like to thank the child and adult participants for taking part in the experiments and for trusting our research group during these difficult times. I am also grateful to all my friends and colleagues who volunteered in the pilot versions of the experiments.

I am much obliged to my colleagues in Verona for their moral support, the stimulating conversations, and for providing happy distractions to rest my mind outside of my research. Also, I am grateful to all the people I met in Konstanz for sharing their knowledge with me and for their hospitality during the months I spent with them.

Most of all I thank my family and friends. This accomplishment would not have been possible without you.

Abstract

While the literature on early language acquisition has mainly focused on nouns and verbs, studies on adjectives are comparatively scarce and have thus far provided contrasting evidence on the timing and mode of acquisition of adjectival meaning in children younger than 4. This thesis presents three eye-tracking studies exploring the online processing of adjectives by children (2;4 – 5;3 years) in comparison to adult controls, providing insights on how linguistic information and visual context interact during real-time comprehension.

Experiment 1 investigated potential differences in the interpretation of three classes of adjectives, intersective (e.g., *red*), relative (e.g., *big*) and absolute (e.g., *full*). 38 Italian monolingual children (2;4 – 5;3) were tested in a visual-world task, where they listened to noun-adjective combinations in a four-picture scenario. Results showed that children as young as 28 months were slower than adults when interpreting noun-adjective combinations, while their looking pattern in the interpretation process was essentially the same. Furthermore, the computation of intersective adjectives was faster than that of absolute and, especially, relative adjectives, showing that children are sensitive to the different ways in which each adjective class is interpreted within different contexts.

In Experiment 2 the complexity of Experiment 1 was reduced by lowering the processing load associated with both the different semantic classes of adjectives and the four-picture scenario. 28 Italian-native children (2;4 – 5;2) were presented with a two-picture display while listening to nouns combined with color-adjectives. In Experiment 2, the visual conditions varied according to the informativeness of the noun or the adjective with respect to the target referent. Results showed evidence that adjective processing develops over time. When the computation of the noun was insufficient and the integration with the adjective was necessary to resolve reference, the youngest children, unlike 3- and 4-year-olds, failed to interpret adjectives correctly and, consequently, task resolution.

Experiment 3 investigated children's incremental processing of prenominal adjectives and their ability to predict the following noun based on the lexical meaning of the adjective. 39 Italian children (2;4 – 5;3) were tested in the online

interpretation of Italian predicative yes/no questions (e.g., *È morbido il cuscino?*, lit. 'Is soft the pillow?') while looking at two pictures on the screen. Here, the informativeness of the adjective was manipulated. Results showed that, when informative (e.g., *soft*, upon looking at a pillow and a bone), the adjective was processed incrementally, i.e., before the noun was heard, indicating that children as young as 28 months are able to predict the upcoming noun based on adjective meaning. Furthermore, children were successful independently of whether the interpretation of the adjective required world knowledge (e.g., being *soft* for a pillow) or the exploration of the visual scene (e.g., being *open* for a window).

Taken together, the three studies provided compelling evidence of a continuous process in children's development of sophisticated, adult-like processing skills. By means of eye-tracking, we were able to reveal that the overall difference between children and adults is mostly attributable to toddlers younger that 36 months of age, whose processing skills are still limited when it comes to the meaning computation of noun-adjective combinations. However, from the age of 3, children's processing abilities improve rapidly and, within a few months, they become successful parsers.

Abstrakt

Während sich die Literatur zum frühen Spracherwerb vornehmlich auf Nomen und Verben konzentriert hat, gibt es vergleichsweise wenige Studien zu Adjektiven, die ferner gegensätzliche Erkenntnisse über den Zeitpunkt und die Art und Weise des Adjektiverwerbs bei Kindern unter 4 Jahren geliefert haben. Die vorliegende Arbeit stellt drei Eye-Tracking-Studien vor, die die Online-Verarbeitung von Adjektiven durch Kinder (2;4 – 5;3 Jahre) im Vergleich zu Erwachsenen untersuchen.

Experiment 1 erforschte mögliche Unterschiede in der Interpretation dreier Klassen von Adjektiven: intersektive (z. B. *rot*), relative (z. B. *groß*) und absolute (z. B. *voll*). 38 einsprachig italienische Kinder (2;4 – 5;3) wurden mit dem *Visual-World* Paradigma mit vier Bildern getestet, bei der sie sich Nomen-Adjektiv-Kombinationen anhörten. Die Ergebnisse zeigten, dass Kinder im Alter von 28 Monaten bei der Interpretation von Nomen-Adjektiv-Kombinationen langsamer waren als Erwachsene, während ihr Blickbewegungsmuster weitgehend identisch waren. Darüber hinaus war die Berechnung intersektiver Adjektive schneller als die absoluter und insbesondere relativer Adjektive, woraus hervorgeht, dass Kinder für die Interpretation der unterschiedlichen Adjektivklassen in unterschiedlichen Kontexten sensitiv sind.

In Experiment 2 wurde die Komplexität von Experiment 1 verringert, indem der Verarbeitungsaufwand für die semantisch verschiedenen Adjektivklassen verringert wurde. 28 italienische Kinder (2;4 – 5;2) wurden mit zwei Bildern konfrontiert, während sie Substantive mit Farbadjektiven hörten. In Experiment 2 variierten die visuellen Konditionen je nach der Aussagekraft des Nomens oder des Adjektivs in Bezug auf den Zielreferenten. Die Ergebnisse zeigten, dass sich die Adjektivverarbeitung mit zunehmendem Alter entwickelt. Wenn die Berechnung des Substantivs unzureichend und die Integration mit dem Adjektiv notwendig war, um den Referenten zu identifizieren, konnten die jüngsten Kinder im Gegensatz zu den 3- und 4-jährigen die Adjektive nicht richtig interpretieren und damit die Aufgaben nicht lösen.

Experiment 3 untersuchte die inkrementelle Verarbeitung pränominaler Adjektive bei Kindern sowie ihre Fähigkeit, das folgende Nomen auf der Basis der lexikalischen Adjektivbedeutung vorherzusagen. 39 italienische Kinder (2;4 – 5;3) wurden in der Online-Interpretation italienischer prädikativer Ja/Nein-Fragen getestet (z.B. *È morbido il cuscino?*, lit. «Ist das Kissen weich?»), während sie zwei Bilder auf dem Bildschirm betrachteten. Hierbei wurde die Aussagekraft des Adjektivs manipuliert. Das Experiment zeigte, dass die Kinder, wenn das Adjektiv informativ war – wie etwa *weich* beim Betrachten eines Kissens bzw. Knochens –, die im Adjektiv enthaltene Information verarbeiteten, bevor sie das Substantiv hörten. Das deutet darauf hin, dass Kinder ab dem Alter von 28 Monaten in der Lage sind, das nachfolgende Nomen anhand der Adjektivbedeutung vorauszusagen. Ferner waren Kinder erfolgreich, unabhängig davon, ob die Interpretation des Adjektivs Weltwissen (z.B. *weich* für ein Kissen) oder die Erkundung der visuellen Szene (z. B. *offen* für ein Fenster) erforderte.

Zusammenfassend liefern die drei Studien Evidenz für die kontinuierliche Entwicklung ausgefeilter, erwachsenartiger Sprachverarbeitungsfähigkeiten bei Kindern. Mit der Eye-Tracking Methode konnten zeigen, dass sich Unterschiede zwischen Kindern und Erwachsenen vor allem auf Kinder im Alter vor 36 Monaten beschränken, deren Verarbeitungsfähigkeiten bei der Bedeutungsberechnung von Nomen-Adjektiv-Kombinationen noch begrenzt sind. Ab dem Alter von 3 Jahren verbessert sich die Verarbeitungsfähigkeit der Kinder jedoch drastisch: Innerhalb von wenigen Monaten werden sie zu erfolgreichen Parsern.

General overview

The present work aims to explore the early acquisition of adjective meaning and to contribute to the literature on processing and integration of nouns and adjectives by children between two and five years of age. Three Visual World eye-tracking studies were carried out, aiming to provide insights on how linguistic and visual information interact during real-time comprehension. In doing so, we examine adjectival acquisition from multiple perspectives. First, we investigate potential differences in the online interpretation of different semantic classes of adjectives. Specifically, we focus on subsective adjectives, and we are interested in comparing the pattern of interpretation of intersective, relative and absolute adjectives. Second, we inquire into the interpretation of adjectives in conditions in which the integration of noun and adjective meanings is either required or unnecessary for reference resolution. Third, we investigate the role of the informativeness of the adjective with respect to the referent noun and the ability to predict a noun based on the lexical meaning of the adjective. To determine a baseline for children's performance, a group of adults has been tested in each study.

This dissertation is organized as follows. In chapter 1 we provide a general overview of the theoretical background concerning the word class of adjectives. We discuss some of the most relevant issues concerning the status of adjectives as a universal word class, as well as the syntactic, semantic, and morphological properties characterizing it cross-linguistically. This chapter serves as an introduction to the theoretical literature on adjectives as a part of speech, with a focus on English – being the most studied language –, and Italian – the language under investigation in the experimental part of this work.

Chapter 2 zooms in the acquisitional aspects of adjective comprehension and production. We outline the main theoretical approaches accounting for the difficulty young learners face in acquiring adjectives in comparison to nouns and we report on the experimental studies exploring adjective comprehension that are relevant for the present work. Particularly, we focus on the main studies exploring the integration of noun and adjective meanings, the influence of the relative order of nouns and adjectives in real-time processing, and the interpretation of different semantic classes of adjectives.

In chapter 3 we introduce the Visual World paradigm, an experimental methodology exploiting eye-movements during listening to investigate language processing online. Since this technique constitutes the methodology at the basis of the present dissertation, we overview the main features of Visual World studies conducted with children and adults. We report on the seminal studies introducing this technique to study language processing and the technological innovations that improved it over time. First, we present the Preferential-Looking paradigm, overviewing its revolutionary role in the understanding of early language processing, and discussing its strengths and weaknesses. Second, we illustrate the Looking-While-Listening task and the qualitative adjustments it made to data collection and data analysis. Finally, we present the eye-tracking technology – whose advent revealed the noteworthy limitations of older procedures – and we highlight the revolution brought about the high-precision measurements of this research system.

Chapter 4 is dedicated to the first Visual World experiment we implemented for this dissertation. This study tested children and adults in the interpretation of Italian noun-adjective combinations in a four-referent visual scenario. Adjectives were either intersective, relative or absolute and were uttered post-nominally. After briefly summarizing the theoretical background at its basis, we present the research questions we aim to answer, the experimental design, the materials and methods, as well as the predicted outcomes. We then discuss the results of the statistical analysis we conducted on eye-movement data, and we interpret our findings within the theoretical and experimental literature. Previous studies testing children younger than age 3 found evidence of a *Noun-Anchoring* mechanism in the interpretation of nouns and adjectives in combination, i.e., that children focused on the noun and delay or omit adjective processing. This was not the case in our experiment. We found evidence that children as young as 28 months showed, at a slower pace, the same process of interpretation of older children and adults. Moreover, our data evidenced that children are sensitive to the semantic properties of subsective adjectives and know to what extent the relationship with the context influences adjective meaning.

In chapter 5 we present a second Visual World study, designed to reduce the complexity of the previous experiment by lowering the processing load associated with both the different semantic classes of adjectives and the four-picture visual-context. We displayed two objects on the screen, and we recorded children and adults' eye-movements while listening to nouns combined with color-adjectives. In this experiment, the visual conditions were manipulated and varied according to the informativeness of nouns and adjectives with respect to the referent object. First, we offer a brief introduction on previous studies exploiting a similar experimental task using the Looking-While-Listening paradigm, and we present the research questions. We outline the methods and materials, and we discuss the results emphasizing the new insights emerging thanks to the use of the advanced eye-tracking technology. We will see that testing children in a wide age-range allowed for a developmental pattern to emerge. We evidenced that, unlike older children and adults, 2-year-olds failed target identification when processing the adjective was necessary to resolve reference, but successfully resolved reference when the noun was informative.

Chapter 6 reports the third Visual World experiment we designed to investigate further aspects of adjective processing, namely the interpretation of adjective-noun occurrences and the role of the prenominal adjective in predicting the following noun. We first review the previous literature on incremental processing and predictive understanding, focusing on the few studies addressing this issue with prenominal adjectives. We put emphasis on the fact that the investigated languages in these studies have mostly been Germanic, exploiting the adjective-noun order of attribution characterizing, e.g., English or Dutch. Since the Covid-19 outbreak prevented the afore-planned data collection with German-learning children, we adapted the experimental protocol to Italian, and we asked predicative yes/no questions which present the adjective-noun order. The chapter is organized along the lines of chapters 4 and 5. We outline the research questions and predictions, and we describe methods and materials. After presenting the results, we conclude the chapter with the discussion of our main findings. We will see that, when the adjective is informative about the following noun, children successfully predict the noun based on the lexical meaning of the adjective and we underline the importance of this finding compared to previous literature in this respect.

In chapter 7, we outline our final remarks and conclusions. We summarize our main findings, we interpret the outcomes of the three experiments taken together, and we discuss how this dissertation contributes to the existing literature in the fields of language acquisition and psycholinguistic processing. We conclude by discussing a few limitations and future directions.

PART I

1 Adjectives as a universal word class

Human language is a universal and recognisable human faculty whose purpose is to communicate meaning. The basic concepts of a language are encoded in words (the lexicon), which are combined thanks to grammar (or syntax, from an Ancient Greek word meaning “arrange together”). The categorisation of lexical words into parts of speech has been a central issue in the study of the language since ancient times, with a continuous tradition which probably started with *Téchnē grammatikē* of Dionysius Thrax, around 100 BC (Robins, 1989). He realised that some words inflect for case (i.e., nouns) and other for tense and person (i.e., verbs). This distinction at the morphological level was also linked to the observation, at the semantic level, that nouns referred to concrete and abstract entities, that verbs signified activities and processes performed or undergone, and that adjectives expressed properties of things (Baker, 2003). Nowadays, when entering their first linguistics class, one of the few things students know is that some words are nouns, some are verbs and some others are adjectives, but little is known about what it means to be a noun, a verb, or an adjective, i.e., about the definition of lexical classes.

1.1 The definition of lexical classes

The definition of lexical classes, or parts of speech, is a key element in research on lexical semantics, syntactic theory, and morphological analysis, as well as in the grammatical and typological description of languages. Nonetheless, lexical classes are one of the least defined and least understood concepts in linguistics. According to Beck (2002), to be truly useful and appropriate tools for linguistic enquiry, definitions of parts of speech must perform two tasks: they must accurately spell out what it means to belong to a word class and predict the properties that all members have. Those definitions must be universal in scope and must be able to cope with the typological variation across languages. In this section, we will provide an overview of the history behind the identification of adjectives as a universal word class, showing that cross-linguistic comparisons gave rise to an active debate on the topic.

1.1.1 Through the history of adjectives

Among the three major open lexical classes (i.e., nouns, verbs and adjectives), the identification of nouns and verbs has hardly been a controversial issue in the individuation of lexical categories, and they have always been considered universal classes. By contrast, the adjective class has generated debates and controversies. Cross-linguistically, the distinction between nouns and verbs is generally considered robust and relatively easy to establish, according to basic grammatical functions. Nouns are prototypically heads of noun phrases functioning as core arguments and appearing in copula complements, while verbs prototypically function as heads of predicates which bear tense, aspect, and mood marking (Hajek, 2004). Adjectives, on the other hand, have always been more difficult to distinguish across languages (e.g., Dixon, 1982a; Bhat, 1994; among others).

In the western tradition of linguistics, which goes back at least to the grammar of Sanskrit, Greek and Latin, no distinction between nouns and adjectives was theorised. Within the European grammatical tradition, adjectives have been subsumed for centuries within the category of nouns. The first partial separation of nouns and adjectives is relatively recent. Only around the fourteenth century, Thomas of Erfurt established a nominal sub-class of adjectives (*nominem adiectivum*) alongside normal nouns (*nominem substantivum*), introducing the criterion of gender: nouns have inherent gender, whereas adjectives do not, as they agree with the noun they modify. This doctrine has been accepted among linguists since it seemed to work for the European languages they knew. However, when the study of language started taking into consideration other languages like Finnish, which has no genders, this criterion started showing its flaws as it could not work universally. It was only in the eighteenth century that the first reference to adjectives as an independent class occurred for French (Girard, 1747), and later for Italian (Vanzòn, 1834), with full acceptance only in the twentieth century (Scarano, 1999).

The study of more exotic languages brought several linguists to the conclusion that no basic grammatical functions are associated with adjectives only, and the grammatical differences that distinguish adjectives from the other word classes are more subtle and diffuse (Hajek, 2004). For this reason, many linguists are disinclined to recognize an independent adjective class in the languages they describe and, consequently, it has been argued that having an adjective class is not a universal property of natural languages (see e.g., Déchaine, 1993). More recent research on adjectives (e.g., Baker, 2003; Dixon & Aikenvald, 2004; Dixon, 2004), however, has defended the idea that in every human language it is possible to distinguish three word-classes: nouns, verbs and adjectives. The distinction between those classes must be aligned with the particular grammar of every single language, but each class can be identified in every language, since it has a prototypical conceptual basis and prototypical grammatical functions (Dixon & Aikenvald, 2004). In the next section, we turn to the difficulties linguists faced in trying to identify the adjective class due to substantial cross-linguistic variation concerning its size and its grammatical as well as functional properties.

1.2 Difficulties in the identification of adjectives

In most languages, distinguishing nouns from verbs is easy in terms of syntactic function and morphological properties. This is not the case for adjectives. The reasons for the failure to recognize the class of adjectives as a universal part of speech can be many. Theoretical models were built on what the American linguist Benjamin Whorf referred to as the "Standard Average European" language type, with its three-way division of the lexicon into the three open classes of verb, noun, and adjective. However, as soon as modern linguistics has expanded its horizons and started dealing with previously undescribed languages, many of the properties and definitions of parts of speech have been challenged. This hampers any possible generalisation, especially in languages in which adjectives show syntactic functions that go beyond being the modifier of noun phrases or being copula complements (Dixon & Aikenvald, 2004). In some languages, adjectives have grammatical properties similar to nouns, in some to verbs, in some to both, and in some to neither (Dixon, 2004).

1.2.1 Cross-linguistic variation of adjectives

As mentioned above, according to Dixon (2004), the cross-linguistic variation adjectives show can concern (i) size, (ii) functional and/or (iii) grammatical properties. We will now discuss these three aspects in more details.

First, the adjective class can significantly vary in size across different languages. While nouns and verbs are open classes, adjectives can either be a small/closed class as well as an open word-class. Indeed, the smallest classes have just three or four members, typically belonging to the same semantic types (see §1.4.1). Regardless of the size of the class, however, most of the members are generally derived from nouns and verbs via derivation (see Givón 1970).

Second, while nouns relate to arguments and verbs to predicates, functional expectations for adjectives are more complex and more varied. Two are the main roles they can assume (Dixon, 2004). First, adjectives can state that something has a certain property, as shown in (1):

(1) The man is *tall*.

An extension of this function is found in those languages in which adjectives have a comparative structure, as in the example in (2):

(2) The man is *taller* than the woman.

The second role adjectives play is that they can be the specification of the referent noun in a noun phrase that relates to a predicate argument, as in (3):

(3) The *tall* man.

In most languages all adjectives have both functions, while in fewer languages they can be limited to one of them (see Bolinger 1967 for English).

Lastly, as far as grammatical properties are concerned, adjectives can borrow the morphological processes available to verbs (concerning tense, aspect, mood, etc.) or to nouns (number, case, etc.). However, in some languages they do not share morphology with the other two classes (Dixon, 2004).

This wide variation explains why many linguists are reluctant in the use of the term "adjectives" and prefer "descriptive verbs" (for example Seki, 1990, on Kamaiurá, Tupí-Guaraní branch of Tupí family) or "stative verbs" (for example Buse, 1965, for this category in Oceanic languages)[1].

1.2.2 The need for identification criteria

What has been said so far indicates that, from a typological perspective, it is crucial to have criteria for recognising adjectives across languages and to distinguish them from other word classes, as well as for identifying different types of adjectives. As already mentioned, Dixon (2004) identified two roles that adjectives can play in the grammar of a language, i.e., state that something has a property (see example (1)) and being the specification of the referent noun (see example (3)). Specifically, the first role can be syntactically expressed in two different ways:

- As "verb-like adjectives". In this case, adjectives function as intransitive predicate, that is, they make use of the syntax and the morphology of verbs. An example can be found in Fijian (Austronesian language spoken in Fiji)[2]:

(4) [e balavu]$_{\text{INTRANSITIVE PREDICATE}}$ [a tama - qu]$_{\text{S}}$
3sgS tall ARTICLE father - 1sg.possessor
'My father is tall' (or something like **My father talls*)

- As "non-verb-like adjectives". In this case, adjectives function as copula complements, as in the English example:

(5) [my father]$_{\text{CS}}$ [is]$_{\text{COPULA PREDICATE}}$ [tall]$_{\text{CC}}$

[1] Since the present work is focused on the Italian language, where the presence of the class of adjectives is well established, we will refer to this word-class with the term "adjective".

[2] In the following examples: S = subject, CS = copula subject, CC = copula complement.

‘My father is tall’

The second role (i.e., provide a specification), allows adjectives to be categorised into two classes, according to how they modify the noun within a noun phrase:

- As “noun-like adjectives”. When it functions within a noun-phrase, an adjective can take some or all syntactic and morphological processes that generally apply to nouns.

- As “non-noun-like adjectives”. In some languages, nouns show a number of morphological processes that do not apply to adjectives.

Finally, in isolating languages, there may be no morphological process associated to nouns, so that the distinction between noun-like vs. non-noun-like adjectives is not relevant.

As Dixon (2004) pointed out, the adjective class does not fulfil both functions, predication, and modification, in all languages; there are some in which only one or the other is fulfilled. Thanks to the combination of those parameters, we can find (see Dixon, 2004, for a full discussion):

1. A large number of languages with verb-like and non-noun-like adjectives
2. A large number of languages with non-verb-like and noun-like adjectives
3. Some languages with verb-like and noun-like adjectives
4. Some languages with non-verb-like and non-noun-like adjectives

Detailed studies on adjectives carried out by Baker (2003) and Dixon (2004) have given evidence for the existence of a lexical category distinct from nouns and verbs in languages that had been analysed as lacking an adjective class. The criteria invoked by Baker and Dixon to identify a class of adjectives include the following:

(6)

a. Adjectives allow direct modification of nouns. (Baker 2003: 252–256, Dixon 2004: 19–20)

b. Adjectives differ from other predicates in the comparative construction. (Dixon 2004: 11,21)
c. Adjectives do not have their own gender, but they agree with the gender of the modified noun (Baker 2003: 247, Dixon 2004: 23)
d. Adjectives can appear without a preposition in resultative predications. (Baker 2003: 219–30)

The proposed criteria, however, do not set apart adjectives from nouns and verbs in all languages, since independent cross-linguistic differences can interfere with the criteria (Baker, 2003; Dixon, 2004). For instance, criterion (6a) is not applicable in languages like Slave (Athapaskan), which does not allow direct modification of the noun (Baker, 2003 citing Rice 1989). Criterion (6b) also needs to be analysed in detail for each language. Comparative constructions may not distinguish adjectives from nouns. For example, not all degree words select adjectives exclusively. Both adjectives and nouns admit comparative construction in Portuguese, Sanskrit, and Dyirbal, whereas adjectives only, but not nouns, admit comparatives in Russian, Finnish, and Hungarian (Baker, 2003). For instance, in English, *how, too, so* and *as* are limited to adjectives, while semantically similar expressions such as *more, less* and *enough* can also combine with other expressions such as mass nouns (e.g. *more/less/enough water*) and verbs (e.g., *I trust her more/less/enough*) (Cabredo Hoffher, 2010). A final example concerns criterion (6d): many languages (like French, Hindi, Hebrew, and Chichewa) do not allow adjectives to form resultative secondary predicates and may only have resultatives with prepositional phrases (Baker, 2003).

All these examples make clear that the criteria in (6) are flawed in many ways, since they do not uniformly and universally isolate the characteristic features of adjectives and, as a consequence, other properties of the language can interfere. However, the criteria provide a useful battery of tests to set apart adjectives from other word classes in a given language (Cabredo Hoffher, 2010). Although different attempts have been made, no universal criteria have yet been found to identify adjectives uniformly and universally.

In the following section, we will analyse adjectives from the syntactic point of view, focusing on English (the most studied language) and Italian (being relevant for the present work).

1.3 Syntax of adjectives

The syntactic distribution of parts of speech is often more closely related to lexical class membership than to any semantic or morphological property (Beck, 2002). The most elementary syntactic approaches define each lexical category on the basis of the syntactic function they take in the sentence: nouns can be subjects, verbs can be syntactic predicates, and adjectives can be attributive modifiers (Chomsky, 1965).

In §1.3.1, we will describe in detail a well-known syntactic dichotomy concerning adjectival functions, i.e., attributive vs. predicative adjectives. In §1.3.2, we will report on pre- and post-nominal adjectives across languages, focusing on Italian. §1.3.3 we will illustrate the underlying syntactic structures of both attributive and predicative adjectives. §1.3.4 will discuss arguments in favour and against the hypothesis that attributive adjectives are derived from predicative constructions via a relative clause.

1.3.1 Attributive vs. predicative adjectives

Even though the formal realisation of adjectives is not homogeneous across languages, many authors associate the definition of this word class with two main functions (e.g., Dixon, 2004): predicative and attributive. It might be possible that the substantial morphosyntactic variation adjectives show cross-linguistically is linked to the double role they play, i.e., adjectives either describe a property of an entity (as in 7a) or they specify the set of possible referents (as in 7b):

(7) a. The beautiful woman (attributive use)
b. The woman is beautiful (predicative use)

In the examples (7a) and (7b), the same adjective performs two different functions. (7a) is an example of attributive use, that is, the adjective is the modifier of the noun within a noun phrase. The example in (7b) shows the predicative use of adjectives; in this case, the adjective is the predicate of the sentence whose subject is the noun phrase [the woman]. In English, the copula is required for the predicative function; however, it is not overtly realised in all languages when a predicative adjective is used (as, e.g., in Russian) (Ramaglia, 2008).

It is worth mentioning that the term *predicative* adjective is often used in the literature to refer to structures different from the one in (7b), as the dichotomy *attributive* vs. *predicative* is not consistently used to indicate the contrast in (7). The two structures are sometimes referred to as *adnominal* vs. *predicative*, which correspond respectively to adjectives that are internal to the noun phrase, and adjectives that are heads of an independent phrase with a predicative function. By contrast, the dichotomy *attributive* vs. *predicative* is seldom used to indicate two different kinds of adjectives, both with the function of modifiers of the noun (e.g., Kamp, 1975: 124). This opposition can be explained with the examples below:

(8) a. The former president
    b. An interesting problem

In both the examples in (8), the adjective modifies the nominal head, but there is a slight difference in the interpretation: only the adjective in (8b) can be used as a predicate in a sentence with a copula, whereas the adjective in (8a) cannot, as shown in (9) below:

(9) a. *The president is former
    b. The problem is interesting

Thus, the use of the terms *attributive* and *predicative* makes clear that the first is used to refer to adjectives that cannot occur as predicates, while the second term refers to those adjectives that can be both modifiers of the noun phrase and predicates (Ramaglia, 2008).

To avoid possible confusion, in the present work the terms *attributive* and *predicative* will be used exclusively to refer to the contrast shown in the example

(7). The opposition in the example (8) will be indicated by the terms *functional* vs. *lexical* (following Bernstein, 1992).

Some authors proposed to divide adjectives into two subgroups:

1. adjectives that can be predicative or attributive
2. adjectives that can only be attributive

These two subgroups show different syntactic behaviour. An approach of this kind has been proposed by Bernstein (1992). In her comparative analysis of adjectives in the Romance languages, she uses the term *lexical* adjectives to refer to the adjectives that can only occur in the attributive position and *functional* adjectives to refer to the ones that can fulfil both functions. In her analysis, as shown in (10) for Italian, Bernstein shows that these two groups are different not only in their syntactic function but also because only *lexical* adjectives can exclusively occur in prenominal position (10a), cannot occur in elliptical noun phrases (i.e., without head-noun) (10b) and cannot be modified (10c).

(10) a. Un mero incidente vs. *Un incidente mero
'A mere accident'
b. *Quello mero
'The mere one'
c. *Un così mero incidente
'Such a mere accident'

All these possibilities are allowed for *functional* adjectives: they can occur pre-nominally or post-nominally (11a), can occur in elliptical noun phrases (11b), can be modified (11c) and (11d) and coordinated (11e) and (11f).

(11) a. Una persona simpatica / Una simpatica persona
'A nice person'
b. Quella simpatica
'That nice (person)'

c. Una così simpatica persona / Una persona così simpatica
'Such a nice person'
d. Il mio [più vecchio] collaboratore
'My oldest collaborator'
e. Voglio sapere la [pura e semplice] verità
'I want to know the pure and simple truth'
f. Questo è un imbroglio [puro e semplice].
'This is a pure and simple scam'

Although the terminology is not consistent in the literature, the distinction which is mainly relevant for the present work is the predicative vs. attributive use of adjectives, as shown in (7). In addition, we will benefit from the functional vs. lexical dichotomy in (8) to account for the syntactic distribution of Italian adjectives discussed in §1.3.2.

1.3.2 Pre- and post-nominal position of adjectives

In languages such as Italian, adjectives appear mostly postnominally or rather can occur both in pre- and postnominal position inside the DP. Theories disagree on whether one position is derived from the other or whether they have different underlying syntactic structures.

Approaches that argue for different syntactic structures motivate the variation in adjective positions by claiming that each position is related to a different interpretation (e.g., Cinque, 2010). In Italian, as in some other Romance languages, lexical adjectives (e.g., adjectives of colour, shape, material) are post-nominal, while functional adjectives (e.g., possessives, ordinals) are pre-nominal, except for relational adjectives (see §1.4.5).

There are cases in which adjectives can be either pre- or post-nominal, associated with a different meaning (Alexiadou, 2004). A frequently cited example in Italian is the one in (12) (e.g., Alexiadou, 2004; Cardinaletti & Giusti, 2010):

(12) a. Un povero ragazzo ('pitiable')
b. Un ragazzo povero ('impoverished')

The interpretation of the adjective *povero* is ambiguous: it can mean both pitiable (12a) and poor/impoverished (12b). The former meaning, moreover, is banned from the predicative uses. Only the post-nominal interpretation (impoverished) is expressed by the predicative use, as shown in (13) (Alexiadou, 2004):

(13) Il ragazzo è povero ('impoverished'/*'pitiable')

Besides examples such as the one in (13), there are cases in which the same adjective in pre- or post-nominal position does not create differences in meaning, as in examples (14) and (15). It is mainly a matter of pragmatic interpretation (Ramaglia, 2008).

(14) a. Un enorme elefante
b. Un elefante enorme
'A big elephant'

(15) a. Un bravo chirurgo
b. Un chirurgo bravo
'A good surgeon'

Some of the adjectives that appear both pre- and postnominally have only one possible reading in the prenominal position, as in (16), but two possible reading in postnominal position, as in (17) (Kupisch, 2014).

(16) una mia grande amica (non-ambiguous)
a my big friend
'a close friend of mine'

(17) una mia amica grande (ambiguous)
i. 'a close friend of mine'

ii. 'a friend of mine who is old(er than I)/tall/big'

While the direct modification reading is possible with both orders in (16) e (17i), the indirect modification reading in (17ii) is unique to postnominal adjectives (Cardinaletti & Giusti, 2010; Cinque, 2010).

In the following section, we will discuss how attributive and predicative adjectives have been syntactically analysed in the literature and what the implications for a cross-linguistic comparison are.

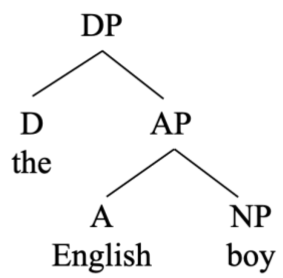
1.3.3 Syntactic analysis of adjectives

1.3.3.1 The syntax of attributive adjectives

Although there is no consensus in the literature as to the analysis of the syntax of attributive adjectives cross-linguistically, two are the main approaches that can be found in the literature. Indeed, adjectives are analysed as either heads or specifiers projections (see Alexiadou et al., 2007, for an overview of the advantages and disadvantages of both analyses).

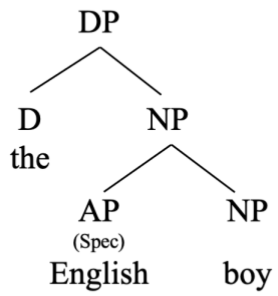
The first type of approach treats adjectives as heads that are selected by the determiner and take the noun phrase as a complement (Abney, 1987):

(18) $[_{DP}$ D $[_{AP}$ A $[_{NP}$ N]]]

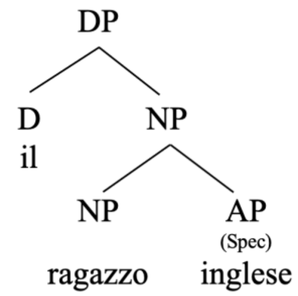


In the second type of analysis, adjectives are phrases that are adjoined to the noun phrase (as in (19), Jackendoff, 1977), or specifiers of dedicated functional projections in the extended projection of the noun (as in (20), Cinque, 1994):

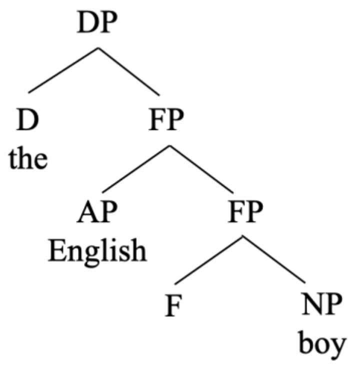
(19) a. $[_{DP}$ D $[_{NP}$ AP NP]] (left-adjoined AP)



b. $[_{DP}$ D $[_{NP}$ NP AP]] (right-adjoined AP)



(20) $[_{DP}$ D $[_{FP}$ AP F $[_{FP}$ AP F $[_{NP}$ N]]]] (AP in spec FP)

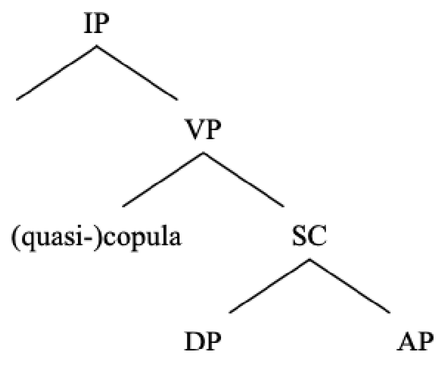


It is worth noticing that, while the syntactic analyses of adjectives proposed in the 1980s-early 1990s were almost exclusively based on English, the inter-linguistic comparison that started in the mid-1990s resulted in further complicating factors that did not emerge before. The approaches that have been suggested in the last decades can be grouped in two main branches on the basis of the presence or absence of noun-movement in the derivation of the structures of adjectival modification. On the one hand, different analyses argued in favour of different types of movement of the noun (Dobrovie-Sorin, 1987; Crisma, 1990, 1993; Cinque, 1994, 2004a, 2005b, 2007b; Menuzzi, 1994; Scott, 1998, 2002a,b; Knittel, 2005), or of the adjective (Giorgi & Longobardi, 1991; Demonte, 1999a; Larson, 1998), or both (Valois, 1991b; Kayne, 1994; Bosque & Picallo, 1996; Fehri, 1997, 1998, 1999; Laenzlinger 2000, 2005a,b). On the other hand, some other analyses excluded the presence of any movement within the determiner phrase in this type of structure (Abney, 1987; Lamarche, 1991; Sadler & Arnold, 1994; Svenonius, 1994; Bouchard, 1998, 2002; Dimitrova-Vulchanova, 2003). Taken individually, none of these analyses can account for all the properties of the adjectival modifiers; for this reason, many authors proposed 'mixed' analyses, where different adjectives are derived in different ways. All these 'mixed' analyses divide adjectives into two sub-classes on the basis of the 'closeness' of the relationship they have with the modified noun and that corresponds to the contrast *lexical* and *functional* adjectives discussed in §1.3.2.

1.3.3.2 The syntax of predicative adjectives

Compared to the syntactic structure of attributive adjectives, the syntax of predicative adjectives seems much less controversial. The most common structure used in the literature to account for predicative relations between a determiner phrase and an adjectival phrase is a small clause (Bowers, 1993):

(21)

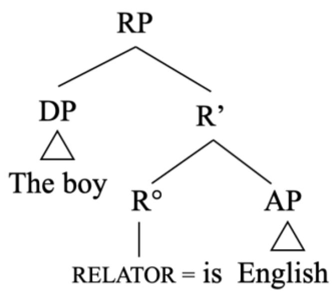


In this structure, a SC is a reduced predicative structure, within which the determiner phrase (the subject) and the adjectival phrase (the predicate) are merged together. Such a structure is selected by the copula or by a verb (like *seem, consider, become*, known as 'quasi-copula') (Cardinaletti & Guasti, 1995; Moro, 1997; den

Dikken, 2006). However, this type of structure only selects lexical adjectival phrases. As discussed above, functional adjectives are excluded since they receive a non-predicative interpretation. Thus, the difference between the syntactic structures of attributive and predicative adjectives is assumed to rely on where the adjective is generated within the structure, i.e., on the left or on the right of the noun phrase respectively. Since many adjectives have both attributive and predicative uses, many scholars have been tempted to reduce attributive adjectives to a single case. In §1.3.4 we will discuss one of the main attempts to explain the relationship between attributive and predicative structures, i.e., to view the attributive use of adjectives as derived from the predicative via a relative clause.

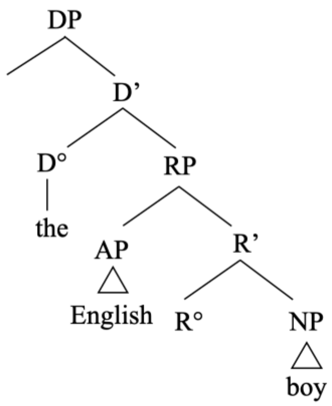
The same configuration for attributive and predicative structures is proposed by den Dikken (2006). He assumed that all subject-predicate relationships are mediated by a *Relator*, an abstract functional head of no designated category. An important assumption of his account is that predication relationships are non-directional. Thus, his analysis accounts for both predicative and attributive structures. In predicative structures as in (22), the subject is in the specifier of the Relator-phrase (RP) and the adjective is in the complement of the Relator-head (R0). The Relator-head is occupied by the copula.

(22) The boy is English.



In attributive structures, as in (23) below, the predicate is in the specifier of RP and the subject is the complement of the Relator-head. In this configuration, the Relator-head is empty. Since the adjective is in the extended projection of the NP and the Relator can be of any category, this configuration can account for the agreement between adjective and noun.

(23) The English boy.



Besides the fact that some adjectives cannot occur in predicative position, another argument against a unified syntactic analysis of attributive and predicative adjectives is that they are associated with differences in meaning. For example, in the Italian example (12) discussed in §1.3.2, *Il ragazzo povero* ('The poor boy') and *Il povero ragazzo* ('The pitiable boy'), the adjective *povero* has two different meanings. According to den Dikken (2006), this problem can be solved by different lexicalizations of the Relator-head.

In summary, many different analyses have been proposed to account for attributive and predicative uses of adjectives and the relation between attributive and predicative structures has not been resolved yet. The observation that most adjectives can occur in attributive and predicative position (or in pre- and postnominal position inside the DP) is taken as an argument for assuming a derivational relation between the two adjective positions. We will discuss this proposal in §1.3.4.

1.3.4 Attributive adjectives as reduced relative clauses

As we have already pointed out, many adjectives can assume both predicative and attributive functions. The observation that most adjectives can occur in attributive and predicative position has brought Port Royal grammarians before and generativist scholars[3] later, to describe one function transformationally in terms of the other. In this view, attributive adjectives are derived from predicative adjectives via a (reduced) relative clause (see Demonte, 2011, for an overview). Thus, an adjectival modifier was considered to be the result of a series of operations called *Whiz Deletion*[4] (or *Relative Clause Reduction*; Berman, 1974; Huddleston, 1976) and *Adjective Shift* (or *Adjective Preposing*; Berman, 1974):

(24) a. The table that was big (Whiz Deletion) →
b. The table big (Adjective Shift) →
c. The big table

As illustrated in (24), a noun phrase whose head is modified by an adjective is analysed as derived from a relative clause such as the one in (24a); an (optional) *Whiz Deletion* is then applied to such structure, resulting in (24b). Finally, to account for the word order within the noun phrase, an (obligatory) *Adjective Shift* is applied, resulting in (24c). This type of analysis is in line with the observation that in many languages predicative adjectives are morphologically simpler than attributive adjectives, e.g., in German, attributive adjectives have agreement morphology, whereas predicative adjectives do not. Though attractive in its simplicity, this analysis encounters a number of problems. As Bolinger (1967), among others[5], has observed, not all prenominal adjectives can undergo such a derivation since many attributive adjectives can never be predicative, as in (25):

(25) The main reason
*The reason is main
derived from: *The reason which is main

[3] Chomsky, 1957, 1965; Smith, 1961, 1964; Katz & Postal, 1964; Lees, 1968; Vendler, 1968; Bach, 1974.

[4] The name *Whiz Deletion* comes from the elements that are deleted in the derivation of English adjectives, i.e., the relative pronoun and the copula *be*. In case the noun modified by the adjective is singular and refers to a [+human] referent, the deleted element corresponds to *who is* (> *whiz*).

[5] See also Winter, 1965; Levi, 1973, 1975; Berman, 1974; Emonds, 1976.

Second, the operation that optionally turns a post-nominal adjective into a prenominal one would have to be made obligatory in some cases (26), but blocked in others (27):

(26) *the house [red] > the [red] house

(27) the people [ready] > *the [ready] people

Third, the possibility of appearing in attributive or predicative position can change when adjectives are modified (Huddleston & Pullum, 2002), as in the examples in (28) below:

(28) a. The patient is awake

b. *The awake patient

c. The wide-awake patient

A final problem for this analysis is that there exist also adjectives that are predicative but rarely or never attributive (at least maintaining the same meaning):

(29) a. The man is asleep

b. *The asleep man

These examples suggest that it is not possible to always correlate the two types of modifications, not even if we consider the attributive adjective as the starting point, since *The man is asleep* should derive from **The asleep man*, which is ungrammatical (Ramaglia, 2008).

Another important observation is that predicative use and attributive use of adjectives may result in differences in meaning (Bolinger, 1967), as shown in (30):

(30) a. The old director (= former)

b. The director is old (= elderly)

If attributive adjectives were uniformly derived from their predicative use, we would expect them to have the same meaning (Bolinger, 1967).

The distinction between the attributive and the predicative use of adjectives is interpreted as reflecting structural differences between two types of modifiers: attributive adjectives are generated to the left of the noun head of the determiner phrase, while predicative ones are taken to be generated to its right, within a small clause (Alexiadou, 2001). The example (31) shows this contrast:

(31) a. a proud person
b. a person proud of his family

It has further been observed that attributive adjectives present a distinction that seems to be related to the attributive/predicative distinction (Bolinger, 1967; Sproat & Shih, 1988). Indeed, attributive modification can be either direct or indirect. Specifically, while direct modifiers are simple adjectival phrases, indirect modifiers are reduced relative clauses (Sproat & Shih, 1988). This syntactic distinction is reflected in two properties. Direct modification allows both intersective and non-intersective modifiers, whereas indirect modifiers are limited to intersective adjectives and are not subject to ordering restriction. If Sproat and Shih's analysis is correct, attributive adjectives cannot be generally reduced to predicative adjectives and, as a consequence, to relative clauses, supporting the conclusion that two types of adjectives have to be distinguished (Cabredo Hoffher, 2010).

1.3.5 Summary

In this section, we have presented an overview of the syntax of adjectives, starting from the two main functions adjectives can fulfil, i.e., attributive and predicative (§1.3.1). Although these functions are shared cross-linguistically, a substantial amount of morphosyntactic variation must be taken into account, including the pre- and post-nominal position adjectives can occupy inside the DP (§1.3.2). We analysed the corresponding syntactic structures of both attributive and predicative adjectives, discussing different approaches accounting for the syntactic difference in adjective use (§1.3.3). Finally, we discussed well-known arguments in favour and against defining the relationship between the two functions as the result of a transformation via a relative clause (§1.3.4). In conclusion, many different hypotheses have been proposed to account for the exact nature of the derivation of adjecive positions. However, the relationship between attributive and predicative structures are still open topics in syntactic research. In §1.4 we turn to the meaning of adjectives.

1.4 Semantics of adjectives

As we have already observed, finding criteria for the identification of the class of adjectives that hold cross-linguistically is elusive. Nonetheless, the most familiar and intuitively appealing approach for distinguishing adjectives from the other word classes is based on their denotational meaning. Nouns generally denote “people, places, and things”, verbs “actions and states”, whereas adjectives denote “properties and qualities”; for this reason, adjectives are referred to as “property concepts” (Thompson, 1988). In what follows, we will discuss the core semantic types of adjectives (§1.4.1), the main classifications that have been proposed in the literature based on semantic criteria (§1.4.2), and the class of relational adjectives (§1.4.3).

1.4.1 Core semantic types

Some prototypical semantic properties are found across most languages, independent of the size of the class in the respective language (Dixon, 1982). For languages that have adjectives as a word class, Dixon (1982) showed that the semantic content, i.e., the concepts expressed by adjectives, is similar across languages. There are four core “semantic types”, usually connected to both large and small adjective classes:

DIMENSION: big/small, tall/short …
AGE: new, young, old…
VALUE: good, bad, lovely…
COLOR: black, white, red…

Only when the adjective class is medium-sized or large (as in languages like Italian and English), it is likely to have also peripheral semantic types:

PHYSICAL PROPERTY: hard, soft, heavy…
HUMAN PROPENSITY: jealous, happy, kind…
SPEED: fast, quick, slow…

Connected only to large adjective classes, some languages include also other semantic types, such as:

DIFFICULTY: easy, difficult, though
SIMILARITY: like, unlike, similar
QUALIFICATION: definite, true, probable
POSITION: high, low, near

Among the three parts of speech, property concepts are the most problematic, since they are not semantically consistent across languages: they show a great deal of intra-linguistic and cross-linguistic variation, which makes a purely semantic definition problematic. As a consequence, a semantic definition of parts of speech does not always predict the lexical classification of a word in a language based on its meaning (Beck, 2002). For instance, the physical property 'hard' is a noun in Hausa, but surfaces as an adjective in Spanish and as a verb in Lushootseed (Schachter, 1985), although the word is a property-concept and should always be defined as adjective according to a purely semantic definition of lexical classes.

1.4.2 Main semantic-based classifications of adjectives

In the literature, many classifications have been proposed to divide the heterogeneous class of adjectives into various subclasses on the basis of different semantic criteria. The following section will illustrate the main classifications, discussing the interpretations linked to the different kinds of adjectives.

1.4.2.1 Intersectivity

Adjectives can be classified according to set-theoretic terms and to the inferences that an adjective+noun combination (or intersection) triggers. Thus, entailment-based typologies focus on the adjective's function as a modifier and classify adjectives according to the inferences they license (McNally, 2016). One appealing hypothesis assumes that adjectives and nouns are both simple one-place predicates that denote sets and their combination indicates the intersection of the two (Kamp & Partee, 2005). As shown by Kamp and Partee, however, this hypothesis does not hold for adjectives in general, but only for intersective adjectives, that are adjectives that licence inferences between the attributive and the predicative use based on both noun and adjective:

(32) Intersective adjectives: Licensed inferences

a. X is Adj N ⇒ X is a N   X is a red house ⇒ X is a house

b. X is Adj N ⇒ X is Adj   X is a red house ⇒ X is red

The intersection hypothesis also explains the recursiveness of modifiers, since intersection can be iterated freely. The result of the intersection between two sets A and B is still a set C, which in turn can undergo intersection with another set D and so on. However, not all adjectives trigger an intersective interpretation. The class of subsective adjectives allows for the licencing of different types of inferences. Subsective adjectives can be interpreted with respect to a class of comparison defined by the noun, i.e., only one of the patterns of inference is fulfilled, precisely the inferences based on the noun:

(33) Subsective adjectives: Licensed inferences

a. X is Adj N ⇒ X is a N   X is a perfect typist ⇒ X is a typist

b. X is Adj N ⇏ X is Adj   X is a perfect typist ⇏ X is perfect

A perfect typist is perfect as typist and not in absolute terms; he could be, for example, not a perfect husband. The fact that only (33a) is triggered explains why two sentences such as (34a) and (34b) do not necessarily imply (34c) (examples from Kamp & Partee, 1995:138):

(34) a. Mary is a skilful surgeon.
b. Mary is a violinist.
c. Mary is a skilful violinist.

In principle, subsectivity and intersectivity are independent properties of adjectives. However, three combinations of subsectivity and intersectivity are attested in the literature on the semantics of adjectives (e.g., Kamp & Partee, 1995):

1. Adjectives that are subsective and intersective.
2. Adjectives that are subsective, but not intersective.
3. Adjectives that are neither subsective nor intersective.

Note that the term "subsective" is sometimes used to refer to non-intersective adjectives (e.g., Kamp & Partee, 1995; Partee, 1995). Following Weiker (2019), in the present work we use "subsective" in its literal sense, namely for all adjectives for which the inference *X is an N* holds: the set denoted by the adjective-noun phrase is a subset of the set denoted by the noun. Adjectives for which intersectivity does not apply are called "non-intersective". Thus, we assume intersective adjectives to be a kind of subsective adjectives, which also includes gradable adjectives that will be discussed in §1.4.2.2.

There is another class of adjectives (such as *former* or *alleged*) which are neither subsective nor intersective: non-subsective adjectives. Non-subsective adjectives can be divided into simple non-subsective and privative (Kamp & Partee, 1995). With simple non-subsective adjectives the adjective+noun combination implies neither the adjective nor the noun:

(35) Simple non-subsective
a. X is Adj N ⇏ X is a N   X is an alleged murderer ⇏ X is a murderer

b. X is Adj N ⇏ X is Adj X is an alleged murderer ⇏ *X is alleged

By contrast, with privative adjectives (such as *fake* or *counterfeit*), the adjective+noun combination implies a negative inference for the noun, that is, this combination is never an instance of the noun alone (Kamp & Partee, 1995):

(36) Privative

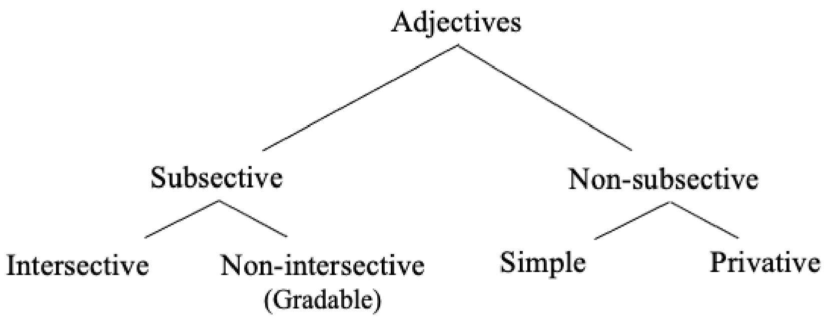
a. X is Adj N ⇒ X is not a N X is a fake diamond ⇒ X is not a diamond

b. X is Adj N ⇏ X is Adj X is a fake diamond ⇏ X is fake

Fake is a privative adjective since a fake diamond is not a diamond, while *alleged* is not because an alleged murderer may or may not be a murderer.

Based on the entaillment-based properties, the following classification of adjectives emerges:

(37)



The distinction based on intersectivity can be partially linked with the syntax of the adjectives. As a matter of fact, only attributive adjectives can allow intersective or non-intersective meaning, while predicative adjectives are always intersective (Cabredo Hoffher, 2010).

(38) Olga is a beautiful dancer

The well-known sentence in (38) is an example[6] of the ambiguity of the noun phrase between an intersective and a non-intersective interpretation. In the first case, the person denoted by *beautiful dancer* is a member of the intersection between the set of elements that are beautiful and the set of elements that are dancers (as paraphrased in (39a)). In the second case, the interpretation of the adjective is adverbial (as paraphrased in (39b):

(39) a. Olga is beautiful and a dancer (inters. interpretation of (38))
b. Olga dances beautifully (non-inters. interpretation of (38))

However, this opposition is relevant only for attributive adjectives. By contrast, a predicative adjective is always applied to the extension of the subject noun.

(40) The dancer is beautiful

In the sentence in (40), the adjective and the referred noun are the same of the sentence in (38); the difference between the two examples is in the attributive vs. predicative function of the adjective itself. Such a difference is associated to a difference in the interpretation. While the adjective *beautiful* in (38) is ambiguous, in (40) it is not, being the adverbial interpretation (in (39b)) impossible: the property of being beautiful is only applicable to *dancer* and not to his/her way of dancing. Moreover, non-intersective adjectives occur strictly pre-nominally in English and in the Romance languages, as shown for Italian in in (41) (from Alexiadou, 2001).

(41) a. Un mero accidente (adjective-noun order)
b. *Un accidente mero (noun-adjective order)
'A mere accident'

[6] See e.g., Vendler (1968); Siegel (1976, 1980); Larson (1995, 1998, 1999); Cinque (2005b, 2007b); Alexiadou et al. (2007).

For the purpose of the present work, we are only interested in adjectives licensing the subsective interpretation. To capture the differences within the class of subsective adjectives, we refine this class by adding another important property of adjectives: gradability.

1.4.2.2 Gradability

Differences within the class of adjectives that we refer to as subsective, i.e., the intersective vs. non-intersective distinction discussed in the previous section (§1.4.2.1), can be captured by their gradability, which has always been considered a prototypical characteristic of adjectives (see e.g., Jackendoff, 1977) and has often been used as a criterion to distinguish adjectives from other word classes such as nouns and verbs (e.g., Dixon, 2004). In addition, Kennedy (2006) points out that ordering and comparing objects is a basic component of human cognition. Natural languages reflect this cognitive ability with syntactic categories that express gradable concepts.

Not all adjectives are gradable. A first distinction that needs to be made is between gradable and non-gradable adjectives. Gradable adjectives like *big* or *beautiful* can occur in comparative constructions (42a) and can be combined with degree modifiers (42b)), whereas non-gradable adjectives such as *Italian* normally cannot (43).

(42) a. The elephant is bigger than the mouse.
b. The elephant is too big.

(43) ? Paolo is more/too Italian.

Bierwisch (1987, 1989) notes that combining adjectives like *Italian* with comparative morphemes and degree modifers is in general possible. However, Bierwisch, among others, classifies *male* as non-gradable. Other non-gradable adjectives are color adjectives. According to Kennedy and McNally (2010), color adjectives have both a gradable and a non-gradable reading. The gradable reading refers to either the color quantity or the color quality. The non-gradable reading refers to the classificatory reading as in *red wine*. Demonte (2011) argued that, although color adjectives can occur with degree modifiers and can accept comparative morphology they differ from gradable adjectives with respect to the content of their scale: scales of gradable adjectives consist of degrees whereas scales of color adjectives consist of prototypes and shades of the respective color. An additional argument for their non-gradability is that they do not come in antonym pairs, which is common for gradable adjectives. We assume color adjectives to be non-gradable[7]. The gradable vs. non-gradable (or non-intersective vs. intersective) distinction is relevant for the present work in acquisition perspective. Implications for language acquisition will be discussed in detail in chapter 2.

As already pointed out, adjectives are classified as gradable if they can occur in comparative constructions and if they can be modified by degree expressions like *very*, *too* or *enough*. However, among the class of gradable adjectives, differences can be found concerning other degree modifiers, the interpretation of borderline cases, characteristics of vagueness and their context-sensitivity (Kennedy & McNally, 2005; Kennedy, 2007; MacNally, 2011). Kennedy and McNally (2005) proposed a semantic typology based on the property of the scales along which predicates order their arguments (or their "scalar structure"). These predicates are classified along two parameters:

i. whether the scale is open or closed
ii. whether the standard of comparison is relative (contextual) or absolute (regardless of context, there is a maximal and a minimal value on a scale; adjectives correspond to the endpoints of the scale).

This results in two different cases:

1. Open scale relative adjectives: *big*
   No upper limit (incompatible with "completely"): **completely big*

[7] It is often possible to coerce a gradable interpretation from a non-gradable adjective. For instance, "The table is more wooden than the door" might be understood to mean the table has more wood in it than the door, but such interpretation is clearly marked (Kennedy, 2007).

Relative standard of comparison (compatible with “very”): *very big*

2. Closed-scale absolute adjective: *undocumented*
   Upper limit on a scale (compatible with “completely”):
   completely undocumented
   Absolute standard of comparison (incompatible with “very”):
   **very undocumented*

Kennedy (2007) identified three important differences between relative and absolute adjectives. A first difference has been observed with respect to characteristics of vagueness. Indeed, relative adjectives give rise to the Sorites Paradox, as exemplified in (42). If the difference in height between two people is very small, it is unlikely that we will judge one as tall, but not the other. More generally, if we assume Premise 1 and given Premise 2, the general form of which is held to be valid for relative adjectives, they bring to the absurd conclusion in (42c) (MacNally, 2011).

(44)

a. Premise 1: A 1.65-meter-tall eleven-year-old is tall (for an eleven-year-old).
b. Premise 2: If X is a tall eleven-year-old and Y is an eleven-year-old who is one millimeter shorter than X, then Y is a tall eleven-year-old.
c. Conclusion: A one-meter-tall eleven-year-old is a tall eleven-year-old.

By contrast, if we try to recreate the same paradox with an absolute adjective like *closed*, Premise 2 will fail; indeed, if we open a door which is closed even the smallest amount, we can easily determine that the door is no longer closed. This happens in general for all absolute adjectives (Kennedy, 2007).

A second difference concerns the fact that relative adjectives allow to find borderline cases, for which it is not possible to establish whether the adjective truthfully holds or not. Borderline cases exist for relative adjectives, but not for absolute adjectives. *Tall* is a case in point: in any given context there is a set of objects that can be clearly judged as tall, and another set of objects that can be judged as *not tall*. For instance, if we discuss the height of a group of children of the same age, but we only take into account the group and we do not consider the general knowledge about the height of children that age, we might have trouble in judging a 1.50-meter-tall kid as *tall*, if the height range of the group varies, e.g., from 1.30 to 1.70. In addition, there are also objects that are less well-defined. Thus, in antonym pairs the negation of one adjective does not entail the assertion of the other adjective, as shown in (45) below.

(45) John is not tall ⇏ Paul is short.

The infelicitous inference in (45) results from the context-sensitive standards of relative adjectives (Klein, 1980): the standard for *tall* and *short* need not to be the same degree (Kennedy, 2007). This is also pointed out by Solt (2011). If the negative antonym is analysed as the direct opposite of the positive antonym, then the standard for both would be the same single degree, that completely divides the scale into tall and short entities. However, there are sizes that are judged as neither *tall* nor as *short*, as suggested by the felicity of sentences like "John is not tall, but he is not short either"[8]. Conversely, no such difficulties arise with absolute adjectives like *closed* (MacNally, 2011). Because borderline cases do not exist for absolute gradable adjectives, the standard should be a single degree on the scale, which is shared by the positive and the negative antonym (Solt, 2011).

A further difference is that relative and absolute adjectives show different degrees of context sensitivity (e.g., Rotstein & Winter, 2004; Kennedy & McNally, 2005; Kennedy, 2007; McNally, 2011; Sassoon & Toledo, 2011; among many others). Take the examples in (46).

(46) a. Marta is tall. (relative)

---

[8] Solt (2011) defined antonym pairs of relative adjectives as "contraries", i.e., there is for instance a range of sizes which count as neither big nor small. Antonym pairs of absolute adjectives are "complementaries" i.e., the standard of comparison is a single degree, which is shared by the positive and the negative antonym.

b. The door is closed. (absolute)

According to Kennedy (2007), the sentence in (46a) is context dependent, while the one in (46b) is not. The addition of an explicit standard of comparison to the sentence in (46a) can render it either possibly true (47a)-(47b) or almost certainty false (47c)-(47d).

(47) a. Compared to her friend Andrea, Marta is tall.
b. Marta is tall for an 11-year-old.
c. Compared to Michael Jordan, Marta is tall.
d. Marta is tall for a professional basketball player.

However, the same addition is not felicitous with (46b), suggesting that a contextually-specified comparison class is not compatible with the adjective itself (as in (48a) and (48b)) (MacNally, 2011).

(48) a. ?? Compared to Door #1, Door #2 is closed.
b. ?? That box of cookies is closed for a box my daughter has gotten into.

This distinction seems to be further supported by the distribution of *for*-phrases. As a matter of fact, they restrict the contextual domain that determines the standard, as shown in (49) (Siegel, 1979; Toledo & Sassoon, 2011):

(49) a. John is tall for a ten-year-old boy.
b. *This shirt is dirty for a T-shirt.
c. *This knife is clean for a kitchen knife.
d. *This glass is full for a wine glass.

These observations led to the conclusion that the interpretation of relative adjectives is always linked to a standard of comparison, which is located at the mid-point of the scale. To count as *tall*, John needs to exceed a contextually determined standard of height (Kennedy, 1999, 2007), a degree that corresponds to an "average" or a "norm" for the scalar concept relative to the salient set, which may be explicit, inferred or implicit in the sentence. According to Kennedy (2007), context-sensitivity with relative adjectives derives from the fact that their scales are open and unbounded at both endpoints. Absolute adjectives, on the other hand evoke a scale which is closed in either one or both thresholds. For this reason, they either have a *minimum standard* or a *maximum standard*, conventionally provided by one of the scale's endpoints. *Minimum standard* absolute adjectives, such as those in (50), simply require their arguments to possess some minimal degree of the property in question.

(50) a. The table is wet.
b. The door is open.
c. The rod is bent.

The sentence in (50a) does not mean that the degree to which the table is wet exceeds some context-dependent of wetness (for a table), it is true as long as there is some amount of liquid on the table. Likewise, (50b) requires a minimal aperture of the door and (48c) requires a rod that has a non-zero degree of bend.

Maximum standard absolute adjectives such as those in (51) require their arguments to posses a maximal degree of the property they describe.

(51) a. The floor is dry.
b. The door is closed.
c. The rod is straight.

The sentence in (51a) typically means that the floor has no moisture on it. (51b) is uttered when a door is completely closed and (51c) is true for a completely straight rod. However, Rotstein and Winter (2004) have accounted for evidence in favour of context sensitivity also for absolute adjectives[9]. Consider the examples (52) and (53) below.

(52) a. This kitchen knife is clean.
b. This surgical instrument is clean.

(53) a. This child's shirt is dirty.
b. This tuxedo is dirty.

In examples (52) and (53) the standard of comparison of *clean* and *dirty* depends, admittedly, on the noun they modify. Hence, the standard for *clean* is higher for a surgical knife and lower for a kitchen knife. Likewise, the standard of *dirty* can vary widely for a child's shirt and a tuxedo (Toledo & Sassoon, 2011).

Toledo and Sassoon (2001) proposed to distinguish between *within-individual* and *between-individuals* interpretations. In the former interpretation, the adjective's argument is compared to its "counterparts". For instance, "the description of a shirt as *dirty* or *clean* is based on a visualization of that particular shirt in various degrees of grubbiness rather than on its juxtaposition with other concrete shirts. […] The nature of the scale (whether it is open or closed), is imposed by the individual under consideration, e.g., we can imagine a maximally clean counterpart of the above shirt" (Toledo & Sassoon, 2011:141-142). On the other hand, *between-individuals* interpretation concerns the comparison of an individual to other distinct individuals. For example, "the comparison class of adjectives such as *tall* or *short* may comprise any of many possible categories, each imposing equally salient natural height bounds, or no bounds at all" (Toledo & Sassoon, 2011:142). Thus, both types of comparison classes depend on contextual considerations (Toledo & Sassoon, 2011). In the present work, we will focus on Kennedy (2007)'s view, according to which, truth-conditionally, absolute adjectives always have endpoint-oriented meanings and, unlike relative adjectives, are context-independent.

[9] Rotstein and Winter (2004) refer to adjectives like those in (50) 'partial' predicates and adjectives like those in (51) 'total' predicates (using terminology introduced in Yoon, 1996)

In summary, subsective adjectives can further be classified according to the property of gradability as relative and absolute adjectives. Relative adjectives show properties of vague predicates. Due to their context-sensitivity, their standard of comparison is assumed to be located around the center of the scale. Absolute adjectives on the other hand are not vague. They evoke a scale which is closed in either one or both thresholds. For this reason, they either have a *minimum standard* or a *maximum standard*, provided by one of the scale's endpoints.

1.4.2.3 Aspectual classes of adjectives

Another aspect which is part of the semantics of adjectives concerns aspectual classes. Aspectual distinctions in the lexicon have to do with the temporal organization of events, as aspect is a grammatical feature expressing how eventualities extend over time. Against this backdrop, it is not surprising that research on aspect has focused on verbs, as it is the word class which encodes events. However, more recent studies on aspectual distinction have explored other word classes including adjectives. We will now analyse two aspects of predicates that have been applied to the word class of adjectives, i.e., the individual- vs. stage-level distinction (Carlson, 1977) and the subevental structure of adjectives (see Smith, 1991; Hay, Kennedy & Levin, 1999; Kennedy & Levin, 2008).

Individual level vs. stage level

Firstly, the distinction between individual-level and stage-level predicates, frequently used for types of stativity, can be also used to distinguish adjectives (Carlson, 1977; Dowty, 1979; Pustejovsky, 1995). Predicates such as *tall*, *intelligent*, and *overweight* might be thought of as properties that an individual preserves throughout his lifetime; they are called individual-level predicates, since they can be identified directly with the person or object. On the other hand, adjectives like *hungry*, *sick*, and *clean* generally refer to non-permanent states of an individual and have been called, consequently, stage-level predicates (Carlson, 1977). This semantic distinction between adjectives is mirrored, for example, in the ability of most stage-level adjectives to enter predicates with the progressive aspect, whereas individual-level adjectives cannot, as shown by the examples (54) and (55) below (examples from Pustejovsky, 1995).

(54) a. The horse is being gentle with her rider.
b. You're being so angry again!
c. Stop being so impatient.

(55) a. *John is being tall today.
b. *Aren't you being beautiful tonight!
c. *Stop being so intelligent.[10]

Starting from Bolinger (1967)'s work, different authors (Ferris, 1993; Sadler & Arnold, 1994; Larson, 1998, 1999, 2000b; Larson & Marušič, 2004; Cinque, 2005b, 2007b) noticed that when an English adjective can appear both pre- and post-nominally (e.g., the class of adjectives in *-able/-ible*), the two positions result in two slightly different interpretations, as shown in (56):

(56) a. The only river navigable
b. The only navigable river

When an adjective like *navigable* modifies the noun *river*, the denoted property is interpreted as temporary (or occasional) if the adjective follows the noun (56a), and as stable (or characteristic) if the adjective is pre-nominal (56b). In other words, the phrase in (56a) refers to a river that is the only navigable river in a specific moment in time (for example, because of adverse weather conditions), whereas the phrase in (56b) indicates a river which always has the property of being navigable. Thus, (56a) contains a stage-level predicate, while (56b) an individual-level predicate.

[10] This sentence could be considered grammatical in particular contexts. It is a case of semantic coercion that "may alter the distributional behaviour of words, allowing them to show up in contexts in which they would otherwise not appear" (Pustejovsky & Jezek, 2008). In this specific case, a stage-level interpretation is applied to the individual-level adjective "intelligent".

The difference between the two interpretations, however, is not linked to the relative position of noun and adjective. Larson (1998) pointed out that in English the post-nominal position of the adjective is linked to a stage-level reading, but the pre-nominal adjective is ambiguous between the two possible interpretations, as illustrated in (57):

(57) a. The visible stars (include Capella, Betelguese, and Sirius) (ambiguous)
b. The stars visible (include Capella, Betelguese, and Sirius) (stage-level only)

In (57a), the visibility of the stars can be interpreted as permanent property, i.e., their brightness makes them visible to a naked eye, or as a temporary state, i.e., during a cloud-covered night, it is possible that (57a) is true and that (57b) is false at the same time. Furthermore, as observed by Larson (1998), "[a]lthough Bolinger speaks in terms of pre- and post-nominal position, in fact the relevant contrast is not one of linear order, but rather of relative closeness to [the] N[oun]". It is possible that two formally identical adjectives linked to two different readings (i.e., stage- and individual-level) co-occur as modifiers within the same noun-phrase, as in (58):

(58) a. The visible$_{[\text{i-level}]}$ stars visible$_{[\text{s-level}]}$ (include Capella)
b. The visible$_{[\text{s-level}]}$ visible$_{[\text{i-level}]}$ stars (include Capella)

In both sentences in (58), the subject is the group of stars that are both inherently and temporarily visible. In (58a), the pre-nominal adjective is interpreted as individual-level and the post-nominal adjective as stage-level, as expected from the pattern in (57). As for the phrase in (58b), native speakers interpret the left-most adjective as stage-level, i.e., corresponding to the post-nominal adjective in (58a) and as individual-level the adjective closer to the noun. This intuition is confirmed by the contrast between the sentences in (59) below (Larson, 1998):

(59) a. The invisible visible stars (include Capella)

b. *The visible invisible stars (include Capella)

The phrase in (59a) makes sense because it refers to the stars that are inherently visible, but that are invisible in a specific moment. On the contrary, the phrase in (59b) is not coherent because it should refer to those stars that are intrinsically invisible, but that are visible in a particular time. Thus, the contrast in (59) confirms that, in a sequence of two pre-nominal adjectives with two different readings on the basis of the dichotomy stage-level vs. individual-level, the first is interpreted as stage-level and the second as individual-level. This contrast is also present in some Romance languages, including Italian (Cinque, 2005). However, English and Italian differ in the position of the adjectives with respect to the noun, as illustrated in (60), the translation of (59) into Italian:

(60)

a. Le invisibili stelle di Andromeda sono molto distanti. (individual-level only)

b. Le stelle invisibili di Andromeda sono molto distanti. (ambiguous)

While in English the pre-nominal adjective is ambiguous and the post-nominal is interpreted as stage-level, in Italian the adjective is ambiguous post-nominally and the non-ambiguous interpretation is the pre-nominal individual-level reading.

The subevental structure of adjectives

In a manner similar to verbal predicates (Smith, 1991), the meaning of adjectives can be analysed in terms of subeventive structure, i.e., the internal structure of the interval for which the state holds, as shown in the examples in (61):

(61) a. He was dead. (transition, no right boundary)

b. He was changed. (transition, right boundary not restricted)

Both adjectives imply that the state holds for an interval having a left boundary, i.e., a moment in which the state began. However, the interval for which *dead* in (61a) is true does not have a right boundary, due to its lexical meaning. On the contrary, *changed* in (61b) is neutral with respect to the length of the interval (Cabredo Hofherr, 2010).

Another element that affects the temporal trace of which the state holds is gradability. It affects the time interval in two respects:

i. The distribution across the interval of the property denoted by the state
ii. The possible transition from state to non-state

Consider the examples in (62):

(62) a. He was drunk/sick.
b. The shop is open/closed.
c. The door is open/closed.

In (62a), *drunk* is compatible with different degrees of drunkenness over an interval, i.e., it is a gradable state, whereas *open/closed* for a shop (62b) holds or does not hold, i.e., it is a yes/no state. Thus, the two adjectives contrast with respect to the distribution of the property over the time interval. However, they also contrast as regards the possible transition from state to non-state. The transition from *sober* to *drunk* is a matter of degree, while that from *open* to *closed* (for a shop) is not. In other words, while the interval for which *open* holds has a right boundary, this might not be true for *drunk*, as the transition is gradual (Cabredo Hofherr, 2010). This analysis, however, does not hold for *open/closed* in all contexts. If we consider the example (62c), the adjective *open* can be used to describe various degrees of openness (from half-closed to wide-open, for a door), while *closed* can only refer to one state (a closed door). Thus, like *drunk*, the adjective *open* is compatible with an interval. As discussed in §1.4.2.2, we assume that both *drunk* and *open* are gradable adjectives and that all gradable adjectives evoke scales. The difference between the two has to do with the boundaries of the scale: while *open* is closed on both ends, drunk is a *minimum standard* gradable adjective.

In summary, a further semantic distinction between adjectives concerns their aspectual classes, which have to do with the temporal organization of the events. More specifically, the individual- vs. stage-level distinction and the subevental structure of adjectives have been taken into account as further criteria.

1.4.3 Relational adjectives

Relational adjectives[11] are denominal adjectives, that is to say, adjectives etymologically or morphologically derived from nouns through a productive synchronic process. In the generative semantics, these adjectives are referred to as *pseudo-adjectives* or *transpositive*, since they are transpositions from non-referential nouns (Bally 1944; Marchand 1966, 1967). Such adjectives do not denote properties, but entities defined by the noun to which relational adjectives are morphologically connected (Bosque & Picallo, 1996). For example, adjectives such as *atomic* (linked to the noun *atom*) and *Italian* (linked to the noun *Italy*) are of this adjectival type (Ramaglia, 2008).

(63) a. The atomic bomb
b. The Italian language

Relational adjectives express a relationship between the noun from which they derive and the element to which they referred (Renzi et al., 2001). In this way, both noun and adjective contribute to the identification of the discourse reference. As a consequence, they are intrinsically restrictive, i.e., they restrict the reference of the head noun (Ramaglia, 2008).

---

[11] For detailed studies on the syntax and interpretation of such adjectives, see Bolinger (1967); Vendler (1968); Postal (1969); Ronat (1974); Levi (1973, 1974, 1978); Bartning (1976); Warren (1984, 1988); Bosque (1993); Bosque & Picallo (1996); Demonte (1999b); McNally & Boleda (2004); Giegerich (2005); Van de Velde (2006).

Bally (1944) identified three properties that characterise relational adjectives. First of all, they occur only post-nominally in Romance languages (examples from Ramaglia, 2008):

(64) a. Il sistema nervoso / *Il nervoso sistema
b. L'invasione italiana / *L'italiana invasione

Second, relational adjectives cannot have a predicative use (65a) (Bally, 1944), nor they can appear as appositive modifiers (65b) (Ramaglia, 2008).

(65) a. *The heat is solar
b. *The heat, solar, …

Finally, they are not gradable. This is related to their 'taxonomic' or 'classificatory' meaning (Bally, 1944).

(66) *The heat is very solar.

As already mentioned, relational adjectives do not denote properties but entities, and they do not constitute a homogeneous class (Bosque & Picallo, 1996). They can be divided into two subclasses on the basis of the kind of relationship they establish with the head noun: thematic adjectives and classificatory adjectives (examples from Bosque and Picallo, 1996, adapted to Italian).

(67) a. Produzione automobilistica (thematic adjective)
b. Escursione automobilistica (classificatory adjective)

As shown by the examples in (67), (67a) is a thematic adjective since it has an argumental role, i.e., that of theme, assigned by the name *produzione* (*produzione automobilistica* means *produzione di automobile*, 'car production'). In (67b), by contrast, the same adjective introduces a domain of relation to which the object denoted by the head noun is classified (*escursione automobilistica* means *escursione con l'automobile* 'trip by car') (Bosque & Picallo, 1996).

The difference between these two types of relational adjectives does not only concern semantics but also the syntactic structure and, in particular, the order of the adjectives with respect to the noun, and the order of adjectives when combined in a NP (Bosque & Picallo, 1996). In (68), two relational adjectives appear within the same noun phrase:

(68) a. La politica estera italiana
the politics foreign Italian
'The Italian foreign politics'
b. *La politica italiana estera
the politics Italian foreign

If we consider the English translation, in which both adjectives are pre-nominal, it is evident that the distribution of the two types of adjectives is determined by the noun-head distance: the classificatory adjective must appear closer to the noun than the thematic one (Bosque & Picallo, 1996). Hence, the order of adjectival modifiers is rigid and depends on the (semantically based) class to which the adjectives belong (Bosque & Picallo, 1996).

1.4.4 Summary

In this section we have discussed the semantics of adjectives and their classification based on meaning. In §1.4.1 we explored the core semantic types of adjectives that are common to all languages and those that are present in languages with a large and productive class of adjectives. In §1.4.2 and §1.4.3 we illustrated classifications on the basis of semantic criteria, i.e., intersectivity, gradability, aspect, and the class of relational adjectives. In all these cases, we observed that the differences at the semantic level correspond to differences in syntactic behavior and, thus, that the two aspects are closely related. We turn now to a last criterion to identify adjectives cross-linguistically, namely the criterion based on their pragmatic function.

1.5 Pragmatic functions of adjectives

In recent years, various scholars have claimed that nouns, verbs and adjectives are universal prototypical word classes (e.g., Baker, 2003; Dixon & Aikenvald, 2004; Dixon, 2004). However, the various attempts to define adjectives as a distinct category are flawed in two ways: they do not account for variation within word classes, and they fail to be general enough to apply across all languages.

An alternative theory of categorisation, the *Prototype Theory*, has been proposed by Rosch (1987) and supported by other scholars (Givón, 1979; Dixon, 1982; Hopper & Thompson, 1984; Croft, 1991; Bhat, 1994). According to the *Prototype Theory*, human categorisation proceeds from central to peripheral instances, with the central instances being prototypical for the category. The more prototypical a member of a category is rated, the more elements it has in common with the other members of that category.

Bhat (1994) proposed the simultaneous use of different criteria in the characterisation of adjectives, taking into account prototypical elements of a given category. As we have already observed, the different criteria examined above are interconnected and form a unified system of explanation (Bhat, 1994). For example, Croft (1991) suggested that word classes can be analysed cross-linguistically according to two independent yet prototypically correlated parameters: (i) the semantic class of the lexical element and (ii) its pragmatic function in the clause structure. The semantic concepts prototypically expressed by nouns, verbs and adjectives are, respectively, object, action and property; whereas their pragmatic functions are, respectively, reference, predication and modification (Bhat, 1994). These functions can be roughly defined as follows: referring denotes what one is talking about, predicating indicates what is being said about and modifying denotes a secondary referring or predicating function. Semantic and pragmatic parameters should remain constant cross-linguistically and, when combined, give the prototypical nouns, verbs and adjectives (Croft, 1991).

Besides those prototypes, there exist also hybrid lexical categories, created through the possible correlations of the above-mentioned semantic classes and pragmatic functions, as shown in Table 1 (Croft, 1991)[12].

Table 1. Parameters for syntactic categories.

	Noun	Verb	Adjective
PRAGMATIC FUNCTION:	Referring	Predicating	Modifying
SEMANTIC CLASS:	Object	Action	Property

The classification of this kind of approach is shown in Table 2, with examples in Table 3.

Adjectives can be defined in terms of their belonging, prototypically, to the semantic class of properties and their primary (categorial) function, that is modification (of a noun), as shown in Table 2 for all categories.

Table 2. Possible combinations of syntactic categories with different semantic classes and pragmatic functions (Croft, 1991).

	Reference	Modification	Predication
OBJECTS	*Core nouns*	Genitives adjectivals, PP modifiers	Predicate nominals
PROPERTIES (QUALITIES)	Abstract de-adjectival nouns	*Core adjectives*	Predicate adjectives
EVENTS (ACTIONS)	Nominalisations, infinitives, gerunds, complements	Participles, relative clauses	*Core verbs*

[12] For a detailed analysis of the use of adjectives in other functions, such as that of participant identification (reference), predication, verbal modification and compound formation, see Bhat (1994).

Table 3. Examples of all the possible combinations of syntactic categories with different semantic classes and pragmatic functions (Croft, 1991) (the examples have been adapted from Croft, 1991).

	Reference	Modification	Predication
OBJECTS	*Dog*	Doglike	Be a dog
PROPERTIES (QUALITIES)	Height	*tall*	Be tall
EVENTS (ACTIONS)	Destruction	Destroyed	*destroy*

Thus, if parts of speech are pragmatically defined, it is possible to identify words in every language that have the modifying property. Accounting for the differences between syntactic categories in terms if associations of transitoriness, isolable properties and stereotypes, Croft (1991) showed that the adjectival function is universal, although it can assume various forms. The different distribution of the words of a language into the different lexical categories is responsible for cross-linguistic variation (Bhat, 1994).

In the last section of the present chapter, we provide an overview of the adjectival phrase in Italian, being the language investigated in the present work.

1.6 The Adjectival Phrase in Italian

The adjectival phrase can be defined in every language according to the type of agreement with the noun it modifies and based on distributional observations. Among all the aspects analysed in the previous sections, the definition of adjectives as a universal part of speech has considered morphological aspects, i.e., the grammatical categories for which adjectives are marked. Hence, nouns have been defined as the lexical items that have grammatical gender and are inflected for number and case; verbs are inflected for tense, aspect, voice, and mood; adjectives are inflected in comparative constructions and, in some languages, also for number, gender, and/or case (Beck, 2002). Again, there is considerable cross-linguistic variation and universal morphological categories that can be used exclusively and univocally for adjectives are difficult to find. A possible solution could be to define a set of grammatical categories that are cross-linguistically typical of the class of adjectives and then decide, language by language, the most appropriate given a grammatical system (Beck, 2002). Another important aspect concerns distribution: the adjectival phrase can also be defined based on its distributional properties, i.e., the set of contexts in which it can appear, especially relative to the referent noun. In the present section, we will summarize the morphosyntactic and distributional properties of the adjectival phrase in Italian, as relevant to this thesis.

Adjectives in Italian can occur in predicative and in attributive structures. When used in predicative structures (e.g., *Il ragazzo è inglese* 'The boy is English'), the adjective is the complement of the copula and agrees with the subject noun in gender and number. In attributive position (e.g., *Il ragazzo inglese* 'The English boy'), the adjective is part of the DP, occurs in the unmarked post-nominal position and agrees with the noun in gender and number.

1.6.1 Morphology of Italian adjectives

According to gender and number, we can distinguish three classes of adjectives: (i) adjectives that vary in both gender and number, (ii) adjectives that only vary in number, and (iii) invariable adjectives (Renzi, et al., 2001). In the first class of adjectives (distinct suffixes for gender and number), the masculine is realised through the ending in *–o* for the singular and *–i* for the plural, whereas the feminine is realised by *–a* for the singular form and *–e* for the plural, as for *bello* ('beautiful'), shown in (69) (examples from Renzi et al. 2010).

(69)	Masculine singular:	*bello*
	Feminine singular:	*bella*
	Masculine plural:	*belli*
	Feminine plural:	*belle*

Always in this first class, there are adjectives which are invariable in the singular form, but vary in the plural: they are adjectives ending in *–ista*, a gender-neutral suffix, like in *ottimista* ('optimistic'), *egoista* ('selfish'), *idealista* ('idealist') (Renzi et al. 2001),

(70) Masculine singular: *ottimista*
Feminine singular: *ottimista*
Masculine plural: *ottimisti*
Feminine plural: *ottimiste*

The second class includes adjectives which end in *–e* in the singular form and in *–i* in the plural, so that they are not distinguished for gender; for example, *abbondante* ('abundant'), *amabile* ('lovely'), *socievole* ('friendly') and so on (Renzi et al. 2001).

The third category comprises compounds made up of an adjective of colour and an adjective which qualifies the colour: *castano chiaro* ('light brown'), *verde sbiadito* ('faded green'), *rosa pallido* ('pale pink') and similar; forms which consist of an adjective of colour modified by a noun, like *giallo oro* ('gold yellow'), *rosso fuoco* ('fire red'), *verde bottiglia* ('bottle green') and some adjectives that denote colours, such as *carminio* ('carmine'), *nocciola* ('nut'), *vermiglione* ('vermillion') (Renzi et al., 2001):

(71) Quel ragazzo ha gli occhi *castano-chiaro*
Due camicette *verde bottiglia*
Un maglione *nocciola*

Adjectives of foreign origin are invariable as well: *beige, blu, chic, snob, standard, tabù*; these adjectives can also be used as nouns (Renzi et al., 2001).

(72) Certe discussioni sono *tabù*.
Una ragazza *snob*.

Lastly, there exists a group of compound adjectives which are created through the combination of two adjectives, such as *agrodolce* ('bittersweet'), *audiovisivo* ('audiovisual'), *idroelettrico* ('hydroelectric'), *francoprovenzale* ('Franco-Provençal') and similar. The agreement with the noun is, in general terms, expressed only on the second element of the compound, while the former does not vary:

(73) I dialetti *francoprovenzali* / *franchiprovenzali
Una camicetta *grigio-azzurra* / *grigia-azzurra.

1.6.2 Syntax of Italian adjectives

As for their syntactic functions, adjectival phrases in Italian can be complements of particular verbs and fulfil a predicative function. We can distinguish different cases.

Firstly, when the AP goes with the verb *essere* or with other copulative verbs such as *diventare, sembrare, restare* and so on, it serves as predicate of the subject (Renzi et al., 2001).

(74) a. [[SUBJECT Gianni][PREDICATE è *fedele* a Maria]]
'John is faithful to Mary'.
b. [[SUBJECT Gianni][PREDICATE sembra *sicuro* di sé]]
'John seems self-confident'.
c. [[SUBJECT Gianni][PREDICATE resterà *fedele* a Maria]]
'John will remain faithful to Mary'.

Secondly, adjectival phrases are predicates of the object when they occur as complements of epistemic verbs, i.e., verbs of knowledge (e.g., *conoscere, sapere* ('know')), verbs of belief (e.g., *ritenere, considerare* ('think/believe')), and causative verbs (e.g., *lasciare* ('leave'), *mantenere* ('maintain')). In this case as well, the morphology of the adjective agrees with that of the noun (Renzi et al., 2001).

(75) Considero [OBJECT Gianni] [PREDICATE *fedele* a Maria].
'I consider John faithful to Mary'.

Lastly, adjectival phrases can carry out the predicative function of the subject (76a) or of the object (76b) of the sentence (Renzi et al., 2001).

(76) a. Gianni ha abbandonato Maria *arrabbiato*.
  ‘John$_j$ left Mary angry$_j$’
 b. Gianni fotografa le persone *sedute*.
  ‘ John takes pictures of people$_j$ sitting$_j$’

The adjectival phrase is compulsory when it is the complement, as shown by the unacceptability of the examples in (77):

(77) a. *Maria ritiene Gianni.
  ‘*Mary considers John’
 b. *Questo gioco rende.
  ‘*This game makes’

By contrast, when it acts as an adverb, its presence is optional (Renzi et al., 2001).

(78) a. Gianni ha abbandonato Maria.
  ‘John left Mary’
 b. Gianni fotografa le persone.
  ‘John takes pictures of people’

Finally, adjectival phrases can be complements of copulative verbs and fulfil the presentative function. The sentences that we obtain are called presentative since they aim at presenting events (such as in (79a) or at introducing in the discourse the syntactic subject (as shown by (79b).

(79) a. [PREDICATE È *sicuro* che Maria verrà alla festa].
  ‘It is sure that Mary will come to the party’
 b. [PREDICATE È *nota* solo la destinazione delle merci].

‘It is only known the destination of the goods’

In Italian, as in other Romance languages, adjectives can have three different functions: attributive, predicative or presentative (Renzi et al. 2001). When the adjectival phrase is within a noun phrase, preceding or following the noun, it carries out the attributive function, i.e., it attributes a quality to the noun it refers to.

(80) a. Un bel libro
b. Un libro *bello*
‘A beautiful book’

(81) a. un’*interessante* discussione
b. una discussione *interessante*
‘An interesting discussion’

By contrast, the adjectival phrase appears obligatory post-nominally when it includes complements, as shown by the ungrammaticality of the sentence (82b).

(82) a. un ragazzo *fedele* a Maria
b. *un *fedele* a Maria ragazzo

In general, if the adjective is used in a distinguishing or restrictive sense, it follows the noun, i.e., when the adjective denotes contrast or establishment of a difference it occurs post-nominally (Vincent, 1998) (see §1.3.4).

(83) a. le pietre *preziose*
b. le *preziose* pietre
‘The precious stones’

In (83a) the stones are contrasted to ordinary ones, whereas in (83b) their value is taken for granted (Alexiadou, 2001). More generally, there is tendency according to which whenever the adjective denotes objectivity, it typically follows the noun (D'Addio, 1974).

1.7 Summary

Chapter 1 of this thesis has explored the lexical class of adjectives under different theoretical points of view. In §1.1, we have dug into the long history of definition of adjectives as a word class and the debate about its status as universal and common to all languages. In §1.2, we have explored the difficulties linguists have found in finding criteria to categorise adjectives and why the individuation of a universal categorisation is challenging, but necessary. In §1.3, we have analysed the class of adjectives as for their syntactic distribution and the syntactic functions they can assume. In doing so, we have focused our analysis on English (the most studied and documented language) and Italian (the language of the experimental part of the present work). In §1.4, the adjectival class has been explored based on meaning and on the main semantic properties that adjectives have, namely intersectivity and gradability. In §1.5, we have presented theoretical approaches for adjective categorisation based on their pragmatic functions. Finally, in §1.6 we have deepened the main aspects concerning the adjectival phrase in Italian, mainly focusing on its morphology and syntactic distribution.

We turn to chapter 2, where we will provide the theoretical background concerning the acquisition of adjectives by young children and will summarise the main findings of previous acquisition studies in this respect.

2 Adjectives in first language acquisition

2.1 Introduction

Language acquisition is an impressive process whose ultimate attainment is what distinguishes humans from all the other species in the world. For this reason, mastering a language has been notoriously identified as the "quintessentially human trait" (Pinker, 1995).

A child learning a language is exposed to two streams of information: the stream of perceptual-cognitive data about the world around him and the stream of the language being spoken. His task, in acquiring word meaning, is to match up these two streams (Gentner, 1982). According to the *two-step model of lexical acquisition*, learning words proceeds in two phases. First, the child has to segment the speech stream into word units, as a result of which the child has a phonological lexicon of word forms (the so-called *phonological bootstrapping hypothesis*, see e.g., Gleitman & Wanner, 1982). Second, this word form must be associated with a meaning. This model does not entail that in all cases the word is first identified and later the meaning is attached to it. In other words, infants can build a prelexical representation of speech based on acoustic cues. Indeed, to identify word forms, infants do not need to already have a lexicon: they draw on other sources of information such as prosody, distributional regularities, phonotactic constraints, and typical word shapes. Thus, by 6 months, infants are sensitive to the typical word shapes of the environment language, by 8 months they develop the ability to notice distributional regularities (see Jusczyk and Hohne, 1997, Stager and Werker, 1997), and by 9 months they are sensitive to the phonetic and phonotactic properties of native-language words (see, e.g., Jusczyk et al., 1993, Saffran et al., 1999, among many others). This representation is largely language specific and is meant to facilitate the extraction of language regularities.

At 9-12 months, young children start to develop an expectation and link words to a broad range of commonalities, including category-based and property-based commonalities (Waxman & Booth, 2003). This initial expectation becomes more and more specific and, by 12-13 months, infants begin to distinguish nouns from other grammatical categories. However, children's mappings for words from other grammatical categories, including verbs and adjectives, take place several months later (e.g., Gleitman et al., 2005).

The present chapter summarises the main findings of previous studies on the acquisition of adjectives. The chapter is structured as follows. §2.2 focuses on the differences between adjective and noun classes, from the point of view of their syntactic distribution and their meaning. In §2.3, we report on different proposals accounting for the earlier and faster acquisition of category words (i.e., nouns) over other word classes, including property terms (i.e., adjectives). §2.4 will provide an overview of the observational studies conducted on children's first words, analysing the main findings and limits. In §2.5, we report on comprehension studies conducted to explore adjective acquisition. More specifically, we will focus on several aspects of the acquisition of adjectives: the ability to use syntactic context in learning adjectives (§2.5.1); the processing of nouns and adjectives in combination (§2.5.2); the different semantic classes of adjectives (§2.5.3). The chapter closes with §2.6, reporting on studies investigating adults' online processing of the class of adjectives.

2.2 Why are adjectives different from nouns?

Studies investigating children's early lexical development mainly focus on the production of children's first words (e.g., Nelson, 1973; Dromi, 1987; Clark, 1993). Although nouns, verbs and adjectives are among the first words children produce, the class of adjectives turns out to cover a comparatively smaller portion of the mental lexicon and to increase less than nouns and verbs. Children typically learn words that name concrete objects sooner and more easily than they learn words describing properties of objects. The reason for this asymmetry lies in the fact that nouns and adjectives differ both semantically and syntactically, hinting at a complex acquisition task for the language learner. The child not only has to detect the syntactic environments in which adjectives can occur but she also has to learn the lexical meaning of adjectives as well as their compositional semantics.

As far as their semantics is concerned, nouns, especially the earliest ones acquired, typically denote discrete entities, such as objects, animals and people, that are salient and tangible. Such nouns are not complex, infrequent nor abstract. In other words, children prefer nouns for cohesive whole objects, referring to basic level categories that surface as count nouns (Mintz & Gleitman, 2002). Modifiers, on the other hand, label states and properties: these are not discrete, cannot be touched, and do not speak (Gleitman et al., 2005). Furthermore, while people and objects might be referred to with some frequency in the discourse context, properties are not often labelled ostensively and, even if they are, the child has to face the demanding task of determining which of the non-discrete aspects in the environment she is supposed to pick out (Gleitman et al., 2005). In addition, common nouns label objects that are similar across inter-related properties, whereas adjectives label objects that are similar on only one property (see Markman, 1989). For example, knowing that an object is a bird allows many different inferences about its properties; on the other hand, knowing that an object belongs to the category of "white-things" triggers predictions only about its colour (Gasser & Smith, 1998). A further difference between these word classes concerns the fact that common nouns classify objects at one level: the way nouns are commonly used is explained by the "one-object, one-name" rule, even though imperfectly (see, e.g., Markman, 1989). "An object is a dog or a house or a watch or a car or a leaf. Thus, the question *What is it?* is answerable by one basic noun" (Gasser & Smith, 1998: 274). On the other hand, adjectives are mutually exclusive within a dimension, but overlap completely across dimensions. "Between-category similarity among nouns is minimal, but between-category similarity among adjectives is great" (Gasser & Smith, 1998: 276).

In addition to the semantic differences, nouns and adjectives vary in their syntactic distribution. When encountering an unfamiliar word in a sentence, listeners make use of a variety of linguistic cues to identify what kind of word it might be. Nouns can be identified on the basis of reliable morphosyntactic and positional markers, such as a determiner preceding it. Determiners such as *the*, *a*, *some*, etc. can also be followed by an adjective. However, since in child directed speech nouns are more frequent[13], children might assume that this syntactic position

[13] Thorpe and Fernald (2006) conducted a corpus analysis of adult speech to five children from the CHILDES database (MacWhinney, 2000). They searched for occurrences of the word *the*, yielding a total of 882 usages, of which 93% were followed directly by a noun.

is reserved to nouns only (Thorpe & Fernald, 2006). In contrast, linguistic cues associated with adjectives are less trustworthy. For instance, typical bound morphemes associated with adjectives like *-ish* or *-y* can also occur in nouns (e.g., *goldfish*, *tummy*). Other cues may rely on the position within a syntactic frame. Again, they might be misleading because nouns and adjectives can both appear in identical contexts. For example, when trying to understand the meaning of the pseudo-word *blane* in the English sentences in (84), (a) and (b) are ambiguous.

(84) a. That's *blane*.
b. That woman is *blane*.
c. That's a *blane* woman.
d. That's very *blane*.

Based on the syntactic position, both in the examples (84a) and (84b), *blane* could either be the name of a substance (e.g., *iron*), a proper name (e.g., *Mary*), or a property term (e.g., *big*). Only when preceding a familiar noun as in (84c) or when following qualifiers as in (84d), it is evident that *blane* must be an adjective. Indeed, the most reliable syntactic marker that a word is an adjective in English is its occurrence in prenominal position or when following an adverb of quantity (e.g., *very*, *too* or *quite*). Although these syntactic cues provide a highly reliable cue to identify the class of adjectives, they also present potential problems for the listener interpreting the speech stream as it unfolds, because the disambiguating information does not appear until after the unfamiliar word has been spoken. Given that determiners in English are more likely to be followed by nouns than by any other types of word, the best bet for the listener relying on distributional cues alone is that any novel word following *the* or *a* is a noun rather than an adjective (Thorpe & Fernald, 2006).

Besides morphosyntactic and positional cues, prosody can provide potentially useful information in identifying nouns and adjectives in continuous speech. Although there are no prosodic patterns uniquely associated with a word class. In English, for example, when an adjective preceding a noun is not given contrastive emphasis, it is typically deaccented relative to the following noun, which might convey vowel lengthening, rise in pitch followed by a fall, and increased amplitude (see, e.g., Crystal & House, 1988). For instance, in a sentence like (84c) spoken with a typical prosodic contour across the noun phrase, the stronger prominence of the familiar final word with respect to the unfamiliar penultimate word could be used by the listener as evidence that *blane* is an adjective. Furthermore, before hearing the final word, the deaccented prosody on the penultimate word might be sufficient to indicate that it is not the last word in the phrase. Indeed, several studies showed that adults use these types of cues to determine whether a word is the last word in the sentence or whether the sentence continues (e.g., Grosjean, 1983). Similarly, Thorpe and Fernald (2006) concluded that 2-year-olds use a combination of syntactic, lexical, and prosodic cues to anticipate what is coming next in the speech stream.

Although syntactic, distributional and prosodic cues are abundantly available in spoken language, children might not make use of these tools to consistently identify grammatical categories. To address this issue, Mintz and colleagues (2002) conducted a distributional analysis of early language input, investigating how the co-occurrence patterns of words in caregivers' speech could yield information about linguistic structure. They managed to categorize 200 common words in a large corpus of child-directed speech, correctly classifying the majority of nouns and verbs. However, adjectives were frequently misclassified as nouns. While lexical co-occurrences alone appear to be insufficient for distinguishing nouns from adjectives, Christiansen et al. (1998) found that the combination of these co-occurrence patterns with prosodic cues, including lexical stress and boundaries between utterances in child-directed speech, was much more effective.

In summary, adjectives differ from nouns as for the meaning they convey, their syntactic distribution and, often, the prosodic contour in spoken language. In acquiring the lexicon, different studies showed that children make use of a combination of semantic, syntactic and phonological cues, that are abundantly available in the environment. We turn now to the various hypotheses that have been proposed to account for the difficulties young children have in acquiring property words.

2.3 Why acquiring adjectives is more difficult than acquiring nouns

In acquiring a language, children face the challenge of interpreting new words under ambiguous observational conditions. However, already by twenty-four months of age, most children use adjectives in their everyday speech and seem to be aware that they refer to properties and not to objects (Mintz & Gleitman, 2002). One possible explanation for this attainment is that children have innate knowledge that labels are used to refer to things in the world and they recognize the same intention in other humans (Bruner, 1978; Macnamara, 1982). Although children naturally know that words are in general used to refer, they need to figure out what words refer to. One conjecture is that toddlers learn word meanings by means of hypothesis formation and testing procedure. According to this view, children make a hypothesis about the meaning of a word on the basis of the associative principle of temporal contiguity: young children develop the ability to recognize co-occurrences between words and references and between words and the act of pointing. Once the hypothesis is made, children test it in different contexts in which the same word is used (see Inhelder & Piaget, 1964; Bruner et al, 1966). Thus, the word-learning task involves a *Word-to-World Mapping Procedure,* that is, young children associate a word with what is perceived when the word is spoken.

There is also evidence that children can make use of non-verbal cues in learning word meaning. For instance, around 18 months of age, toddlers use the direction of the speaker's gaze to detect the focus of attention when uttering a word or sentence (Baldwin, 1991). Although this may work in many cases, this procedure does run into problems, the most evident of which is the so-called *Indeterminacy problem* (Quine, 1960). Although the ability to attend to non-verbal cues may help identifying the meaning, a given scene is open to multiple interpretations, such that the identification of the correct meaning of words is a hard task if the child can only rely on the observation of the accompanying extralinguistic context. This task is even more complex considering that the circumstances under which adjectives are uttered vary drastically; therefore, the child needs to be able to operate cross-situationally (Mintz & Gleitman, 2002). Moreover, these cues may not always be available or may not be sufficient. For example, given a situation with a dog, when hearing the word *dog*, the child might assume that it refers to the dog, or that it is its name, or that it labels one of its parts (e.g., the tail or the nose) or one of its properties (e.g., the size or the colour) and so on. However, children easily learn noun terms such as *sugar*, and they do not use them to mean *white* or *sweet*, but rather, they mistake *white* or *sweet* to mean *sugar* (Mintz & Gleitman, 2002). Thus, before children have mastered the noun/adjective contrast, both nouns and adjectives enhance infants' attention to kinds of objects (Markow & Waxman, 1995).

A first account explaining why nouns are extensively more prominent than adjectives in early lexicons is Gentner's (1982) so-called *Natural Partitions Hypothesis*. According to this view, nouns are earlier acquired since they are particularly accessible to children. Indeed, nouns are conceptually more basic than the concepts referred to by predicates such as verbs and adjectives. According to Gentner's (1982) hypothesis, the linguistic distinction between nouns and predicate terms relies on a pre-existing perceptual-conceptual distinction between concrete concepts and predicative concepts. Furthermore, the category corresponding to nouns is conceptually more basic than those corresponding to predicates (i.e., adjectives and verbs); for this reason, the particularly transparent semantics of nouns facilitates the word-to-world mapping procedure. In other terms, the child already has the object concept and only needs to match it with the novel word (Gentner, 1982). Verbs and adjectives, on the other hand, have a less transparent relation to the perceptual world, a characteristic referred to as *Relational Relativity*: verbs, adjectives, prepositions, and other relational terms are cross-linguistically variable in how they map from concept to word. In contrast, concrete and proper nouns refer to things that are perceptually coherent and easily individuated even by children. Thus, infants acquiring a language are less able to guess those meanings relying on their knowledge of the world, because "object concepts are given to us by the world and can be learned one at a time; predicate concepts [on the other hand] form a system that must be invented or, from the child's point of view, discovered" (Gentner, 1982:54). Gentner (1982) concluded that the role of the *Natural Partitioning* account prevails in the early distinction between nouns and other predicates, but its role decreases during later acquisition of distinction between classes of predicates.

A second explanation for the asymmetry between nouns and adjectives is the hypothesis proposed by O'Grady (1987), who classified lexical categories using the notion of dependency. According to this view, adjectives belong to a "dependent class"[14], since they express a function that applies to another element in the sentence. More specifically, adjectives are always combined with an independent nominal category. To make this distinction clearer, O'Grady pointed out that nouns are typically used as arguments or *primaries*, referring to entities or classes of entities; while verbs and adjectives, or *secondaries*, are predicates that depend on *primaries*[15] and express something about them. O'Grady (1987) suggested that children are able to use *secondaries* (e.g., the adjective *big*) only if they have some nominal argument in mind (e.g., "the thing that is big").

A similar account for the later acquisition of adjectives came from Mintz and Gleitman (2002), who argued that the conflation of a property under a lexical category is more arbitrary than natural. Differently from properties, which can be referred to by different lexical categories, whole concrete objects fall under the noun category cross-linguistically. Besides, also in languages like English or Italian with an open class of adjectives, there are properties that do not fall under the adjective category. For example, the word *triangle* is a noun, but it means "something triangle-shaped". Moreover, adjectives do not always describe properties as in the case of English *mere, similar, former,* and so on (Mintz & Gleitman, 2002). Additionally, according to Mintz and Gleitman (2002), adjective learning is parasitic on and subsequent to the acquisition of a stock of nouns, since the information on the adjective concerns the noun it modifies, and the interpretation of the adjective is conditioned by the noun with which it co-occurs. Compare, for example, the meaning of *good* in the sentences in (85) (examples from Katz, 1964):

---

[14] According to O'Grady (1987) there are two classes of lexical categories: a dependent class consisting of verbs, adjectives, and adverbs and an independent class consisting of nouns.

[15] "Adjective: A secondary dependent on either an N or an NP (and not compatible with tense and aspect)" (O'Grady, 1982, p. 13).

(85) a. a good knife
b. a good wife
c. a good pipe
d. a good life

According to Pustejovsky (1995), the meaning of all lexical items, including adjectives, is relational in different degrees. Word meaning (expressed in terms of predicative structure, or *qualia*) is made up of four essential aspects, which he incorporates in the "Qualia Structure":

CONSTITUTIVE: the relationship between an object and its parts (material, weight, …)
FORMAL: what distinguishes the object within a larger domain (shape, colour, …)
TELIC: the object's purpose and function (aim of an activity, purpose in performing an act…)
AGENTIVE: the factors involved in the object's origin (creator, natural kind, …)

As Pustejovsky (1995: 86) put it, "[t]he qualia provide the structures template over which semantic transformations may apply to alter the denotation of a lexical item or phrase". These transformations, such as coercion and co-composition, map the extension of a meaning to new meanings and apply in relation to the syntactic and semantic environment in which the phrase appears. Thus, qualia are interesting "since they not only structure our knowledge of words, but also 'suggest' interpretations of words in context" (Pustejovsky, 1995: 87). Consider, for example, how the nouns in (85) contribute towards the interpretation of the adjective *good* in the phrase (85a) - (85d) above. In (85a), the knowledge of knifes allows us to infer that a good one is sharp and cuts smoothly, hence permitting an ellipsis of its particular properties. Likewise, in (85b), the adjective *good* refers to a wife's faithfulness and patience. In (85c), the interpretation of *good* is made possible by our knowledge of what pipe smokers may look for, a durable pipe, for example. Lastly, in (85d) the "sense in context" of *good* could be referred to happiness, to a healthy lifestyle and so on. To conclude, two are the factors used to determine the appropriate sense in each case: the qualia structure of the lexical items, and co-composition, that is, qualia unification resulting in a specific meaning in context, without the need to list all individual senses (Pustojevsky, 1995). Although all adjectives depend on the nouns they modify to a degree, different semantic subclasses of adjectives vary in how they combine with nouns and how much context is necessary for interpretation (e.g., Partee, 1995), leading to different patterns of acquisition that will be discussed later in this chapter.

An influential approach is the one proposed by Markman (1990). She observed that, although syntactic cues may help children in the identification of word classes, constraints on word meaning may be more critical for babies who have not yet learned enough syntax to rely on grammatical structure to limit their hypotheses as to a word's meaning. She argued that, by the time they are ready to acquire vocabulary, children place constraints on possible word meanings, thereby greatly reducing the hypotheses that need to be considered. Infants can quickly zoom in on some potential meanings that they are predisposed to prefer thanks to an *a priori*, built-in knowledge. According to this view, children are endowed with cognitive constraints that help them narrowing down the number of hypotheses they could formulate whenever they are confronted with a novel word and have to assign it a meaning (Markman, 1990: 58).

The first assumption helping children constraining the meaning of words is labelled as the *whole object* assumption: “A novel label is likely to refer to the whole object and not to its parts, substance or other properties” (Markman, 1990: 59). Let’s consider the dog example once again. Whenever an adult points to an object and labels it by saying *dog*, in a situation in which there is a dog that is wagging its tail, children should take into consideration all the possible meanings that the word could convey, namely “dog”, but also “tail”, “hairy object” and so on. However, this first assumption guides them towards the whole-object hypothesis, that is, to assume that *dog* refers to the animal.

Once children have decided that a novel word refers to the whole object and not to its parts, substance, or other properties, they still have to decide how to extend it to other objects; in this case the second assumption, i.e., the *taxonomic*, plays a fundamental role. According to this bias, labels refer to objects of the same kind rather than to objects which are semantically related. Indeed, at a broader cognitive level, it has been independently proved that children are more sensitive to thematic relations than to categorical or taxonomic ones. For instance, children are more sensitive to the relation between a dog and its bone, than to the relation between cats and dogs. Nevertheless, when an adult points to a dog and calls it *dog*, the child is biased to reject possible meanings such as "the dog and its bone," or "Mum petting the dog" or "the dog under the tree". As Di Sciullo and Williams (1987) point out, "words are generic in meaning in a way that phrases are not"; thus, even though children may find thematic relations more salient, languages rarely encode them in single words (Markman, 1990).

In early word learning children are also guided by the third assumption, which favours *basic-level* representations. In other words, children prefer objects that are identified at a basic level (Mintz & Gleitman, 2002). Indeed, children often extend *Bud* to name all dogs, but they do not go any further, i.e., they never use it to refer to animate things or to include all animals. What is more, as Clark (2003) observes, "Conceptually, basic-level categories are privileged: they cohere internally, so their members are readily perceived both as members of their category and as distinct from nearby categories. The basic level is also privileged linguistically: the terms assigned to basic-level categories tend to be simpler in form than those assigned to lower levels and so are easier to learn and to remember" (Clark, 2003: 125).

These three assumptions can explain the predominance of nouns referring to concrete objects in children's early vocabulary. However, the child needs a fourth assumption that helps her override the three previous mentioned constraints, thereby enabling her to acquire terms for events and properties, i.e., the *mutual exclusivity* assumption. According to this bias, "Words are mutually exclusive and each object will have one and only one label" (Markman, 1990). The child, therefore, initially assumes that words are mutually exclusive, i.e., there are no synonyms, and each object has one and only one label. For instance, when the child already knows the word *dog* and the novel-word *brown* is mentioned (referred to the dog itself), the mutual exclusivity bias guides him in excluding the meaning "dog" for the novel word and helps him assuming that it refers to something else, e.g., one of the dog's properties (Panzeri, 2011).

To sum up, in learning the lexicon children have to deal with ambiguous observational conditions and figure out what novel words in the input refer to. When they hear a new word, young children seem to make hypotheses about their possible meaning and need to be able to operate cross-situationally to test them in different contexts. In order to do so, in the first phases of language learning, children seem to be guided by innate cognitive constraints, which favour the acquisition of nouns over other lexical categories, including adjectives. Thereby, children learn category labels earlier and faster than property terms. We turn now to the studies that have observed this asymmetry cross-linguistically in children's early vocabularies.

2.4 Asymmetry between nouns and adjectives in the early vocabulary

The priority of nouns over other word classes has been observed in many corpus analyses and elicited production studies. These studies showed that, in early child vocabularies, nouns are massively favoured in comparison to adjectives and verbs both in comprehension and in production (Gentner, 1982; Gleitman, 1990; Caselli et al., 1995; Gleitman et al., 2005), and this observation holds cross-linguistically (Gasser & Smith, 1998).

Different observational studies have tried to categorize children's first words into parts of speech. The main problem in doing so is linked to the fact that, in children's single-word utterances, the same word can be used for different purposes. For example, an adjective such as *hot* can refer to the property of having a high temperature but can be used as a word for "oven" or "Be careful! Don't touch it!" (Clark, 1993). Thus, various criteria for the classification of early words have been proposed. Adjectives, in particular, have been classified in various ways. Some studies merged adjectives and verbs into the category of predicates (see, e.g., Gentner (1982)'s proposal in §2.3 above). In other approaches, adjectives are classified as modifiers (e.g., Nelson, 1973; Dromi, 1987).

Nelson defined modifiers as "words that refer to properties or qualities of things or events" (1973: 17). This definition holds for attributes, states, locatives and possessives. In order to investigate the early lexicon with respect to the use of modifiers, Nelson (1973) analysed the first 50 words of 18 English-speaking children until the age of 25 months. She found that all word classes were among children's first 50 words, although not equally distributed. Across children, the largest class were nominals (65%) followed by action words (13%) and modifiers (9%). Interestingly, the proportion of modifiers did not seem to increase over time, contrarily to what she observed with nominals.

Similarly, Dromi (1987: 93) defined modifiers as "words that refer to properties and qualities of things and events". Examining the speech of one Hebrew-speaking child during the one-word phase between 10 and 17 months of age, Dromi (1987) found that modifiers constituted only 4% of the produced words, corresponding to 11 words, 8 of which were analysed as the adult part-of-speech adjective. In addition, Dromi (1987) observed that the first adjective was produced at 16 months and that the child often used adjectives with various reference. For example, the word *hot* was used for "being hot" or for "heater" or "oven", thereby confirming the hypothesis of the priority of objects over properties.

In analyzing the spontaneous speech of 32-German speaking children between 13 and 36 months of age, Kauschke (2000) created a list of produced words in the target-like form (that is, without morphological marking) rather than the children's specific realization of the word. Similar to Nelson (1973), Kauschke (2000) found that, while nouns and verbs increased with age, the proportion of adjectives increased significantly less: from 13 to 36 months (3,6% for types, 1% for tokens).

Contrary to Kauschke (2000), Tribushinina et al. (2013) found a significant increase in the adjective token frequency of Dutch speaking-children between age 2 and 3. However, the difference in the results may be attributed to the cross-sectional data analysis conducted by Tribushinina et al. (2013), in contrast to Kauschke (2000)'s longitudinal study.

Carrying out an analysis using adults' parts of speech, Caselli and her colleagues (1995) analysed the early vocabulary of hundreds of Italian- and English-speaking children between 8 and 16 months of age. The results showed that there are similarities in the acquisition of these two different languages, with a dominance of nouns and a virtual absence of verbs and adjectives up to when vocabulary exceeds 100 words.

Very detailed analyses of adjectives were carried out by Blackwell (2005) in a longitudinal study with two English-speaking children between the age of 2;3 and 4;11 years of age. Tribushinina et al. (2014) extended this line of research to other languages with an open adjective class (Dutch, German, French, Hebrew, Turkish) and to an earlier time window (1;8 to 2;8 years). Both studies employed semantic adjective classes to define the concepts adjectives express. They both found that physical property, colour, spatial property, evaluative and conformity adjectives were among the early occurring classes and emerged before internal state, behaviour, temporal and modal adjectives, and ordinal numbers. Moreover, their findings imply that adjectives are acquired in an orderly fashion. According to Blackwell (2005), adjectives that are independent of the noun meaning (the so-called "categorematic adjectives") may be used earlier and more frequently than adjectives that are dependent on the noun meaning (the so-called "syncategorematic adjectives"). However, dimensional adjectives (such as *big* or *long*) do not fit this explanation: they are acquired early although their interpretation is dependent on the noun. In contrast, Tribushinina et al. (2014) concluded that adjectives with so-called "concrete semantics", that is, adjectives that refer to immediately perceptible concepts (such as colour, size, physical properties), emerge earlier and are more frequent than adjectives referring to less accessible abstract properties (such as behavioural or temporal adjectives). Furthermore, Tribushinina et al. (2014) pointed out that the occurrence of adjectives in child-directed speech is relevant: adjectives that emerge early are highly frequent in the input and important for child-parent interactions. On the contrary, Blackwell (2005) claimed that input properties can account only partially for the variance in the age of acquisition for the different adjectives.

The observation that nouns are highly favoured with respect to adjectives in children's early vocabularies holds cross-linguistically and adjective acquisition seems to be dependent on adjective meaning. In this respect, it is also relevant to consider in which syntactic frames adjectives occur in the input and whether different structures can help children infer adjectival meaning.

A few studies have investigated the development of attribute and predicative uses of adjectives in production (e.g., Nelson, 1976, Menyuk et al., 1995, Becker, 2000, Saylor, 2000, and Blackwell, 2001, for English; Bittner, 1998, Clahsen et al., 1994, and Eisenbeiss, 2000, for German; Tribushinina et al., 2013, for Dutch). These studies suggest that, during the third year of life, noun phrases start to become more complex. More specifically, children start to use both a determiner and an adjective instead of either of them (Bittner, 1998). Moreover, from two years of age, copular verbs are realised in utterances containing adjectives (Becker, 2000) and, already at age 2, both attributive and predicative structures are attested, but attributive adjectives are more frequently used than predicative adjectives (Tribushinina, et al., 2013). With respect to the distribution of attributive and predicative use, previous results are inconclusive for multiple reasons. First of all, it is difficult to categorize the produced structure as either predicative or attributive, because children tend to reduce syntactic complexity by omitting some elements (e.g., the copula or the determiner). Thus, their categorisation is more often speculative and based on adults' categories (Ninio, 2004). Moreover, children's meaning may be different from adults', even when the form is the same. Lastly, if what children actually say mirrors their knowledge or interest cannot be known; that is, looking at production may be misleading for a functional analysis (Nelson, 1976). Indeed, it is impossible to evaluate negative evidence: "if the child does not say something, we cannot conclude that he could not" (Nelson, 1976: 15).

To conclude, various studies have tried to describe the acquisition of adjectives by observing children's early production. All studies have reported an asymmetry in the first vocabulary between nouns and adjectives, with the first being massively favoured cross-linguistically. Nevertheless, the problem of categorisation of children's first words and syntactic uses makes it clear that a more effective way to assess children's language skills is to focus on comprehension.

In the next section we will review the relevant literature on adjectival acquisition, focusing on the comprehension studies that constitute the groundwork for the present thesis. More specifically, we will report on studies investigating the word-to-world mapping problem adjectives raise, the difficulties in the interpretation of noun-adjective combinations and the influence different semantic classes of adjectives exert on real-time processing.

2.5 Comprehension studies on adjective acquisition

Research on the acquisition of word classes has primarily focused on the development of nouns and verbs, whereas the acquisition of adjectives has received comparatively little attention. Indeed, it has been difficult to demonstrate adjective learning experimentally with children as young as two years old. However, at this age most children use a number of adjectives in their everyday speech and appear to understand that, in so doing, they are referring to properties of things.

The studies conducted in the last few decades confirmed the observation that there is a disparity between young children's ease in learning novel object names and their difficulty in learning new property terms. Indeed, in line with the observational studies mentioned in §2.4, children are more likely to interpret novel words as referring to categories and not to properties. Moreover, the extension of novel adjectives happens more slowly and seems to be determined by different variables (Gasser & Smith, 1998).

In the present section we will report on the literature on comprehension studies concerning adjective acquisition. More specifically, we will provide an overview of the studies analysing the use of syntactic context in learning adjectives (§2.5.1); studies investigating whether toddlers are able to interpret nouns and adjectives in combination and whether they use the adjective to establish reference (§2.5.2). Furthermore, we will report on studies that investigated the interpretation of semantically different classes of adjectives by children (§2.5.3) and adults (§2.5.4).

2.5.1 The use of syntactic context in learning adjectives

The studies on adjective acquisition have mainly addressed the word-to-concept mapping problem raised by adjectives, concentrating on the identification of the circumstances under which children are able to extend a novel adjective to a new context (see, among others, Gelman & Markman, 1984; Waxman & Markov, 1998; Hall & Graham, 1999; Klibanoff & Waxman, 1999; Waxman & Booth, 2001). These studies aimed at testing the so-called *Syntactic Bootstrapping Hypothesis* (Gleitman, 1990), according to which children are able to use the observed syntactic structures as evidence for deducing the meanings. "The learner observes the real-world situation but also observes the structures in which various words appear in the speech of the caretakers. Such an approach can succeed because, if the syntactic structures are truly correlated with the meanings, the range of structures will be informative for deducing which word goes with which concept. This *Sentence-to-World Mapping* will be quite handy if, as I have argued, *Word-to-World Mapping*[16] cannot succeed over the full range of meanings that we know are acquired" (Gletiman, 1990: 30).

The typical procedure in this type of study is for the child to be presented with an object with a special property such as a nubbly texture described by a pseudo-word (i.e., non-existent but phonologically possible in the target language). That word could be recognised as belonging to the category of nouns or adjectives on the basis of morpho-syntactic cues. For instance, descriptions such as *This is a(n) N* (e.g., "This is a *dax*") were given for nouns and *This is a(n) A one* (e.g., "This is a *dax* one") were given for adjectives. Children were asked to find either "another *dax*" or "another *dax* one" (see, e.g., Klibanoff & Waxman (2000) for English; Waxman et al. (1997) for a comparison between English, French and Spanish; Waxman & Guasti (2009) for Italian). The presence of the nonce word within different syntactic structures should trigger the two different interpretations.

[16] The *World-to-Word Mapping* procedure (discussed in §2.3) is often referred to as the *Semantic Bootstrapping Hypothesis*, devised by Grimshaw (1981) and considerably elaborated by Pinker (1984, 1987). "The two bootstrapping proposals [syntactic and semantic] are much alike in what they claim about correspondences between syntax and semantics, and they are also alike in proposing that the child makes significant use of these correspondences" (Gleitman, 1990:27). "The difference between *semantic bootstrapping* and *syntactic bootstrapping*, then, is that the former procedure deduces the structures from the word meanings that are antecedently acquired from the observation of events, while the latter procedure deduces the word meanings from the semantically relevant syntactic structures associated with a verb in input utterances. Note that although the hypothesized procedures are distinct, to hold that one of them is implicated in learning is not to deny that the other one is too. Quite the contrary. It is very likely that they operate in a complementary fashion." (Gleitman, 1990:30).

If the novel word in the phrase "a *dax* one" is interpreted as an adjective, the child should choose the property-match over the kind-match. While children under 3-years-old reliably interpret "This is a *dax"* as referring to a category name, they fail to interpret "This is a *dax* one" as referring to a property, extending the novel word to the object that matches in kind rather than in texture (see, among others, Taylor & Gelman, 1988; Waxman & Kosowski, 1990; Waxman & Markow, 1998; Klibanoff & Waxman, 2000). In other words, these studies showed that pre-schoolers tend to interpret novel words as labels for object categories (i.e., nouns), rather than as properties of objects (i.e., as labels for adjectives). Thus, despite the presence of morpho-syntactic cues, children (at least up to three years of age) need further clues to overcome the noun preference (Mintz & Gleitman, 1998). Although consistent with the observation that adjectives are generally harder to learn than nouns, the magnitude of these negative findings are somewhat surprising given that by 30 months of age children are typically speaking up to fifty descriptive words (Dale & Fenson, 1996).

Presenting children with a novel word that is framed syntactically as either a noun (*a dax*) or an adjective (*a dax one*) is used to test whether children can use syntactic cues to figure out the meaning of unfamiliar words. This question arises from the assumption that young children have two capacities. First, that they are able to make use of surface structure cues, including word morphology and syntactic position, in the identification of novel words as either nouns or adjectives. Second, that these children have abstract grammatical categories for both word classes, i.e., for both nouns and adjectives, such that categorising the novel word leads to inferences about its meaning as object name or property label. However, other studies found that children are more likely to map a novel adjective onto an object property if the object is familiar; in other words, children use of syntactic cues varies as a function of the familiarity with the object kind (Hall et al., 1993). Findings like these reveal that preschool-aged children are sensitive to syntactic distinctions (e.g., count noun, adjective) and expect that the syntactic form of a novel word, applied to an individual, provides information regarding its meaning.

To strengthen the possibility to detect the emerging relationship between adjectives and object properties, Waxman and Markov (1998) selected a group of twenty-one-month-old children. At this age, most of them have already acquired and are able to produce at least 50 words, are acquainted with words describing object properties and have begun to produce a few multi-word utterances. Children were presented with three objects from the same basic level category, even though not identical (e.g., cars); one differed from the others for a property (e.g., two yellow cars and one green car). Children participated in either the Novel-adjective condition or the No-word condition. In the Novel-adjective condition, children were shown one of the two yellow cars and were told *This is a(n) X one, can you find another X one?* (e.g., "This is a *dax* one, can you find another *dax* one?"). In the No-word condition, children were not introduced to novel words, but heard the question *Look at this one, can you find another one like this?*. The results showed that children exhibit a preference for the matching property in both the Novel-adjective and the No-word condition, that is, they were able to pick another object with the same property. Hence, the results confirmed children's emergent ability to map a novel adjective to *within-category* familiar objects. However, when asked to extend the novel adjective to another basic-level category (e.g., a plane), children failed to reveal a preference for the matching property (Waxman & Markov, 1998).

The same results were found by Klibanoff and Waxman (2000). 3- and 4-year-old children were presented with a forced-choice task involving objects from basic-level categories (e.g., cups) and were either exposed to a novel word (Novel-adjective group) or not (No-word group). In addition, children were split into two experimental conditions, namely the within-basic condition and the across-basic condition. In the within-basic condition, the two objects were members of the same basic-level object category (e.g., a yellow cup and a blue cup). Here the two target objects varied in a single novel dimension, therefore the process of comparison and identification of the property should be facilitated. On the contrary, in the across-basic condition, the objects varied in several ways since they belonged to different basic-level categories (e.g., a yellow cup and a blue bottle). The process of comparison alone could not support the identification of a unique property and, consequently, did not facilitate the choice. The results showed that 4-year-olds successfully extended novel adjectives from the target to the matching test object whether these objects were drawn from the same, or different, basic level categories. In contrast, 3-year-olds' extensions were more limited. Indeed, children in the within-basic condition identified the target property, were able to map it to the novel adjective and to extend it to the properties of objects belonging to different basic-level categories. Instead, if the initial comparison involved members of different basic level categories (i.e., in the across-basic condition), or if the target property was not named (i.e., in both No-word conditions), children failed in the identification of the target property. They concluded that basic level object categories serve as an initial foundation in the process of mapping novel adjectives to object properties (Klibanoff & Waxman, 2000).

Mintz and Gleitman (2002) exposed 24- and 36-month-old children to novel adjectives, providing syntactic and referential information about the adjective meaning; in other words, they used the novel word (e.g., *zavish*) to describe various objects (e.g., a cube, a ball and an elephant) sharing the salient property within full noun-phrases (i.e., not in conjunction with pronouns). Under this condition, children were able to identify the correct reference. However, when using a general and vague reference (like "one" or "thing"), both groups failed. Mintz and Gleitman (2002) suggested that, in order to identify the relevant property, young word-learners require access to the taxonomy of the object type. These findings favour an account of lexical acquisition in which both the syntactic and the referential information are required in word learning.

Although empirical support is quite robust in this respect, it derived almost exclusively from English-speaking preschool-aged children. In order to test this *noun-category* bias[17] cross-linguistically, Waxman et al. (1997) developed a study from the observation that English, French and Spanish offer and interesting set of cross-linguistic comparisons, primarily concerning constructions with adjectives. Speakers of these three languages may distinguish objects with noun phrases that include a determiner, an adjective, and an overt noun (e.g., *the blue cup, a blue cup, the blue one*). But under certain circumstances, the noun (either lexically specific,

[17] This phenomenon has been described by Waxman et al. (1997) as the *noun-category* bias, rather than the *taxonomic* bias (e.g., Markman, 1990; Markman, 1994) to highlight the fact that the tendency to focus on categories of objects is evident specifically with novel words that are presented as count nouns.

as *cup*, or generic, as *thing* or *one*) can be omitted from the surface of the sentence (e.g., *the wealthy*). As a result of this noun-dropping, the surface structure of the noun phrase only includes a determiner and adjective, but no overt noun, as in Spanish *un azul*, literally "a blue", corresponding to the English *a blue one*. These constructions are possible to some extent in all three languages, what differs is the prevalence of these structures in each language. In Spanish, where both adjectives and articles are marked for number and gender, determiner-adjective constructions (*det-A*[18]) are ubiquitous and such constructions are produced by two years of age (MacWhinney & Snow, 1990; Snyder, Senghas, & Inman, 2001). In contrast, such constructions in French and English are less frequent. As a consequence, in languages like Spanish, *det-A* constructions frequently overlap with the surface constructions in which nouns appear. Whether children acquiring such languages may have more difficulties in distinguishing between nouns and verbs has been the focus of recent investigations with Spanish or English monolingual. Waxman and Markow (1998) found that, by 21-23 months of age, infants acquiring English successfully extend adjectives, but not nouns, on the basis of property-based commonalities (e.g., color, texture). Waxman and Weisleder (2007) found that infants acquiring Spanish reveal this link between adjectives and object properties between 23-29 months.

Based on these differences, Waxman et al. (1997) tested 87 monolingual French-speaking children (ranging from 2,4 to 5;0) and 45 monolingual Spanish-speaking children (ranging from 3;1 to 4;9). Children were assigned to one of three experimental conditions. In the Novel-Noun conditions, novel words were presented as count nouns (e.g., *See this fopin? Can you find another fopin?*). In the Novel-Adjective condition, novel words were presented as adjectives (e.g., *See this fopish one? Can you show me another one that is fopish?*). In the No-Word condition, no novel word was introduced (e.g., *See this? Can you find another one?*). They found that French- and Spanish-speaking children in the Novel-Noun condition, consistently with English learners, extended count nouns to other category members, and concluded that the mapping between count nouns and object

[18] We use this terminology following Waxman and Guasti (2009).

categories may be a universal phenomenon. However, the extension of novel adjectives in the Novel-Adjective condition varied cross-linguistically. French-speakers, like English-speakers, performed at chance, i.e., they did not extend novel adjectives to other members of the same category. These results suggested that, for them, adjectives were linked to neither category-based nor thematic relations. In contrast, Spanish-speakers selected the category-based alternatives in the Novel-Adjective condition, as they had in the Novel-Noun condition, suggesting that the mappings between adjectives and their associated words vary across languages. More interestingly, Spanish-speaking children's extension of novel adjectives to the category-based alternative help up when the adjective was both presented in combination with a noun (e.g., *la cosa pequena*, "the thing little") and in det-A phrases in which the noun was dropped from the surface of the phrase (e.g., *la pequeña*, "the little (one)"). Thus, also in the unambiguously adjectival context, i.e., in conjunction with a noun, Spanish-learners consistently selected the category-based alternative. Waxman et al. (1997) concluded that experience with different languages leads children to establish different tacit expectations regarding the use of adjectives.

To understand whether the pattern exhibited by Spanish-learners is spurious or whether it is indeed tied to the grammatical features of the Spanish language, Waxman and Guasti (2009) adapted and replicated Waxman et al. (1997)'s study with children acquiring Italian. They argued that Italian is an ideal test case because, like Spanish, it has a rich grammatical gender and number agreement, the agreement markings are overt, and, perhaps as a consequence, *det-A* constructions are ubiquitous and frequent in child-directed speech[19]. Results showed that Italian-learning children favoured category-based alternatives both when presented with nouns and with adjectives, confirming Waxman et al. (1997)'s results on Spanish. These findings are consistent with the claim that children's expectations for novel

[19] Waxman and Guasti (2009) conducted analyzed eight corpora from the CHILDES database (MacWhinney, 2000), in which Italian-speaking caregivers interacted with children from 16 to 40 months. They found that, within the 974 utterances out of 3054 where adjectives were used in conjunction with a determiner (either with noun-dropping or within a noun-phrase), *det-A* constructions constituted 27% of the utterances.

adjectives reflect the behaviour of adjectives in the ambient language. This outcome suggests that, although there are strong universals in the treatment of adjectives (and nouns) across languages, there are also cross-linguistic differences in the treatment of adjectives. These differences might seem subtle on the surface, but clearly have developmental consequences.

In conclusion, several studies showed that children under the age of three make use of syntactic context in learning the meaning of novel adjectives. The ability to extend newly learnt words to new contexts seems to appear later in development, i.e., over the fourth year of age. Children show better performances when dealing with basic-level and familiar objects. A cross-linguistic comparisons also highlighted that, in learning a language, the target language plays a major role in adjective acquisition. Children learning languages like Spanish or Italian, where constructions compatible with both nouns and adjectives are frequent in child-directed speech, show different developmental behaviour than children learning languages where these constructions are infrequent.

The studies presented in this section have addressed the word-to-concept mapping concerning adjectives. We turn now to the studies that tackled the processing and interpretation of adjectives and nouns in conjunction and the problem of integration such combinations raise.

2.5.2 Interpreting nouns and adjectives in combination

The studies reported so far focused on the question of whether children are able to use the syntactic context to infer the meaning of novel adjectives in attributive and predicative uses. Such findings are interpreted as evidence that the knowledge of the described object can provide support for the child who is trying to interpret a novel adjective, helping in the focalisation of the relevant property. However, in order for children to master the task of interpreting noun-adjective combinations, the language learner not only has to map a single word to the relevant property; this process also involves the integration of the noun and the adjective meanings in order to identify a possible referent (Fernald et al., 2010).

In developmental studies, children's understanding of nouns and adjectives in combination has rarely been framed in terms of problems at the level of integration, even though, as pointed out by Nelson (1976: 14), “they enable the child to make distinctions among referent objects and classes of objects on the basis of both general and specific properties. That is, they provide him with a linguistic means to generate new reference classes”.

In this section, we will report a detailed summary of the studies investigating noun-adjective combinations that are mainly relevant for the present work, i.e., Ninio (2004), on the offline processing noun-adjective combinations (§2.5.2.1); Thorpe et al. (2006), on the offline processing adjective-noun combinations. (§2.5.2.2); Fernald et al. (2010), on the online[20] comprehension of informative vs. uninformative prenominal color-adjectives (§2.5.2.3); Weisleder and Fernald (2009), on the online interpretation of informative vs. uninformative nouns preceding color-adjectives (§2.5.2.4); Tribushinina and Mak (2016) on children’s online interpretation of informative vs. uninformative prenominal adjectives (§2.5.2.5).

2.5.2.1 Ninio (2004)

To investigate the comprehension of attributive adjective structures in Hebrew-speaking children (between 1;6 and 4;4 years), Ninio (2004) developed an off-line picture-pointing task asking children to point at a picture described by a noun-adjective combination. Hebrew presents a large open class of adjectives, some derived from verb roots and many from nouns, e.g., *rashi*, ‘chief’ from *rosh*, ‘head’. As in Italian and English, Hebrew adjectives can have a predicative and an attributive function. Like Italian, but unlike English, in both cases the adjectives agree in gender and number with the head-noun. Moreover, the attributive adjective is post-nominal. For instance, the phrase ‘big teddy’ translates to the Hebrew *dubi gadol*, *dubi* is ‘teddy’, whereas *gadol* is ‘big’. Although word order is quite flexible in Hebrew, the relative order of noun and adjective within a NP is quite rigid and its change results in an ungrammatical phrase. Furthermore, Hebrew does not require a copula in the present tense, and the unmarked order of the predicative

[20] We use the term *online* here to refer to Fernald et al. (2010)’s Looking-While-Listening paradigm. This methodology, together with its status as an online method, will be discussed in detail in chapter 3.

construction is the same as in the attributive one, resulting in an almost identical word combination. The only difference is in the articles: in the attributive construction, when the referent is definite, both noun and adjective get the pro-clitic article *ha-* as in *ha-dubi ha-gadol* ‘the big teddy’, whereas in the predicative construction the subject-noun but not the predicate adjective get the article: *ha-dubi gadol* ‘the teddy’ [is] big’. In Ninio’s (2004) experiment, children were presented with a cardboard showing four coloured pictures of familiar objects with varying natural attributes (such as big and small teddies). Each foursome presented two objects crossed with two attributes (e.g., a big teddy, a small teddy, a big clock and a small clock). The experimenter described one of the four pictures using a noun-adjective combination, e.g., *Give me the Noun-Adjective* (e.g., *Give me the big teddy*) and children were asked to point to the picture matching the description from among the four alternatives.

This study was based on the observation that attribution has a complex logical structure. In order for noun-adjective combinations to be processed, the noun must be interpreted as a category of objects; afterwards, the adjective must be interpreted as a subset of the object-set that also possess specific properties. We will refer to this proposal as the ***Noun-Anchor Hypothesis***[21]. Thus, Ninio (2004) raised the possibility that children have a problem in mastering the attributive relation between nouns and adjectives because it requires a complex process involving a two-step logical semantic integration.

Results showed that children under the age of 4 have difficulties in comprehending adjective-noun combinations. Even when tested with basic properties referred to familiar everyday objects, in 41% of the trials children did not select the correct picture. When faced with an attributed noun, their most

[21] In the present work we will adopt the label *Noun-Anchor Hypothesis*, introduced by Weisleder & Fernald (2009) on the basis of Ninio’s two-step proposal. What the *Noun-Anchor Hypothesis* predicts “is that due to the asymmetry between nouns and adjectives, noun-adjective phrases are always easier to interpret when the noun comes first. In this case, because the noun must be interpreted before the adjective, [learners of a language with noun-adjective order] have an overall advantage: they can interpret the noun phrases sequentially regardless of whether the first or second word is informative.” (Weisleder & Fernald, 2009: 4).

frequent error was to select the object from the correct category, but with an incorrect property (i.e., when asked for “the big teddy”, they selected the small teddy). Ninio (2004) interpreted these results as mirroring the problem children have with the two-step procedure: indeed, when making a mistake, children focused on the noun, omitting the second step, that is, they were unable to go through the second step. Moreover, the most common mistake (i.e., selecting the *correct category – wrong property*) was subject to spontaneous self-corrections, and most children eventually picked the correct referent. This was interpreted as reflecting the fact that children do have the adjective information, but they are not able to integrate it simultaneously with the noun.

From Ninio’s results, three questions arise. First, her research could not account for cross-linguistic variation concerning adjective position in attributive use. Indeed, children acquiring a language with adjective-noun order could show a different pattern of behavior. Studies addressing this issue will be discussed in §2.5.2.2, §2.5.2.3, §2.5.2.4, and §2.5.2.5.

Second, the adjectives used in the experiment had different semantic properties, but these differences have not been taken into account, although they might affect the processing of noun-adjective combinations in different ways. Studies exploring the interpretation of different semantic classes of adjectives in acquisition will be reported in §2.5.3.

Third, offline tasks can only provide measurements of the accuracy, which comes *at the end* of the process of interpretation. Online processing studies can instead account for what happens *during* the integration. This methodological issue will be discussed in detail in chapter 3. In general, online studies investigating adjective acquisition are scarce, and the acquisition of different semantic classes of adjective has not been investigated based on online tasks, so far.

2.5.2.2 Thorpe et al. (2006)

Ninio’s (2004) *Noun-Anchor Hypothesis* was tested on Hebrew, a language that, like Italian, has post-nominal adjectives. Thus, the sequential steps of integration in her proposal are aligned with the order in which words are presented in Hebrew. However, this two-step process might be more challenging for, e.g., English-speaking children. Since in English the order is the opposite of, e.g., Hebrew or Italian (i.e., adjective-noun), processing these combinations might be even more demanding and could entail an increase in working memory load. Indeed, in interpreting noun phrases containing adjectives, English-learners would have to wait until the noun to perform the first step of noun identification while holding the prenominal adjective in memory. Then, after the interpretation of the noun, they would have to retrieve the adjective from working memory, integrate it with the noun and identify the appropriate referent. If this was the case, the task designed by Ninio (2004) would be more difficult for young English-speaking children than for Hebrew learners. A first attempt to address this issue came from Thorpe and colleagues (2006), who conducted a replication of Ninio's (2004) study on English, using the same off-line picture-pointing task. Three age groups were tested, namely 1;6, 2;6 and 3;6 years old, spanning the same age range as Ninio.

Similar to Hebrew-speaking children, English learners identified the correct referent 60% of the times, although accuracy varied dramatically by age. Indeed, children older than 36 months were close to ceiling on this task, but younger children showed great difficulty with adjectives. Only children in the 2;6 years group showed evidence for the noun-biased error, i.e., picking the *correct category – wrong property* referent, observed with Hebrew-speaking children. This error pattern was interpreted by Ninio (2004) as evidence that children process noun phrases containing adjectives by first determining the noun category, but then failing to complete the identification of the correct picture based on the adjective interpretation and integration with the noun. However, the picture-pointing task used in both studies can only provide information about the final interpretation, since only accuracy can be measured and analyzed. As Thorpe et al. (2006) observed, it is impossible to tell whether the English data were the result of processing the noun first and then leaving off the adjective, or whether the errors were the result of being more familiar with the noun or remembering it better than the preceding adjectives. Fernald et al. (2010)'s study has been developed to tackle this issue.

2.5.2.3 Fernald et al. (2010)

Fernald et al. (2010) developed a looking-while-listening paradigm, in which 30- and 36-month-old English-speaking children looked at pairs of pictures while listening to speech stimuli. A later frame-by-frame analysis of eye movements yielded a record of the time-course of looking behavior as each auditory stimulus unfolded (see Fernald et al., 2008). The experiment was designed to determine whether children were able to make use of potentially informative prenominal color-adjectives in establishing reference. In order to do so, the speech stimuli (e.g., *Can you find the blue car?*) were paired with visual stimuli divided into three conditions summarized in Table 4.

Table 4. Examples of the stimuli for each condition in Fernald et al. (2010).

Condition	**Visual stimuli**	**Speech stimuli**
Same Color/Different Objects	blue car – blue house	***Where is the blue car?***
Different Colors/Different Objects	blue car – red house	***Where is the blue car?***
Different Colors/Same Object	blue car – red car	***Where is the blue car?***

First, the control condition was *Same Color/Different Objects*. In this condition children saw, for instance, a blue car and a blue house. Here, the adjective is uninformative and only the noun can be used to identify the correct referent. In the second condition, i.e., *Different Colors/Different Objects*, children saw e.g., a blue car and a red house. Thus, the adjective is sufficient to pick the correct referent, although not necessary since the noun alone is also sufficient. Finally, in the *Different Colors/Same Object* condition, children saw two cars, one red and one blue. Here, the adjective was the only cue that could be used for the correct target identification. What they predicted is that, on the basis of the two-step *Noun-Anchor Hypothesis*, children would wait for the noun to respond, taking as long to shift to the target in the two informative-adjective conditions, as in the control condition. Interestingly, Fernald et al. (2010) found different results for the two age-groups. Indeed, 30-month-olds failed to take advantage of the informative prenominal adjective to identify the name object more rapidly, and also showed difficulty in the integration of nouns and adjectives in combination. Moreover, in the *Different Colors/Same Object* (e.g., blue car vs. red car), some children responded by looking back and forth between the two cars as they heard the noun, before finally settling on the target picture. On the other hand, 36-month-old were more successful in this task. When the adjective was informative (i.e., in the *Different Colors/Different Objects* and in the *Different Colors/Same Object* conditions), they were more likely to orient to the target picture as soon as they heard the word *blue*, without waiting to hear the entire noun-phrase. Thus, target identification was faster in both informative-adjective conditions than in the control condition.

The findings for 36-month-olds showed that children make impressive gains over the third year in understanding sentences in real time, being able to interpret language incrementally. In contrast, the results from 30-month-old confirmed the *Noun-Anchor Hypothesis*, because children actually waited for the noun to start the interpretation.

2.5.2.4 Weisleder and Fernald (2009)

The evidence so far presented might suggest that having pre-nominal adjectives might be sub-optimal for children learning English, due to the working memory load they have to face with until after hearing the noun. In commenting Fernald et al.'s (2009) results, Weisleder and Fernald (2009) argued that the fact that children looked back and forth between the two pictures as they heard the noun in the *Different Colors/Same Object* condition suggested that the uninformative noun might have "disrupted" their processing of the adjective-noun phrase. Similarly, when asked a question like *¿Dónde está el carro azul?* ("Where's the car blue?") Spanish-speaking children could face a situation analogous to the one experienced by English learners in the *red car – blue car* condition: an informative noun (e.g., *carro*, "car") is followed by an uninformative adjective (e.g., *azul*, "blue") and, although the first word is sufficient for target identification, the second word, relevant to both images, might disrupt children's processing. Weisleder and Fernald (2009) refer to this proposal as the ***Sequential-Integration Hypothesis***: children have difficulties with the sequential integration of any two content words. According to this view, the integration of two words is easier when the informative word comes last for two reasons: first, the disruption effect is left out and second, this arrangement is less taxing on working memory.

To test this hypothesis, Weisleder and Fernald (2009) tested 25 Spanish-speaking children between 3;2 and 3;8 years in the looking-while-listening procedure conducted by Fernald et al. (2010), adapted to Spanish. Examples of the stimuli for each condition are summarized in Table 5.

Table 5. Examples of the stimuli for each condition in Weisleder & Fernald (2009).

Condition	Visual stimuli	Speech stimuli
Different Objects/ Different colors	blue car – red plane	***¿Dónde está el carro azul**?* *(Where's the car blue?)*
Different Objects/ Same color	blue car – blue plane	***¿Dónde está el carro azul**?* *(Where's the car blue?)*
Same Object/ Different Colors	blue car – red car	***¿Dónde está el carro azul**?* *(Where's the car blue?)*

Weisleder and Fernald (2009) found that Spanish learners interpret noun-adjective combinations incrementally. In the conditions in which the noun was informative in relation to the visual scene (i.e., in the *car blue – plane red* and in the *car blue – car red* situations), children were able to identify the target early, and did not seem to suffer any disruption from the adjectives, regardless of whether it was informative or not. In trials in which the noun was uninformative, children waited until the adjective to shift to the target picture. These results, together with Fernald et al. (2010)'s findings on English learners, showed that young children's processing of phrases with adjectives seems to be facilitated when the noun is heard before the adjective, i.e., in languages like Spanish or Italian. Taken together, these studies confirmed that the noun-adjective order is less taxing on working-memory because children can make use of the information in the order in which it is heard, identifying the category before its modifying property. Thus, consistent with the two-step process proposed by Ninio (2004), young children start the interpretation of noun-adjective combinations when they hear the noun, independently of whether it is the first or the last word of the combination.

2.5.2.5 Tribushinina and Mak (2016)

In a Visual World eye-tracking study[22] on Dutch, Tribushinina and Mak (2016) addressed the question of whether young children are able to process adjective-noun phrases incrementally and to predict the noun on the basis of the adjective meaning. According to their view, Fernald et al. (2010)'s finding cannot tell whether 3-year-olds are able to integrate the meanings of nouns and adjectives or if they just look towards "blueness" based on hearing the word *blue*. Indeed, in the informative conditions, only one blue object was displayed and, hence, hearing 'the blue one' would have been enough, being the noun irrelevant for the identification of the target object.

Tribushinina and Mak (2016) designed a paradigm where the target property was never apparently visible on the screen. Twenty-one Dutch-speaking children between 3;0 and 3;5 years were presented with two pictures on the screen (e.g., a grey butterfly and a grey stone), described by an adjective-noun phrase in either an Informative or Uninformative condition. In the Informative condition, the adjective term was enough to establish a reference (e.g., "The heavy stone"), while in the Uninformative condition the property term could match either picture (e.g., "The grey stone"). Examples of visual stimuli are given in Figure 1.

Figure 1. Example of a picture combination in Tribushinina and Mak (2016)'s experiment. In the Informative condition participants heard "The heavy stone". In the Uninformative condition they heard "The grey stone".

The results showed that, when the adjective was uninformative, the proportion of looks to the target picture did not differ from chance; however, when the adjective was informative about the following noun, looks to the target

[22] Tribushinina and Mak (2016)'s study is the only study at our knowledge making use of the eye-tracking technology to address the interpretation of noun-adjective combinations by children. See chapter 3 for a detailed overview of this methodology.

increased while hearing the adjective well before the onset of the noun. Moreover, children did not need more time to direct their look to the referent if compared to adults. Tribushinina and Mak (2016) concluded that children younger than four are able to process adjectives as they hear them and even predict the noun on the basis of the adjective meaning and, contrarily to results of Ninio (2004), attributive adjectives are not *always* processed after the head-noun.

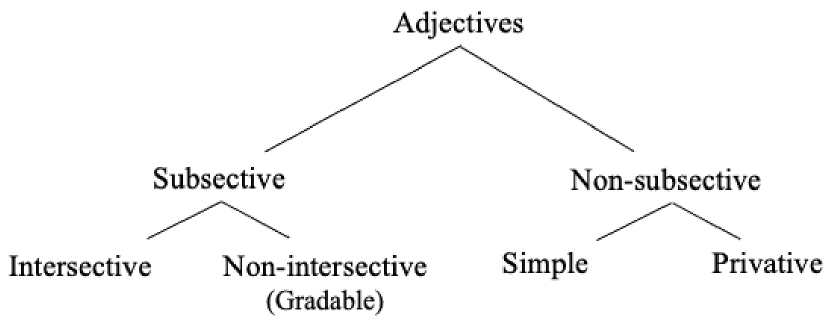
Although these findings seem to indicate an adult-like adjective interpretation in young children, we believe, however, that Tribushinina and Mak's results do not entirely disprove the *Noun-Anchor Hypothesis* and that different interpretations for their findings are possible. It is possible that, when hearing a property that is prototypically linked to the referent noun (e.g., *heavy* for a stone), children may retrieve information from their knowledge of the world and compute the most probable referent of that property in the visual world. In addition, young children may rely on patterns of adjective-noun co-occurrences in child-directed speech. Looking at the stone upon hearing the adjective *heavy* in a visual scenario with a stone and butterfly might be caused by the fact that, in the input, they are more likely to hear the combination *heavy stone* rather that *heavy butterfly*. Therefore, it might be the case that children are simply relying on their experience of the world and that they are not actually predicting the following noun on the basis of the adjective meaning per se. What is more, we believe that a further problem with Tribushinina and Mak's paradigm concerned the design of the auditory stimuli. Noun phrases were recorded with a 3-second interval between the onset of determiner and the onset of the adjective, and with another 3-second interval between the onset of the adjective and the onset of the noun. We believe that such a design does not mirror natural language processing, as long breaks between words are not common in adult natural speech nor in child-directed speech. Although by means of eye-tracking it is possible to study language comprehension online, we believe that this paradigm does not provide a comprehensive insight of real-time speech processing and adjective interpretation.

The studies presented in this section tackled several issues concerning adjective acquisition both developmentally and cross-linguistically, yet their main limit concerns the fact that they investigated mainly color adjectives (i.e., intersective adjectives, discussed in §1.4.2.2), and when they did not, semantic differences were not taken into consideration. The question of whether and how different semantic classes of adjectives influence children's online interpretation of noun-adjective combinations still remains open and will be addressed by the present work. We turn now to the relevant literature concerning the acquisition of adjectives. More specifically, we will report on offline studies examining the interpretation of intersective, relative and absolute adjectives both in isolation and in comparison. Finally, we will present the few online studies conducted on adjective classes exploring adults' processing.

2.5.3 Intersective, relative and absolute adjectives

As discussed in §1.4.2, adjectives can be classified semantically in set-theoretic terms and in relation to the interferences a noun-adjective combination can trigger (Kamp & Partee, 2005), leading to the following classification:

(86)



This classification allows for the distinction between subsective (e.g., adjectives of colour, size, material, etc.) and non-subsective adjectives (e.g., *alleged, fake, counterfeit*, etc.) (see §1.4.2.2 for a full discussion). Among subsective adjectives, a further distinction can be made between intersective and gradable adjectives. This last distinction will be discussed in detail as it will be relevant for the experimental part of the present work.

Combining an **intersective** adjective with a noun puts the described object at the intersection between the set of objects denoted by the noun and the set of objects possessing the property labeled by the adjective. Color adjectives fall under this category. For example, the referent of “red fruit” is at the intersection between the set of fruits and the set of red objects in the world. By contrast, subsective gradable adjectives (e.g., *tall, heavy, cheap*) trigger a non-intersective interpretation, but can only be interpreted with respect to a class of comparison defined by the noun. Indeed, their core meanings involve a relation to a scalar concept on the basis of which objects can be ordered (such as height, weight, cost). This class of adjectives is usually referred to as gradable adjectives (Kennedy, 1999). Among gradable adjectives, a distinction between **relative** and **absolute** adjectives is made, based on the property of the scale along which predicates order their arguments. Thus, relative adjectives (e.g., short/tall) evoke an open scale, while absolute adjectives have at least one closed boundary (e.g., clean/dirty) (see §1.4.2.1 and §1.4.2.2). For the purpose of the present work, intersective, relative and absolute adjectives will be taken into account and analyzed as for the differences they show in syntactic distribution and in their relation to the context of use. Furthermore, we will discuss the implications on acquisition by young children, to their interpretation and their integration with the referred noun.

A first difference between intersective, relative and absolute adjectives concerns their syntactic distribution. Intersective adjectives do not always appear felicitously in comparative constructions (e.g., $^{?}$*redder than,* $^{?}$*as red as,* $^{?}$*less red than*), while relative adjectives do, since they evoke a scale open at both ends (e.g., *taller than, as tall as, less tall than,* the negative counterpart being *shorter than, as short as, less short than*). For absolute adjectives a further distinction needs to be made. Indeed, the scale absolute adjectives evoke can either be “closed” at one (e.g., *clean/dirty*) or both ends (e.g., *open/closed*). Thus, comparative constructions are possible only for the open boundaries (e.g., *dirtier than, as dirty as, less dirty than,* but **more open than, *as open as, *less open than*).

Intersective, relative and absolute adjectives also differ in the way they are tied to the context. Intersective adjectives are context-free, in the sense that they can be interpreted independently from the context of utterance and do not require any class-based dependence[23]. For instance, when described as a “red fruit”, a strawberry can easily be interpreted as being red or not, independently from other fruits or other red objects in the context of use. On the contrary, relative and absolute adjectives show different degrees of context sensitivity (e.g., Rotstein & Winter, 2004; Kennedy & McNally, 2005; Kennedy, 2007; McNally, 2011; Sassoon & Toledo, 2011; among many others).

Relative adjectives are context-sensitive in the sense that different contexts pick out different sets of objects for which the adjective can be truthfully predicated. For instance, what counts as *tall* can vary depending on different aspects, so that the extension of *tall* is context dependent. The relevant factors for the interpretation of the extension of the property include the denotation of a modified noun (e.g., *tall snowman* versus *tall building* versus *tall mountain*); an implicit or explicit comparison class (e.g., *tall for a 10-year-old*); extra-linguistic knowledge (e.g., a snowman built by third graders versus fraternity brothers; see Kamp & Partee, 1995); and the interests and expectations of the participants in the discourse (Graff, 2000). Thus, for the interpretation of relative adjectives a standard of comparison is always needed, and that standard is located at the mid-point of the scale: to count as *tall*, a building needs to exceed a contextually determined standard of height (see, e.g., Bartsch & Vennemann, 1972; Kennedy, 1999, 2007), a degree that corresponds to an “average” or a “norm” for the scalar concept relative to some salient set, which may be explicit, inferred or implicit in the sentence.

The interpretation of absolute adjectives can also vary from context to context. According to some theories (e.g., Lassiter & Goodman, 2013), contextual sensitivity is reduced to how much distance from the end-point interpretation is tolerated within the context. Hence, the standard of comparison is more rigid and

[23] Color adjectives have been claimed to manifest various types of context-dependence (see, e.g., Kennedy & McNally, 2010). Indeed, different studies showed how the interpretation of adjectives can be different according to the modified noun. For example, Medin and Shoben (1988) asked participants to rate a phrase, e.g., *grey clouds,* as more similar to e.g., black clouds or white clouds, and found that the results were opposite when participants had to rate e.g., grey hair. Nevertheless, color adjectives have not been argued to show the same kind of comparison class-based context dependence of relative and absolute adjectives, especially not in the language acquisition perspective that we are interested in here.

tends to lay very close to the end-point itself (see, e.g., Kennedy & McNally, 2005; Kennedy, 2007; McNally, 2011). For example, to count as *empty*, a glass needs to be very close to the maximum degree of emptiness. In contrast, other views argue that, although absolute adjectives are demonstrably gradable, they have fixed and context-independent standards. In the present work, we will focus on one possible implementation of this view, that of Kennedy (2007). Building on previous work on the topology of adjectival scales (Kennedy & McNally, 2005), Kennedy (2007) traces the relative/absolute distinction back to the structural properties of the scale denoted by the adjective. If, on the one hand, context-sensitivity with relative adjectives derives from the fact that their scales are open and unbounded at both endpoints, the scales evoked by absolute adjectives are closed in either one or both thresholds. Thus, whereas with relative adjectives the value of the threshold cannot be provided by the scale itself and in most cases must be computed with respect to a contextually provided comparison class, absolute adjectives live in closed scales. In general, they either have a *maximum standard* or a *minimum standard*, conventionally provided by one of the scale's endpoints. *Maximum standard* gradable adjectives, such as *full*, *straight* or *dry*, require their arguments to possess a maximal degree of the relevant property. For instance, *full* is true for a container just in case it is completely full; *straight* is true of an object just in case it has no bend and so on. In contrast, *minimum standard* absolute adjectives, such as *spotted*, *wet* or *bent* only require their arguments to possess some minimum, non-zero degree of the property. For example, *spotted* is true of an object as long as it has some spots; *bent* is true of an object if it has some degree of bend and so on. In conclusion, the context-sensitivity displayed by relative and absolute adjectives is fundamentally different: relative adjectives are semantically context-sensitive, whereas absolute adjectives are pragmatically context-sensitive.

2.5.3.1 Experimental studies on color-adjectives

Although color and gradable adjectives are recurrent in adult language, in children's early vocabulary adjectives are infrequent (Gasser & Smith, 1998). Parents report that by the age of 30 months, 85% of English-speaking children correctly produce most color words, according to norms for the MacArthur-Bates

Communicative Development Inventory (Dale & Fenson, 1996). However, the pattern of acquisition of color terms seems to be particularly problematic.

It is very well known that Charles Darwin, among other parents observing their children's first words, was convinced that his child was color-blind because he showed no response to colors (see Landau & Gleitman, 1985). Both observationally and empirically, the literature has repeatedly shown that, in comparison with other types of words, the course of acquisition for color is protracted and errorful (see, e.g., Rice, 1980; Bornstein, 1985).

Rice (1980) taught children who knew no color words the words *red, yellow* and *green*. Starting with *red*, she showed children a series of red objects and asked the child to name the color of each. It took an average of 85 trials to reach success and an average on 1080 trials for the children to be able to respond correctly to all three colors. Carey and Bartlett (1978) and Heibeck and Markman (1987) taught children a new color word using a contrastive technique, that is, they contrasted a new-color word with a known-color word. Overall, children succeeded in comprehending the new word and recognizing it as a color term after one single trial. Soja (1994) explained the difference between the latter two studies by suggesting that, in the second study, children already knew at least one color word. Thus, children might experience some sort of obstacle to the color word lexicon but, once they overcome it, they learn color words more easily.

In learning color words, children need to make a mapping between a word and a color. In other words, they must represent the word, represent the color, and then make an association between the two. Different studies showed that children have perceptual color categories that seem to be exactly like adults' (e.g., Bornstein et al., 1976) and know that some color words are subordinates to the word *color*, even though they do not know the meaning of any color words. Indeed, when asked to answer the question *What color is it?* with a color term, the color terms children younger than 3 provide seem to be randomly chosen and unrelated to the property in question (e.g., Backscheider & Shatz, 1993, Sandhofer & Smith, 1999). One possible explanation is that children first understand that color words refer to color without knowing how color words map to specific color categories (Carey, 1982). Another possibility is that, at this early stage, children might only know that words such as *blue* or *red* are appropriate answers to *What color is it?,* without knowing why. In other words, toddlers answer with a color because of mere linguistic association, and not conceptual knowledge. This second option is also supported by the fact that children blind from birth, who never experienced colors, also answer this question in the same way (see, e.g., Landau & Gleitman, 1985).

Soja (1994) argued that children have color concepts and can make inferences based on colors but only later infer that a color word maps onto this conceptual representation. In contrast, Kowalski and Zimiles (2006) confronted the issue of color-language versus color-concepts and proposed that children initially do not have abstract conceptual representations of color but they quickly acquire color concepts after they learn some color words. However, stating that children possess a conceptual representation at a specific point in time might be misleading, since it may be dependent on which task is presented to children. Thus, the value of addressing this issue using different types of conceptual color tasks is that a broader picture emerges, taking into account the various aspects entailed in color adjective acquisition. In sum, the developmental progress in learning color words seems to move from learning words to learning the properties those words refer to (Sandhofer & Smith, 2001).

2.5.3.2 Experimental studies on gradable adjectives

Gradable adjectives are among the first adjectives to be produced by children (Nelson, 1976; Blackwell, 2005). In a study conducted on the Brown corpora within the CHILDES database, Blackwell (2005) found that the majority of the adjectives used in an age range between 18 and 35 months are either relative (e.g., *big*, *little*, *cold*, *hot* and so on) or absolute (e.g., *dirty*, *clean, full, empty* and so on).

Early studies on gradable adjectives mainly focused on relative adjectives only, investigating their acquisition across different dimensions (e.g., size, length) and to the different poles of scales (e.g., the use of *big* vs. *small*).

Blackwell (2005) noticed that as children build up their language repertoire, they show adult-like behaviour in the interpretation of relative adjectives when the standard of comparison is easily identifiable (e.g., when asked to judge the size of a ball among a group of balls), but they are reported to make a consistent series of errors, including the overuse of dimension terms (e.g., using *tall* for *big*), or extreme labelling (e.g., when presented with a series of seven items that decrease along a relevant dimension, e.g., seven blocks decreasing in size, they tend to label only the extreme item of the series as *big* or *small).*

Eilers et al. (1974) tested English learners between 2;6 and 3;6 years in a forced-choice task. They presented children with pairs of objects, differing either in size, length, or width, examining the comprehension of *big-little, long-short,* and *wide-narrow*. The experimenter made a request such as *Give me the big one*. Each object pair was presented twice: once with a positive pole adjective (e.g., big) and once with its negative counterpart (e.g., little). Children made fewer errors (i.e., picked the wrong object) for *big-little* than for *long-short*, and fewer errors for *long-short* than for *wide-narrow*. Eilers and colleagues (1974) concluded that more specific dimensions such as length or width are acquired later than more general dimensions such as size. In addition, their findings do not suggest a difference in the acquisition of positive pole and negative pole adjectives. Bartlett (1976) extended Eilers et al.'s (1974) research to a wider age range, testing children between 2;1 and 4;10 years and added objects differing in height. In line with and colleagues (1974)' results, they found that more general dimensions are acquired earlier than specific ones and that performance for *wide-narrow* at age 5 was not yet at ceiling.

Taken together, these studies show that at age 3, children are able to compare two objects with regard to some dimensions and to decide on the basis of this comparison for which object an adjective is true. Sera and Smith (1987) argued that this ability can be regarded as a first instance of a relational interpretation of relative adjectives. However, Clark (1970) proposed that children start with a so-called "nominal" interpretation of relative adjectives, i.e., initially, they interpret adjectives like *big* as "having extent", and not as relative to a standard of comparison. Consequently, children either apply *big* only to the biggest object in a group because it is the most typical exemplar, or they call *small* also objects that are big. Possible evidence for this assumption is provided by Panzeri and Foppolo (2012) and Huang and Snedeker (2013).

Panzeri and Foppolo (2012) tested 3- and 5-year-old Italian learning children in a Truth-Value Judgment task by presenting objects in isolation (e.g., a wooden rod). The main assumption behind the experimental design is that this presentation does not provide any standard of comparison and, because no other objects were present in the visual context, participants could not use a perceptual standard. Participants were asked to judge whether a sentence such *This is tall* was true for the presented object. Since no contextual standard was provided, children were expected to answer *I don't know* or to show chance distribution of "yes" and "no" responses. Surprisingly, 3-year-olds replied 'yes' in 80% of the trials, and the 5-year-olds in 60% of the trials. Since there is ample evidence in the literature that children know adjectives like the ones tested in this experiment from age 2 on (cf. Syrett, 2007), Panzeri and Foppolo (2012) excluded that the high rate of acceptance was the result of the lack of knowledge the meaning of the adjectives combined with an alleged. Moreover, children rejected test sentences in the filler trials that were specifically designed to prompt a *I don't know*-response (e.g., a toy-zebra was described as *This is obedient*; 3-year-olds split between "yes" and "no": around 50% accepted the description while the other 50% rejected it). Thus, Panzeri and Foppolo (2012) concluded that they could not have a tendency to opt for a yes-answer in case of uncertainty. However, this experimental setting cannot exclude that children may have some conceptual representation of the objects used that provided the standard of comparison.

Huang and Snedeker (2013) conducted an online Visual-Word study with 5-year-old English-speaking children. Participants saw 4 pictures and heard test prompts such as *Point to the big coin*, in the context of two coins with different sizes (i.e., a big coin and a small coin), a distractor of the same size as the target object (e.g., a big stamp), and an irrelevant object (e.g., a marshmallow). Huang and Snedeker (2013) found that children looked to the target object more rapidly if the display contained two coins rather than in the control condition with only one coin. This finding demonstrates that at age 5, children typically evaluate whether a property denoted by a relative adjective is true for an object by comparing it to another object.

Another reflex of children's so-called "nominal" interpretation of relative adjectives (i.e., the interpretation of such adjectives as having an extension and not as relative to a standard) is that children apply the adjective only to the object which displays the respective property most. Hence, when presented with a series of objects, children younger than age 3 seem to prefer to label only the extremes of the series (cf. Clark, 1970; Tribushinina, 2013). For example, *big* is only accepted for the biggest object in a series, *small* for the smallest and so on. If the relative gradable adjective is interpreted as "having extent" or "having the property denoted by the adjective", the extreme objects can be viewed as the best exemplars representing the respective property.

When presenting 2- and 3-year-old Dutch-speaking children with a series of seven objects increasing in size, Tribushinina (2013) observed exactly this interpretation. The objects were either prototypically big (elephants, hippos, houses, planes), prototypically small (mice, chickens, gnomes, babies) or prototype-neutral (balloons, cakes, monkeys, umbrellas). Each of these series was paired with a question, e.g., *Which one do you find big/small?*. In more than 50% of the trials children selected only respectively the biggest and the smallest object. However, it is crucial to note that the way of asking (i.e., a subjective question) may have favored answers in which the participants chose only the best exemplar.

A further case in point for children's "nominal" or relational interpretation of relative adjectives is that they are reluctant in switching the standard of comparison. This means that young children do not re-label an object when the context changes. Sera and Smith (1987) investigated whether 2- to 4-year-old English learners are able to re-label objects as *big* and *little* de- pending on the visual context. This means the standard of comparison had to be shifted according to changes in the perceptual context. The perceptual context was created using three circles of different sizes: one small, one medium, and one big. In the first part of each trial, the experimenter presented, e.g., the small and the medium circle and the child had to label the two circles *big* and *little*. In the second part, the small circle was removed and replaced by the big one. The child again had to label the two circles *big* and *little*. Hence, the medium circle should have been labelled *big* in the first part and *little* in the second part. If they said a circle was *big* in a context of smaller objects, 2- and 3-year-olds still claim it is *big* in a new context where it is compared to bigger objects. However, given large size differences between objects, children shifted the labels above chance-level, showing that their performance increased when the size differences between the objects were large. Finally, Sera and Smith (1987) observed that the re-labelling ability seems to emerge at age 4.

Around four years of age, there seems to be an interesting turn: if, by the age of three, children can correctly pick out a big object from a series, at four they appear to rely too much on the vertical dimension (i.e., height) in their judgment (e.g., when asked to locate *the big one*, they may select the tallest one). This decline in understanding gradable adjectives does not seem to last long and, by five or six years of age, children are able to consider the correct dimension when judging object size using the correct term (e.g., Sera & Smith, 1990).

Few studies compared the acquisition of relative and absolute adjectives.

Using the Scalar Judgement Task, Syrett et al. (2006), Syrett (2007) and Syrett et al. (2010) investigated whether 3- to 5-year-old children have different standards of comparison for relative and absolute adjectives and whether they distinguish these two classes with respect to their context-dependency. In the Scalar Judgment Task English-speaking children saw a series of seven objects (e.g., seven containers) displaying the same property (e.g., length), but to different extents. Children were asked to judge whether each object could be described by a certain property and were asked a question about each of them, i.e., *Is this ADJ?*. The adjective was either relative (*long*, *big*) or absolute (*bumpy*, *full*, *spotted* or *straight*). As predicted, children located the standard for relative adjectives in the middle of the scale, that is, children gave a “yes”-response to the three longest and biggest objects more often than they did for the other four objects. Conversely, children’s standard for absolute adjectives was located at one of the endpoints of the scale, but only for the minimum standard adjective *spotted*: they answered “yes” to all objects with a non-zero degree properties, i.e., to all six objects with at least one spot. Contrary to what the authors expected, with the maximum-standard adjectives such as *full*, children accepted also those containers that were not maximally full, suggesting a non-endpoint standard. Although these findings point to a context-dependent interpretation of maximum standard gradable adjectives, Syrett and colleagues (2010) claimed that this apparent context-sensitivity is not part of the lexical meaning of the adjective as it is for relative adjectives. Instead, absolute adjectives can be used in a context-sensitive way, resulting in imprecise meaning.

Foppolo and Panzeri (2013) extended the same Scalar Judgment Task to antonym pairs (e.g., *tall-short, straight-bent, full-empty*), as shown in Figure 2. They tested 38 Italian-speaking 3- and 5-year-old children by using positive and negative pole adjectives. For each element in the series, children were asked *Is this ADJ?*. Foppolo and Panzeri (2013) tested both relative and absolute adjectives and analyzed whether children divide the scale between the antonym pairs or whether they exhibit a "gap" between the positive and the negative poles, failing to judge the objects in the middle of the scale.

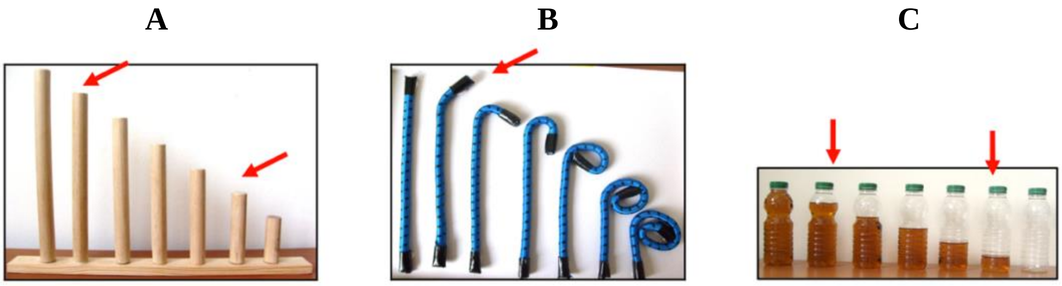


Figure 2. Example of the materials used by Foppolo and Panzeri (2013) to test relative and absolute adjectives in the Scalar Judgment Task. The set in (A) was used to test the relative adjectives *tall* and *short*; the second object from the left was used for *tall* and the second from the right was used for *short*. The set in (B) was used to test the absolute adjectives *straight* and *bent*, projecting on a scale closed on one end; the second object from the left was used for both antonyms. The set in (C) was used to test the absolute adjectives *full* and *empty*, projecting on totally closed scales; the second object from the left was used for *full* and the second from the right was used for *empty*.

They found that children do exhibit the gap between the positive and negative pole, more strongly for *clean-dirty* than for *full-empty*. However, since participants were presented with either the positive or the negative pole, it is not possible to compare their individual standards for positive and negative adjectives.

Syrett (2007), Syrett et al. (2006), and Syrett et al. (2010) also found a difference between the interpretation of relative and absolute adjectives with respect to the relevance of the comparison class: relative adjectives were interpreted relative to a comparison class, whereas absolute adjectives were not. In their studies, the adjective did not modify a noun, hence the comparison class was implicit and had to be inferred from the context.

2.5.3.3 Experimental studies on color vs. gradable adjectives

In the present section, we report on studies investigating the acquisition of color vs. gradable adjectives. As we have already mentioned, although children are able to use color terms for object categorization, the production of color words seem to lag behind many subsective adjectives. Indeed, the pattern of acquisition of gradable adjectives seems to be very different from that in the case of intersective adjectives and, more specifically, colors. Many differences between color terms and gradable adjectives could be responsible for the different learning trajectories.

A first possibility is linked to their meanings. Indeed, while color terms refer to categories of objects (i.e., a red object is red in all contexts), gradable adjectives refer to relations between objects (i.e., a big object is big in some comparison contexts and little in others) (e.g., Clark, 1970).

Another possible explanation for this asymmetry was proposed by Landau and Gleitman (1985). They argued that children expect property terms to refer to relative and binary contrasts, such as *slow/fast* or *big/little*. According to their view, this expectation supports the acquisition of subsective adjectives because they have this property, contrary to intersective adjectives who do not meet this expectation.

In their analysis of the organization of adjectives, Gross, Fischer, and Miller (1989) also assume that adjectives refer to binary properties. The color lexicon, by contrast, is organized differently, because there are many labels for different attributes of the color dimension and their meaning does not depend on a comparison with a standard. Consequently, this expectation inhibits the acquisition of color words.

A further hypothesis is that color adjectives are learned at different stages because of the way dimensions are commonly talked about and taught in the input (Sanhdofer & Smith, 2001). Color terms are usually taught by presenting a single object and asking *What color is this?*. This type of question emphasizes color words as category labels and teaches toddlers that the questions about colors demand an answer from a small set. By contrast, size words might be taught by presenting multiple objects and asking *Where is the big one?,* thus encouraging children to compare items in a set and select the one that best fits the provided label. To evaluate their hypothesis, Sanhdofer and Smith (2001) tested native English-speaking adults in a series of studies in which novel dimensional adjectives were taught in two experimental conditions. The first simulated the production task structure and categorical meaning of color terms (*What color is this?*), whereas the second condition simulated the comprehension (pointing) task structure and relational meanings of gradable adjectives (*Where is the big one?*). They found that, when taught in the "color-way", mastery of words succeeded before correct matching. By contrast, in the comprehension condition gradable adjectives are usually taught, success in matching happened before mastering the words. Thus, novel word learning depended only on the training condition and not on the meanings to be learned. Sanhdofer and Smith (2001) concluded that what makes the developmental pattern in children's learning of size terms and color terms is the way they are heard in the input.

A last explanation for the asymmetry between color and subsective adjectives emerged from Wagner et al.'s (2013) study, which examined the interpretation of relative adjectives compared to color adjectives. They found that children between two and seven years of age (except for the 4-year-olds), interpreted color terms relative to a comparison class. Contrary to predictions, and similar to relative adjectives, children changed their judgements for *blue* depending on the objects in the perceptual context. Wagner et al. (2013) interpreted these results as reflecting a default reading of adjectives as dependent on a comparison class that children initially have. Note, however, that the experimental design involved the visualization of color adjectives as gradable, which may have affected children's interpretation.

Deutsch and Pechmann (1982) explored which adjectives children use in referential description. They showed 3-, 6-, and 9-year-old Dutch learners an array of eight objects and were instructed to describe the object they want the experimenter to select. The objects differed in category, size, and color. Thus, children had to modify the noun by an intersective adjective (describing the color) and a relative adjective (describing the size) to refer to the object the experimenter should select unambiguously. The experimenter started by asking the child *Which one do you want?*, the unequivocal redundant description of the target reference being, for example, *The big red ball*. Whenever the participant used an ambiguous description (e.g., *The ball*), the experimenter repeated the description in question-form (e.g., *Which ball?*). First descriptions were ambiguous 87% of the times for 3-year-olds, 50% of the times for 5-year-olds and 22% for 9-year-olds. In other words, some information was often missing. Although the noun was mentioned in most descriptions, children mainly omitted information regarding the object's color, size, or both. Especially in younger children, the relative (size) adjective was more often omitted than the intersective (color) adjective.

To conclude, children have troubles in acquiring all classes of adjectives for different reasons. On the one hand colors seem to be particularly challenging and young children learn labels before meanings. On the other hand, gradable adjectives are acquired earlier, but children keep making mistakes with their interpretation up to four years of age.

All the studies summarized in this section made use of offline methodologies to investigate children's interpretation of adjectives. Although interesting results have emerged, a more in-depth investigation is necessary to understand how these differences affect the interpretation in real-time. The next section will report on the few studies investigating this issue with adults examining the online processing of the adjectival semantic classes at issue.

2.6 Adults' processing of noun-adjective combinations

The main focus of the present work is the acquisition and the online interpretation of adjectives by young toddlers. In order for a developmental study to provide accurate hypotheses about children's pattern of acquisition, it is necessary to understand what the expected outcome of acquisition is, i.e., to investigate adult processing. To this end, in this section we report on studies examining adult's online interpretation of the class of adjectives from two perspectives, namely the incremental processing of prenominal adjectives (§2.6.1) and the interpretation of different adjective classes (§2.6.2).

2.6.1 Incremental processing

There is growing evidence suggesting that adults process language incrementally, that is, words are processed as they are being heard and immediately integrated with the information from the visual scene (e.g., Tanenhaus et al., 1995; Hagoort et al., 2009; Rayner & Clifton, 2009). Furthermore, adults can predict how discourse will unfold on the basis of semantic information associated with a word currently being processed (e.g., Altmann & Kamide, 1999; Sedivy et al., 1999; Kamide et al., 2003).

Although incremental processing seems effortless to the fluent adult and has several cognitive advantages, such as avoiding burdening short-term memory, it inevitably brings an important implication: temporal ambiguities (Tanenhaus et al., 2000). Temporal ambiguities lead readers and listeners to occasionally make incorrect commitments that require revision at a later point in the sentence, being led down the "garden path". Thus, incremental processing raises the question of how these ambiguities in language are resolved to interpret the meaning of spoken utterances (Tanenhaus et al., 2000).

Because speech unfolds from moment to moment, it is frequently the case that multiple alternative interpretations are possible at a particular moment in time, only one of which is correct. For example, in a sentence like *Susie fixed her back porch*, there are ambiguities at multiple levels of the linguistic structure at different time points. For instance, halfway through the word *fix*-, it could continue as *fixes* or *fixates*, or on hearing *Susie fixed her back–*, one might infer that Susie consulted a doctor (Thorpe & Fernald, 2006).

Many studies have monitored adults' response to temporal syntactic ambiguities using speed of processing tasks. A famous example is the one from Bever (1970), *The horse raced past the barn fell*. When interpreting this "garden path" sentence, listeners first interpret the word *raced* as the verb of the main clause, although in fact it functions as the beginning of a reduced relative clause. When listeners parse the sentence incorrectly at the point of ambiguity, they need to revise their initial interpretation based on subsequent information (e.g., Frazier & Rayner, 1982; Crain & Steedman, 1985).

This process of recognizing words and resolve ambiguities while listening to spoken sentences is, for adults, remarkably fast and efficient, being enabled and facilitated by the use of predictive cues from multiple information in the speech stream (Thorpe & Fernald, 2006). For instance, lexically specific syntactic information, semantic plausibility, frequency of lexical co-occurrence and referential context have all been found to rapidly constrain adults' on-line commitments to interpretation, even in the presence of syntactic ambiguity (e.g., Altmann et al., 1994, Britt, 1994, Tanenhaus et al., 1995, Trueswell et al., 1993, Trueswell et al., 1994; Trueswell, 1996). Such information, either linguistic or contextual, help adults in constraining the possibilities of interpretation, so that understanding happens with no effort and accurately as the speech unfolds (e.g., Trueswell & Tanenhaus, 1994; Spivey et al., 2002). Contextually dependent expressions are quickly evaluated with respect to the context and help with a variety of on-line decision processes, including those pertaining to the structural analyses of a sentence (Trueswell et al., 1999).

As already mentioned in §2.5.2.4, this problem has been often addressed in studies on Germanic languages (mostly English) using prenominal adjectives. Indeed, they can be potentially ambiguous until the following head noun is uttered later in the sentence. Such ambiguity can lead to misinterpretation, thus reducing the speed and efficiency of understanding. English adjectives often serve to restrict the domain of possible reference; thus, a series of studies have been developed to investigate whether listeners could process adjectives incrementally to identify a referent within a larger set of possibilities.

In a series of experiments reported in Eberhard et al. (1995), participants were given spoken instructions to manipulate a set of real objects, while their eye movements to the objects in the visual display were monitored throughout the instruction. The visual array was manipulated in a such a way as to vary the point in the linguistic stimulus in which the instruction becomes unambiguous. For example, the sentence *Touch the plain red square* was paired with three different visual conditions, namely the Early-disambiguation, the Mid-disambiguation and the Late-disambiguation conditions shown in Figure 3.

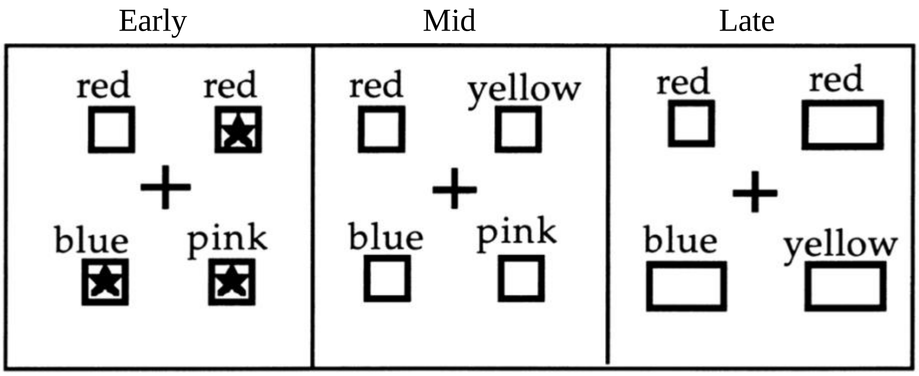


Figure 3. Example displays from the three experimental conditions in Eberhard et al. (1995)'s study. The accompanying instruction to this example was 'Touch the plain red square'.

In the Early-disambiguation condition, the visual array presented three objects marked with a star and a single object with no marking. Thus, the target could be identified already at the word *plain*. The Mid-disambiguation condition had a display consisting of four plain objects, only one of which was red. Hence, target identification was possible after hearing *plain red*. In the late disambiguation condition, the visual array consisted of four plain blocks, two of which were red. Of the red blocks, one was square in shape, and the second was rectangular. In this last condition, identifying the target was only possible after hearing the whole sentence *plain red square*. Hence, by manipulating the displays, it was possible to alter the disambiguation word. The results showed that eye movements occurred generally well before the end of the referential expression and that, more interestingly, were time-locked to the point in the speech stream where it became possible to pick out a unique object from among the alternatives in the display.

Like the processing of structural representations for a linguistic string, the process of establishing reference to the speech stream is also incremental, resulting in local indeterminacies (Sedivy et al., 1999). As Altmann and Steedman (1988) pointed out, "The process of incremental evaluation involves having available representations of 'partially evaluated' referents. These are simply the members of the set of referents which satisfy the available constraints. This set gradually becomes more and more refined as the analysis proceeds, until just the candidate referent remains" (Altmann & Steedman, 1988: 196). Sedivy et al. (1999) argued that, in cases of local indeterminacies, information from the context of utterance is available to pin down the meaning of a linguistic expression.

The question of whether also children interpret language incrementally and are able to predict how a sentence will unfold based on semantic information and context is still open to debate. Fewer studies investigated the online speech processing by young language learners with respect to adults, mostly because of limited experimental tools. With the refinement of the eye-tracking techniques for use with infants, in the last decades it has been possible to monitor the time course of spoken language understanding by very young language learners in comparison with experienced adults. However, to our knowledge, only one study has explored adjectives in this respect using the Visual World paradigm (i.e., Tribushinina & Mak, 2016, see §2.5.2.5).

Although it is interesting to notice that prenominal adjectives are a suitable test for incremental processing, studies on the topic did not take into account the semantic differences between adjective classes. Indeed, interpreting intersective, relative or absolute adjectives has different implications as for the relationship with the (visual) context and the identification of a standard of comparison. We turn now to the studies on adults' interpretation of different semantic classes of adjectives.

2.6.2 Studies on adjective classes

In a well-known Visual-World Paradigm, Sedivy et al. (1999) conducted a series of experiments to investigate the online semantic interpretation of different English adjectives. The first experiment focused on intersective adjectives, including adjectives for colors, materials and shapes. Participants were instructed to touch one of the four objects on a horizontal workspace, while a head-mounted eye-tracker recorded eye-movements. The trial either belonged to the Early-disambiguation or the Late-ambiguation condition. Upon hearing, e.g., *Touch the blue pen*, in the Early-disambiguation condition participants saw a blue pen (i.e., the target), a yellow rubber duck, a red notebook and a pink comb. Thus, target identification was possible at the adjective *blue*. In the Late-disambiguation condition, the same instruction was paired with a display containing the blue pen, a blue bowl (i.e., the competitor), a yellow duck and a red notebook (i.e., the distractors). Hence, the point of disambiguation was at the noun *pen*. The same experiment was conducted by substituting one of the two distractors with an object-competitor (e.g., a yellow pen). Results replicated Eberhard et al. (1995) findings by showing evidence for incremental semantic processing with respect to a visually available set of potential referents. In other words, these data indicate that nouns that are modified by adjectives are interpreted incrementally, with immediate mapping onto a referential model. However, with intersective adjectives, it is possible to evaluate the set of entities that are denoted by the adjective *yellow,* or *rectangular,* independently of knowledge of the noun that is being modified.

A further experiment was thus conducted to investigate the incremental interpretation of relative adjectives. There were four objects in the display, belonging either to the Contrast or the No-Contrast condition. For instance, when hearing the instruction *Pick up the tall glass,* in the Contrast condition participants saw the target object (e.g., a tall glass), a contrast-object (e.g., a short glass), a competitor (e.g., a pitcher) and a distractor (e.g., a key), as shown in Figure 4. Thus, the Contrast condition supported a contrasting interpretation of the adjective by including, alongside the target object, a contrast object that could be described by the noun but not by the adjective in the instruction. In the No-Contrast condition a second distractor (e.g., a file folder, that could not be described either by the head noun or the modifier in the instruction) occurred in the place of the contrast object.

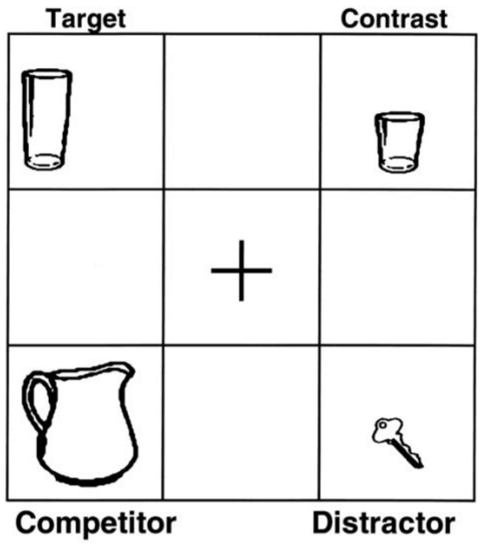


Figure 4. Example of the visual display used by Sedivy et al (1999) in the Contrast condition. In the No-Contrast condition the contrast object was replaced by a second unrelated distractor (e.g., a file folder).

Sedivy et al. (1999)'s main finding of this experiment was that participants' fixations zoomed to the target object faster in the Contrast condition than in the No-Contrast condition. Upon hearing the adjective, in the Contrast condition participants' fixations converged to the target, at a point in which the adjective had not been processed yet. Although the linguistic instruction was still compatible with both the target (e.g., the tall glass) and the competitor (e.g., the pitcher), participants were able to identify the target, suggesting that the information about the contrasting object was used very quickly. Sedivy et al. (1999) proposed two possible interpretations for these results.

First, the effect of referential contrast is driven by the use of the restrictive prenominal modifier. The presence of the adjective in the Contrast condition might trigger a pragmatic reasoning, such that hearers assume that the Gricean Maxim of Quantity is observed and thus, uttering the prenominal modifier would be maximally informative only in those cases in which the adjective is used to disambiguate between two possible referents. As long as the context supports a restrictive interpretation, this view predicts that effects of referential contrast should arise with any prenominal modifier.

A second possible interpretation is that aspects of the lexical semantics of gradable adjectives drive the effect of contrast. Indeed, since the threshold for relative adjectives must be resolved via a contextually supplied comparison class, participants could be integrating the contextual cues as part of the processing of the adjective. Thus, target identification if the No-Contrast condition is slow because it reflects the difficulty of processing a relative adjective in the absence of a proper contextual support. This account predicts that, for the adjectives that do not have context-sensitive thresholds (i.e., intersective adjectives), the presence or absence of a contrasting object in the visual display should not affect adjective processing. This second option, however, was rejected in a subsequent eye-tracking study replicating the same method and design, using intersective adjectives.

Sedivy (2005) tested color and material adjectives in the Contrast and No-Contrast condition and predicted that, since these types of adjectives are not context-sensitive, no effect of contrast should arise. Although this prediction was confirmed for color adjectives, material adjectives did give rise to a Contrast effect. Indeed, in this condition, target identification was faster. For this reason, Sedivy (2005) discarded the hypothesis that adjective processing is lexically driven in this task and argued that pragmatic reasonings and perceptual salience of colors must be responsible for the contrast effect.

In a follow-up eye-tracking study on color adjectives, Sevidy (2003) found that the presence of the facilitating contrast effect in target identification occurred only when the color adjective was predictable from the head noun, e.g., when the target object was a yellow banana, and the contrastive object was a brown banana. Thus, the prototypicality of the property played a role and helped participants in predicting the target already in the adjective-window. Sedivy (2003) argued that the use of predictable adjectives is highly informative for participants, as otherwise the speaker would be overspecifying by using a modifier that is redundant given the noun. This account is also compatible with the absence of contrast effects for color adjectives found in Sedivy (2005), where all color adjectives were non-prototypical and hence unpredictable from the noun. As Aparicio et al. (2015) pointed out, the fact that Sedivy (2005)'s findings point to a pragmatic-based interpretation of adjective-noun combinations does not necessarily entail that lexical meaning does not play a role in the integration of contextual information during online processing. It is possible that certain classes of adjectives, like color adjectives, only show informativity-based effects of contrasts, while other adjectives, such as relative adjectives, a combination of both lexical-semantics and informativity-based effects.

To address this issue, Aparicio et al. (2015) replicated Sedivy et al.'s (1999) design, extending it to absolute adjectives and substituting real world objects with geometric shapes, in order to avoid world-knowledge bias effects. Similar to Sedivy et al.'s (1999) study, two visual conditions were tested. In the Contrast condition, participants saw four objects on the screen: a target (e.g., a tall cylinder), a competitor sharing the target property, but presents a different shape (e.g., a tall spiral), a contrast object belonging to the same object-class as the target, but with a different property (e.g., a short cylinder) and a distractor, a different shape with an unrelated property (e.g., a wavy line). On the contrary, in the No-contrast condition, the contrast object was substituted by a second distractor. Examples of the four objects in each adjective type in the Contrast/No-Contrast conditions are shown in Figure 5.

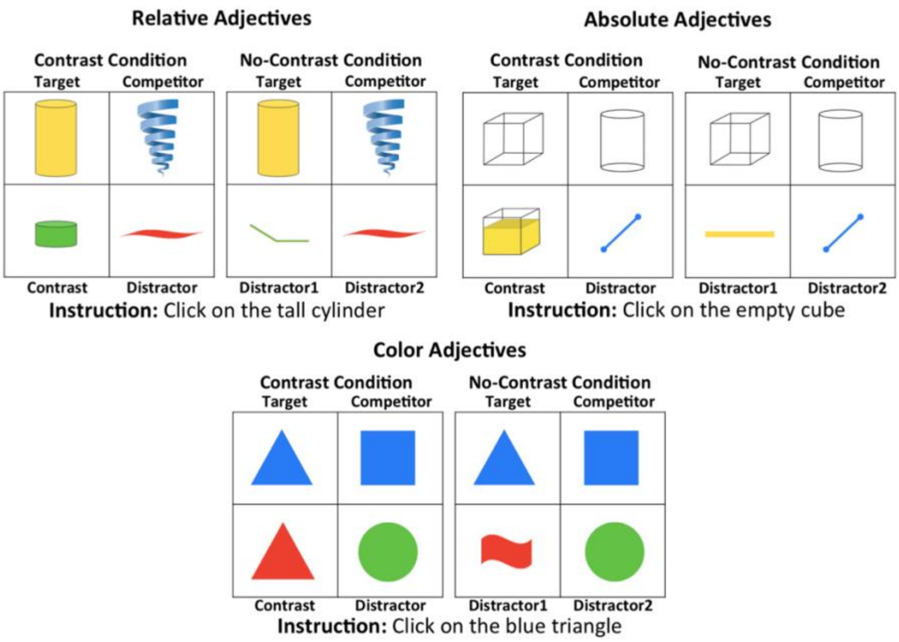


Figure 5. Trial examples of the visual stimuli used in the Contrast and in the No-Contrast conditions, divided by adjective class in Aparicio et al. (2015)'s study.

Aparicio et al. (2015) predicted that all adjectives would at least show an informativity-based effect of contrast and that relative adjectives would be integrated early thanks to the presence of the contrastive object, since the computation of relative thresholds should be facilitated by a contextually supplied comparison class. As for absolute adjectives, if the thresholds are set via the same mechanisms as relative adjectives, eye-movements should be the same. Nevertheless, a divergence between relative and absolute adjectives would suggest that context-sensitivity of these two kinds of adjectives is driven by different mechanisms.

Results showed that target identification was faster in the Contrast condition than in the No-Contrast condition for all three classes of adjectives. However, the contrast effect for each adjective class varies significantly. For color and relative adjectives, the benefit from the contrast takes place in the adjective-window, whereas for absolute adjectives it occurs during the noun-window. To interpret their results, Aparicio et al (2015) used color-adjectives as a baseline, since it was assumed that color adjectives do not evoke a contextually salient comparison class in their interpretation. However, the effect of contrast was observed also for this adjective class. They took these results to suggest that the contrast effect is driven by informativity-based pragmatic reasoning, which lags behind the lexical semantic processing identifying the adjectival property itself. The reason why these results are not consistent with those reported in Sedivy (2003), where the effect of contrast was only found for prototypical color adjectives (e.g., yellow for a banana), is still to be understood.

The contrast effect found for relative adjectives was interpreted as mostly driven by the presence of a contrasting object, which provides a contextual comparison class facilitating the computation of the adjective threshold. Although the effect of contrast was observed in the adjective window, it occurred later than for color-adjectives. This was interpreted as evidence for the fact that the contrast effect has a difference source: "the contrasting object provided a contextually salient comparison class and facilitated the lexical semantic processing for relative adjectives" (Aparicio et al., 2015: 429). Absolute adjectives showed the most complex pattern of results. Indeed, the effect of contrast only occurred during the noun-window, when also the information about the object category was available. Thus, the interpretation of absolute adjectives is overall delayed with respect to the other two classes. Like relative adjectives, they show context-sensitivity. However, they are unique in that they involve higher processing costs and significant effects did not take place until the head-noun information was available.

Aparicio et al. (2015) concluded that relative and absolute adjectives interact with the context in different ways. While for relative adjectives the contextual information is integrated early as part of the lexical processing, for absolute adjectives a different picture emerges. Like relative adjectives, absolute adjectives show sensitivity to the context. However, additional processing costs in their interpretation might be due to the absence of a comparison class variance (see van Rooij, 2011). Indeed, absolute adjectives' interpretation remains constant across contexts and is always precise and endpoint-oriented. "From a processing perspective, committing to a precise (or close to precise) interpretation of the absolute predicate is costlier than committing to an imprecise interpretation, as a variety of contexts can support an imprecise interpretation, but very few situations can actually license a precise one". Thus, these results are compatible with assuming that absolute thresholds are context-independent (see §2.5.3) and contextual sensitivity is only attributable to imprecision calculation[24].

In conclusion, prenominal adjectives have often been exploited in experiments investigating adults' online language processing. These studies demonstrated that adults process language incrementally and that, in the interpretation of noun-adjective combinations, they never wait for the noun when the prenominal adjective is informative. Furthermore, various studies showed that the effect of contrast facilitates target identification. Studies examining different semantic classes of adjectives provided inconsistent results for intersective adjectives, even though their interpretation has always been found to be easier and faster with respect to relative adjectives. Absolute adjectives, on the other hand, seem to be the most problematic and show a different context sensitivity compared to relative adjectives, resulting in additional processing costs. If and how these differences affect adjective acquisition is still to be understood and will be addressed by the present work.

2.7 Summary and conclusion

Although the majority of the literature on children's early vocabulary focuses on the class of nouns, in recent years many scholars have investigated the class of adjectives, trying to identify the reasons why they are so challenging for young toddlers who start to comprehend and use them later, with respect to nouns and verbs. The present chapter aimed to provide an overview of children's acquisition of the adjective class under different points of view.

In §2.2, we underlined how the class of adjectives differs from the class of nouns as for its lexical meaning, syntactic distribution and prosodic contour within spoken language. Unlike nouns, adjectives usually denote abstract properties of objects, making it more difficult for young toddlers who more easily identify words

[24] For further studies investigating differences between adjective classes in online and offline tasks with adults, see Frazier et al. (2008); Heller & Chambers (2014); Liao & Meskin (2015); Leffel et al. (2016); Ziegler & Pylkkänen (2016); Hansen & Chemla (2017); Aparicio et al. (2018).

for concrete objects and people. In addition, nouns are usually recognized thanks to a series of morphosyntactic cues, such as a preceding determiner, while adjectives can appear in a wider range of syntactic contexts. Furthermore, adjectives generally carry deaccented prosody.

In §2.3, we examined various different theories explaining why adjectives are harder to learn than nouns, and the subsequent delay in producing and understanding property terms. We underlined that the concepts referred to by adjectives are conceptually less salient than those labeled by nouns, that adjectives are a dependent class in that they are always combined with a referent noun and that, in learning novel words, children are guided by a series of cognitive constraints favoring nouns over adjectives.

In §2.4, we reported on corpus studies observing that nouns are massively favored with respect to adjectives both in production and in comprehension across languages. In addition, we pointed to the difficulties many researchers faced in the process of categorizing words in children's first vocabularies.

In §2.5, we turned to a review of the literature exploring adjective comprehension by young language learners. First, we reported on the studies examining the use of syntactic context in learning novel adjectives. Second, we analyzed the relevant studies investigating the integration of nouns and adjectives in combination, concentrating on the relative order of noun and adjective in different languages, on the incremental interpretation of informative vs. uninformative adjectives and on different semantic classes of adjectives.

We concluded with §2.6, an overview of the studies investigating adults' processing online, investigating the incremental interpretation of pre-nominal adjectives and the interpretation of semantically different adjective classes.

In chapter 3 we will elaborate on some common experimental methods that have been used to study language comprehension by infants and older children, namely the *Preferential-Looking Paradigm* and the *Looking-While-Listening* task. We will illustrate how, though revolutionary, these techniques present some intrinsic limits that have been overcome by the advent of the Visual World paradigm used in eye-tracking experiments. This paradigm will be presented in detail as it is the methodology employed in the experimental design of the present work.

3 The Visual World paradigm and the study of language acquisition

3.1 Introduction

Months before they produce their first words, young children reveal their developing language knowledge by responding meaningfully to the speech they hear. However, since the comprehension of the linguistic input can only be inferred through children's behavior in a specific context, receptive language competence has been less accessible than their speech production skills. In the last four decades, many valuable experimental techniques have been employed to investigate the emergence of language comprehension. Research on early cognitive abilities has examined how infants become attuned to sound patterns in the environment language over the first year (e.g., Werker & Tees, 1984; Kuhl, 2004) and how they attend to speech patterns relevant to their native language structure (e.g., Jusczyk, 1997; Saffran, 2002). These studies made it possible to explore how first year infants become skilled learners, how they are able to make distributional analyses of phonetic features of their language and how they form acoustic-phonetic representations based on frequently heard sound patterns (e.g., Hallé & de Boysson-Bardies, 1994).

It is widely accepted that children begin to learn their native language by discovering features of its sound system and not by learning words. Indeed, as opposed to sound perception, learning words is said to come later, between 6 and 15 months of age, when infants start to understand other people's intentions and reveal this progress though increasingly differentiated verbal and behavioral responses to speech. Most early scientific studies on developmental language comprehension made use of four main methodologies (Fernald, 2002). *Diary studies* provided observational data on early comprehension abilities, describing how children reacted to and interpreted the ambient language in everyday scenarios (e.g., Lewis, 1936; Bloom, 1973). Parental-report checklists are used in *Studies of vocabulary growth*, which track changes in the estimated size of the child's lexicon, which correlates with later grammatical development (e.g., Fenson et al., 2007). *Naturalistic experiments on comprehension* analyze behavioral responses to test the ability to understand familiar words (e.g., Benedict, 1979) and words in combination (e.g., Shipley et al., 1969). Lastly, *Experiments on novel word learning* constitute the largest area of research. These studies focused on how cognitive biases (e.g., Markman, 1989; see §2.3) and linguistic knowledge (e.g., Katz et al., 1974) guide inferences about word meanings. All these approaches rely on off-line measures of comprehension. In such tasks, researchers assess understanding by observing the child's behavior *after* hearing a linguistic stimulus. Moreover, the judgment that a child does or does not understand a particular word is often made by the experimenters in a controlled situation with a defined control measure, such as success in pointing to a picture given different alternatives. However, young children are known for their resistance and noncompliance, especially when asked to act on command (e.g., Brown, 1958).

Undoubtedly, the focus on language production provided a rich source for language acquisition theories (e.g., Braine, 1963; Brown, 1973; Shatz, 1978). However, such off-line tasks do not tap into real-time processing of spoken language and, as a consequence, reveal less about the child's developing efficiency and knowledge. "Language production reflected the observable half of children's language ability; comprehension was the other, inaccessible half of what children knew about language. Just as astronomers were not satisfied to study only the light side of the moon, researchers in language acquisition recognized that the dark side—language comprehension—held secrets to a process that had to be unlocked" (Galinkoff et al., 2013: 317).

To overcome these hurdles, different techniques have been developed to study language processing in real time. Among the most common, we focus on those that derived from Cooper (1974)'s simple assumption that "The probability of looking at an object increases when the object is mentioned". These experimental methodologies record participants' eye-movements during reading and listening. This has allowed research on spoken language comprehension to show that eye-movements are time-locked to lexical access of words in isolation (e.g., Allopenna et al., 1998) and to the identification of referents for syntactically ambiguous phrases (e.g., Tanenhaus et al., 1995), both among children (e.g., Trueswell et al., 1996) and adults (e.g., Brown-Schmidt et al., 2002).

In the present chapter, we provide an overview of the different methodologies exploiting the relation between language and vision to study online language processing by infants and older children, and how they have improved over time. In §3.2 we will present the *Preferential-Looking Paradigm*; §3.3 will discuss the *Looking-While-Listening* task; §3.4 will be an overview of the *Visual World Paradigm* used in eye-tracking experimental designs and in §3.5 we will discuss the rationale behind the use of such paradigm to study spoken language. §3.6 will summarize the chapter.

3.2 The *Preferential-Looking Paradigm*

In 1963, the developmental psychologist Robert Fantz published the first study using the preferential-looking method with young children, showing that newborns looked selectively at some visual stimuli (patterned images) over others (uniform images). In 1974, Horowitz asked whether visual fixations to images could be used as a window onto language development (Galinkoff et al., 2013; see Colombo & Bundy, 1981). Then Spelke (1976, 1979), in a dynamic version of Fantz's (1958, 1964) paired-comparisons method, developed the preferential-looking paradigm to study intermodal perception in infants. She presented 4-month-old infants with two visual stimuli (e.g., a person clapping hands and a donkey falling onto a table) while presenting an auditory stimulus (e.g., the sound of hands clapping) and found that children looked more at the screen in which the event matched the picture than at the screen in which it did not. Although Spelke's study did not test children's language knowledge, this inspirational study led to the adaptation of this auditory-visual matching procedure to investigate the development of language comprehension in the early years of life.

At the end of the 1970's the first experimental procedures for testing infants' knowledge of object words were introduced. Benedict (1979) found that 12-month-old children reliably directed their gaze to a familiar object when it was named, even when nonverbal behaviors of the speaker (including pointing and gaze) were eliminated, which often allowed toddlers to appear more linguistically capable than they really were (Galinkoff et al., 2013).

A second innovative study was the one conducted by Thomas and colleagues (1981), who used eye-movements as an index of word recognition comparing the ability of 11- and 13-month-old infants to identify a familiar named object among a series of competitors. Although finding that children at that age understand familiar words is not surprising, this method enabled the assessment of word recognition more objectively, unlike informal observational studies and parental reports, by exploiting as dependent measure the time spent looking at a particular item over the other (or an alternative measure such as the single longest look) (see Ambridge & Rowland, 2013). Thus, this technique allowed for the standardization of stimulus presentation, for a careful definition of what counted as a correct response and for the elimination of nonverbal cues. In addition, for questions that focus on the comprehension of verbs and events requiring motion, the advent of videotaped stimulus displays opened a new window into the exploration of language knowledge (e.g., Naigles, 1990; see Hoff, 2011).

3.2.1 The intermodal preferential-looking paradigm

The groundwork studies of Spelke (1976) and Thomas et al. (1981) provided a starting point for later research in which preferential-looking measures were further adapted to assess early language comprehension (e.g., Golinkoff et al., 1987; Reznick, 1990). More specifically, Golinkoff et al. (1987)'s version of the method, known as the *intermodal preferential-looking paradigm* (IPLP), was revolutionary in that it combined use of visual and auditory stimuli simultaneously[25], soon becoming one of the most used experimental designs using this technique.

In the procedures of these studies, infants sit on a parent's lap in the middle of two television monitors. A concealed audio speaker midway between the two monitors plays a linguistic stimulus that matches only one of the displays shown on the screens. Mounted atop the speaker is a light that comes on during each intertrial

[25] There exist also *intra*modal experiments, in in which two visual events are paired and presented with no accompanying auditory stimulus (Hirsh-Pasek & Golinkoff, 1996).

interval to ensure that the infant makes a new choice about which screen to look at on each trial. A hidden camera records the child's visual behavior (see Figure 6).

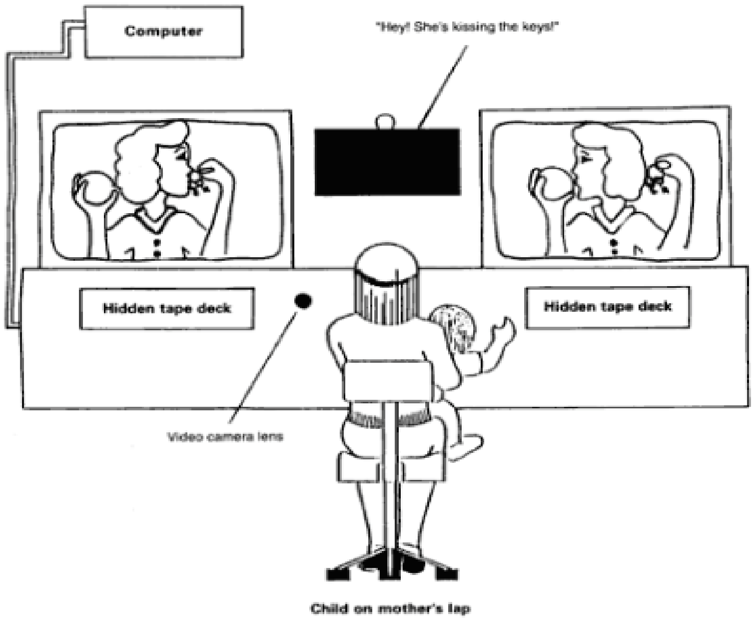


Figure 6. Example of setting for the intermodal preferential-looking paradigm used in Hirsh-Pasek and Golinkoff (1996). The infant is seated on a parent's lap in front of two televisions or one large television with a split-screen. The linguistic stimulus matches only one of the displays shown on the screen. To allow for online and offline coding, a hidden camera below the television records infants' visual fixation. Here, a woman kissing the keys vs. an apple are shown in a study on children's perception of constituent structure.

Researchers typically use the total looking time to the target picture and the duration of the longest look to the target picture as indexes of comprehension. With these measures, the IPLP allowed for the investigation of several different questions about burgeoning knowledge in the areas of phonology, semantics, syntax, and morphology in infants not yet speaking.

In the first paper using the IPLP, Golinkoff et al. (1987) conducted a series of experiments. In the first two, they examined whether 16-month-old infants could understand nouns (e.g., dog, shoe) and verbs (e.g., drink, wave). For example, in the noun experiment infants saw two static objects (e.g., a shoe and a boat) and heard, *Where's the boat? Find the boat!*. In the verb experiment, infants saw two dynamic actions carried out by the same person (e.g., a woman drinking from a coffee cup and the same woman blowing on a sheet of paper) and heard *One is drinking and one is blowing. Which one is drinking?*. The infant's task was to look at one of the two video screens. As in other studies, gaze patterns were coded in real-time through the use of a button box recording fixations on the target vs. a distractor object and shifts between the two. Both the noun and the verb experiment showed that 16-month-olds looked significantly longer at and oriented faster to objects or events matching the linguistic stimulus they heard. Interestingly, although these participants had not begun to produce any verbs, they appeared to comprehend the verbs in the experiment.

In the third experiment, Golinkoff et al. (1987) also found that 28-month-old children, who already produced multi-word sentences, could use word order in a sentence to find which member of a pair of dynamic actions matched the language they were hearing. Since the visual events were constructed to differ only by who was performing an action and who was acted upon, using verbs expressing reversable actions (e.g., *tickle*, *feed*), the task was very difficult for children who had to first analyze the visual stimuli to determine which character was the agent and which was the patient and then use the language to find the particular event described. For example, on one monitor, toddlers saw Cookie Monster tickling Big Bird from behind while Big Bird held a box of toys; on the other monitor, toddlers saw Big Bird tickling Cookie Monster. In the test trial, children heard, *Where's Cookie Monster tickling Big Bird?*. Note that since both characters were moving, children could not just look to the event where the named character was in motion to solve the task. They found that 28-month-olds looked longer at the event that matched the sentence they heard over the event that contained the same participants and same action but depicted a reversed relationship between the participants. Moreover, using a similar paradigm, Hirsh-Pasek and Golinkoff (1996) found that17-month-old infants were able to use word order in processing the described event.

These studies were the first reliable tests providing evidence of word order comprehension. Prior to these studies, researchers could only speculate about whether young children were sensitive to the grammar of their language before they actually started talking. Golinkoff et al. (1995) investigated 3-year-old ability to use Principles A and B of the Binding Theory. As was the case for other syntactic and lexical phenomena studies, the intermodal preferential-looking paradigm found evidence for comprehension of these principles earlier than most other assessments. Naigles (1990) and Naigles and Kako (1993) tested 2-year-old children's knowledge of verb meaning and sensitivity to meaning implications of transitive and intransitive sentence frames. Finally, Naigles and Gelman (1995), Fernald et al. (1994) and Golinkoff et al. (1992) have used the IPLP to investigate lexical comprehension and production.

3.2.2 Advantages and disadvantages

The intermodal preferential-looking paradigm is capable of revealing linguistic knowledge in young children for two reasons. First, unlike many other tasks used to explore language comprehension, this paradigm does not require children to point, select objects, answer questions, or act out commands. Children merely need to employ fixations in order to fulfill the task requirements. Second, the paradigm usually does not set natural cues for understanding in conflict with each other, and it does not omit the contribution of these sources. In other words, in this paradigm infants have access to syntactic, semantic, prosodic, and contextual information. When all these cues are provided, children may take advantage of what Hirsh-Pasek and Golinkoff (1996) called the "coalition" of cues, normally used in language comprehension to demonstrate the upper limits of their knowledge. Furthermore, this paradigm also made the general point that language development occurred more rapidly than previously thought. Language comprehension is ahead of language production and can be used as a vehicle to study burgeoning language knowledge (Golinkoff et al., 2013). The use of this methodology allowed researchers to find experimental evidence that infants analyze sentences they hear to find specific events in the world (e.g., Golinkoff et al., 1987; Hirsh-Pasek & Golinkoff, 1996), that they are sensitive to the grammar found in sentences (e.g., Golinkoff, 1995; Lidz et al., 2003) and even use sentence structure to glean something of the meaning of novel words (e.g., Naigles, 1990).

Although it undoubtedly is a powerful laboratory tool, like all methods of investigation, the intermodal preferential-looking paradigm has its weaknesses.

First, this paradigm can overestimate children's knowledge. This is because it always presents two alternatives, thus children could solve the task through elimination of alternatives or mutual exclusivity. In other words, children might be tested on vocabulary or sentence structure and use their knowledge to discard one alternative to find the correct alternative (e.g., Halberda, 2006; Markman & Wachtel, 1988). Some of this depends on children's age though, as infants are less likely to use this strategy (e.g., Hollich et al., 2000).

A second limitation is linked to the restricted number of items researchers can study, given children's short attention spans, and the inability of the method to investigate individual differences in grammatical development (Bates, 1993). "Although this method works well for group studies, it has proved impossible (at least so far) to adapt the preferential-looking technique for use with individual children. In the experiments they have conducted to date, Golinkoff and Hirsh-Pasek could obtain no more than four to six crucial target trials for any linguistic contrast. Although the results are quite reliable at the group level, the predicted pattern (i.e., preferential-looking at the pictures that match the language input) is typically displayed by only two thirds of the children with looking biases that average 66% for individual subjects. It should be clear why this kind of hit rate would be unacceptable for individual case studies." (Bates, 1993: 228). Note, however, that one reason children do not look exclusively at the screen depicting the event matching the picture is that the tapes are specifically designed to be equally salient and to encourage active looking. Hence, the IPLP may not be suited to study individual differences by its very design (Hirsh-Pasek & Golinkoff, 1996).

In short, the potential for use of the intermodal preferential-looking paradigm is great. Although the paradigm shows some evident weaknesses, the advantages of the paradigm seem to outweigh the disadvantages. Indeed, it has been used to test children of various ages in order to investigate a wide range of linguistic phenomena, from lexical development to syntactic development to the role of prosody in language learning. In the next section, we will overview the *looking-while-listening procedure*, an evolved version of the intermodal preferential-looking paradigm, designed to overcome its weaknesses thanks to more refined technologies.

3.3 The *Looking-While-Listening* task

Based on the studies by Thomas et al. (1981), Golinkoff et al. (1987), and Reznick (1990), in the 1990s Anne Fernald's research group developed a modified version of the preferential-looking method, the *looking-while-listening* procedure, to investigate whether particular features of child-directed speech might facilitate the identification of familiar words in fluent speech. Their initial goal in modifying the paradigm was "to increase the sensitivity, reliability, and validity of the measures, by making minor modifications to the procedure that served to eliminate confounding variables" (Fernald et al., 2008: 188). According to them, earlier preferential-looking studies might potentially confound object salience with target status since they used different stimuli as target and distractor objects. Moreover, some studies failed to counterbalance the side of target object presentation, which made it difficult to interpret infants' selective looking behavior unambiguously.

Thus, a first potentially influential change they undertook was to make sure that all target objects were also presented as distractors, to reduce the influence of object preference. Second, a major change concerned the measures used to capture infants gaze pattern in response to linguistic stimuli. Rather than coding eye-movements in real-time using a button box, they began to code eye-movement from videotapes, frame-by frame in slow motion. This change was introduced to eliminate the noise in the measurements due to the ca. 300 ms latency of the observer to press the button, a first step to achieve a greater precision, though requiring several hours of coding.

In Golinkoff et al. (1987)'s paradigm, word recognition was operationalized as a tendency to look longer at the target picture vs. the distracter, averaged over a 6-second measurement window following the offset of the linguistic stimulus. More recent psycholinguistic research with adults showed that experienced listeners can process language incrementally, generating hypotheses about the meaning on the basis of what they heard up to that moment (e.g., Tanenhaus et al., 2000) (see §2.6.1). This led Fernald's group to assume that children are simply considerably slower than adults and that a 6-s time-window after the offset of the speech stimulus was necessary to give infants the time to process the language they were hearing.

In the early studies investigating the influence of prosodic features on their ability to recognize familiar words, using this modified preferential-looking method Fernald et al. (1992) obtained counterintuitive results. When using a percent-correct measure averaged over a 6-s measurement window, English-learning 24-month-old children apparently performed less well than the 18-month-olds. When reduced to a 2-s measurement window, the predicted improvement in word recognition became clear, being around 60% for 18-month-olds and 80% for 24-month-olds. Adopting the 6-s window, which was a standard for the *intermodal preferential-looking paradigm*, accuracy was greatly underestimated. Indeed, 24-month-olds had oriented quickly to the target picture upon hearing the speech stimulus, had looked at it for 2 to 3 seconds, but then tended to look at the other picture or to look away, as they were losing interest. This behavior has to be interpreted as a sign of rapid processing, since look-backs and look-aways followed a correct response in most cases. Thus, this post-response "noise", when averaged into the percent-looking-to-target over a 6-s window, made 24-month-olds appear less accurate than 18-month-old children (Fernald et al., 2008).

At this point, two important procedural changes were made in the looking-while-listening procedure. First, they started to measure eye-movements from the onset of the target word and not from the offset. Second, they coded eye-movements at the possible finest level of resolution instead of coding eye-movements based on average looking time over an arbitrary time window (which can be suitable for one age but not another). Thus, Fernald et al. (1998) were able to code eye-movement with a 100ms resolution. In subsequent studies, resolution increased to 33ms, the duration of a single video frame, enabling a more precise measure of reaction times, being able to capture child's latency to shift from one to the other picture (see, e.g., Fernald et al., 2010). Thanks to this improvement, the looking-while-listening procedure has become an increasingly powerful method for monitoring real-time language processing, enabling to measure both accuracy and reaction time in word recognition.

3.3.1 Procedure and limits

The looking-while-listening procedure is superficially similar to the preferential-looking procedure in that infants are presented with two pictures on each trial and hear a linguistic stimulus naming one of them, as gaze patterns are recorded and manually coded frame-by-frame. What is new in this procedure is that the interest is not on a single preference score based on total looking at the target object over a time window, but rather on the time course of looking to the referent as the sentence unfolds. As Fernald et al. (2008: 190) put it, “the static notion of ‘preference’ is irrelevant for our purposes. Rather than construing infants’ looking behavior in response to spoken language as motivated by *preference*, we are interested in how children establish *reference* by making sense of spoken language from moment to moment”.

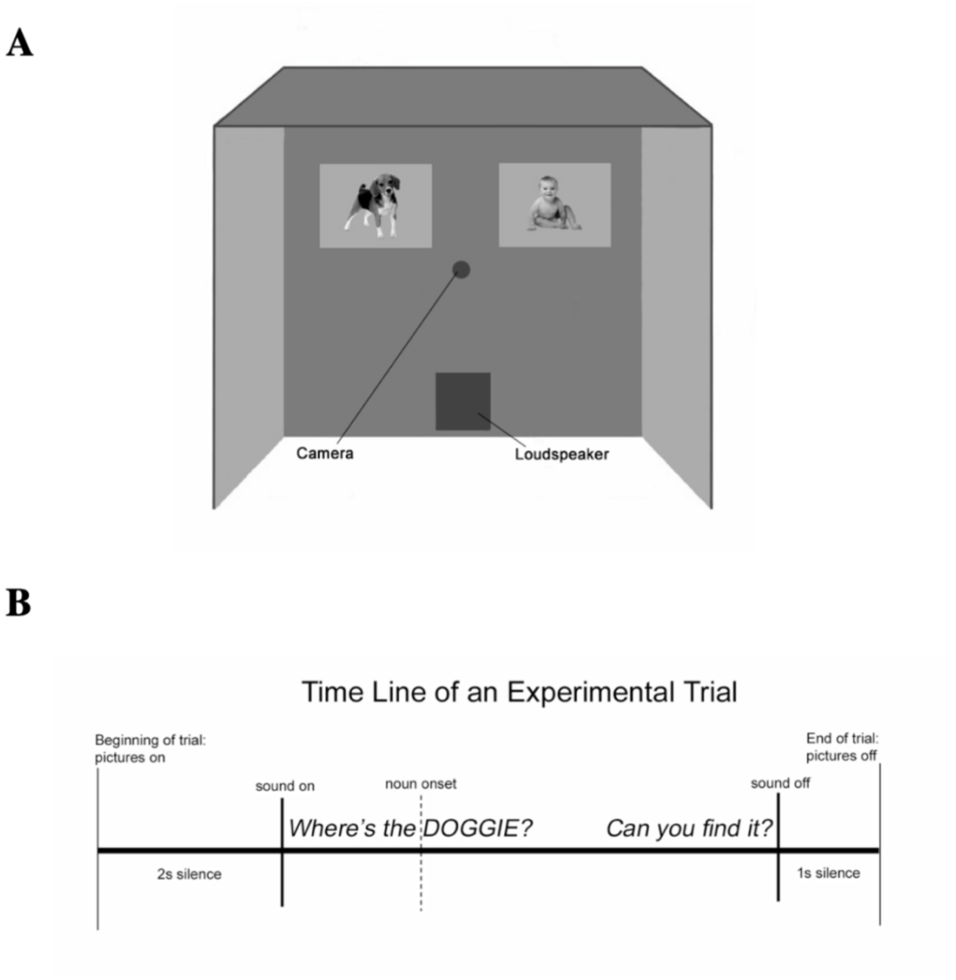


Figure 7. (A) Configuration of test booth with rear-projection screen used in the looking-while-listening procedure. (B) Schematic time line for a typical trial (Fernald et al., 2008).

Experiments using this paradigm take place in a testing room with dimmed lights. The standard procedure consists in presenting pairs of images horizontally on two different screens. The target and distractor pictures are shown for 2 seconds prior to the onset of the speech stimulus, as shown in Figure 7B. A trial is thus divided in a pre-naming and a post-naming phase. Trials last on average 6-8 seconds. The entire experiment lasts about 5-6 minutes.

In their comprehensive review of this paradigm, Fernald and colleagues (2008) listed a series of factors that need to be controlled when developing a study using this paradigm. First, both images need to be matched for size and salience. However, as demonstrated by Arias-Trejo and Plunkett (2010), choosing a distracter image which is perceptually close to the target image (e.g., a balloon paired with an egg) can result in uninvited interference effects so that 18- to 24-month-olds failed to identify the target image. A second recommendation is that across all participants both objects in a given trial should be used as target and as distractor, in order to avoid any preference for one stimulus over another. Although desirable, such a control is not always possible given the restricted choice of items in young children and the need for a sufficient number of trials per participant. In the literature, experiments using the looking-while-listening procedure have controlled for this possible preference effect (Mani & Plunkett, 2007; Swingley & Aslin, 2000, 2002), presenting the same visual stimuli at least twice, while others have not (Durrant et al., 2014; Floccia et al., 2012; Mani et al., 2008) but still found comparable results.

As already mentioned, the looking-while-listening methodology differs critically from the preferential-looking paradigm in terms of the quantitative methods used for data reduction and analysis, yielding measures of speech processing with higher resolution. However, the advent of the eye-tracking technology shows three noteworthy limits of the procedure developed by Fernald and her research groups. First, the looking-while-listening procedure typically uses visual display with only two alternatives, rather than more complex scenes involving three or more displays, which are possible with the eye-tracking technology. Second, no automated eye-tracker is used, and eye-movements are manually coded, thus resulting in lower accuracy and reliability of data analysis.

Lastly, this procedure, as well as the preferential-looking paradigm, was developed for experiments with infants and young children. This does not allow for a comparison with an adult control group which would establish a target baseline in speech processing. Thus, to investigate the relationship between language and vision the eye-tracking technology is the most suitable methodology, allowing for a greater spatial and temporal resolution and more complex visual displays. In the following section we will discuss the so-called *Visual World Paradigm*, a family of experimental methods developed for eye-tracking studies.

3.4 The *Visual World Paradigm*

In 1974, Cooper asked a group of adults to listen to a short text while looking at a display showing common objects, some of which were named in the spoken narrative. Although listeners were not given a task and were told that they were allowed to look anywhere they wanted, Cooper found that the participants' eye-movements were drawn to the objects that were mentioned or were associated with the text. For example, listeners were attracted to a picture of a dog while hearing "my scatter-brained dog Scotty..." than during other passages, or they were more likely to look at a picture of a camera when hearing "During a photographic safari...". Most interestingly, Copper found that participants' gaze was closely time-locked to the text and that more than 90% of the fixations to the critical objects took place while listening to the corresponding word or within 200ms after the word offset. He believed to have found a "practical new research tool for the real-time investigation of perceptual and cognitive processes and, in particular, for the detailed study of speech perception, memory, and language processing" (Cooper, 1974: 84). But it was only after Tanhenhaus et al. (1995)'s study using a similar methodology that the psycholinguistic community began to exploit the systematic relationship between eye-movements and speech processing on a large scale. They used a head mounted video-based eye-tracker to monitor participants' eye movements as they followed spoken instructions to pick up and move objects arranged on a table (e.g., *Put the apple that is on the towel in the box*). Their task-based approach was used influenced by previous studies which used eye movements to study vision in natural tasks (see Hayhoe & Ballard, 2005, for a review). Tanhenhaus and colleagues found evidence for rapid integration of visual and linguistic information in word recognition and syntactic processing. A few years later, Allopenna et al. (1998) developed the first study using a screen-based presentation to investigate the time-course of spoken word recognition in continuous speech and coined the term *Visual World Paradigm*.

3.4.1 Procedure and variants of the Visual World Paradigm

Setting up a Visual World experiment involves making predictions about the distribution of fixations to a target object relative to other elements in the visual display at some critical point in the speech stimulus (Sedivy, 2010). On a typical trial using this paradigm, participants hear an utterance while looking at an experimental display. During each trial, eye-movements are recorded with an eye-tracker. There are two common variants of Visual World experiments: *look-and-listen* studies and *task or action-based* studies.

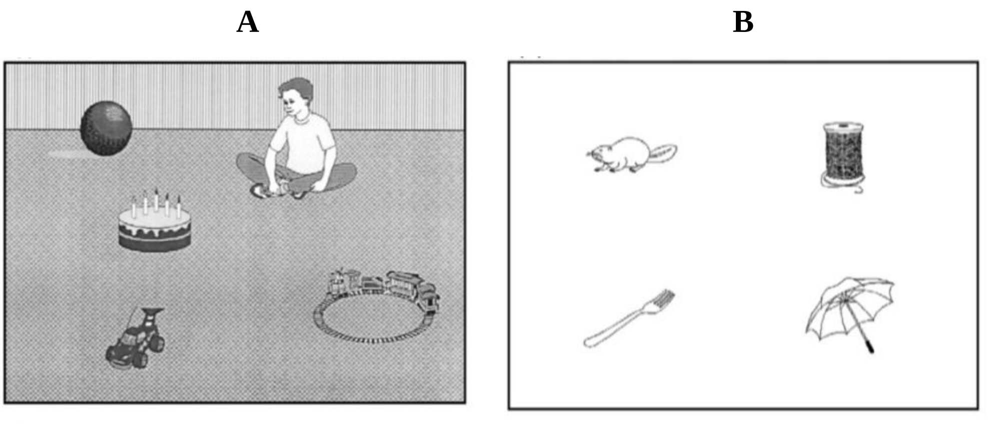


Figure 8. Examples of typical displays in Visual World *look-and-listen* studies. (A) is an example of a semi-realistic scene, in which participants either heard "The boy will eat the cake" or "The boy will move the cake" (Altmann & Kamide, 1999). (B) is an example of a four object display in which participants hear a sentence where an object is mentioned, as for instance "She looked at the fork that was in front of her" (Huettig & McQueen, 2007).

Look-and-listen studies (sometimes misleadingly called *passive listening studies*) do not require participants to perform an explicit task. In a popular version of this study, introduced by Altmann & Kamide (1999), the visual stimuli consist of line drawings of semi-realistic scenes displayed on a computer screen and auditory stimuli are utterances describing or commenting upon the pictures (e.g., "The boy will eat the cake", Altmann & Kamide, 1999), as shown in the examples in Figure 8. The screen usually shows object mentioned in the sentences (e.g., a boy and a cake) and distractors, which are never mentioned. Because the interpretation of the language is co-determined by information in the visual scene, the listener's attention is drawn to referents, including pictures that the participant anticipates will be mentioned or pictures associated with specific implied events.

Task or *action-based* studies display sets of objects, either laid out on a workspace (e.g., Tanenhaus et al., 1995; Nadig & Sedivy, 2002), or shown as drawings on a computer screen (e.g., Allopenna et al., 1998). In such tasks, participants interact with real-world objects or screen-based pictures to perform a motor task such as clicking and dragging pictures to follow explicit instructions (e.g., *Put the clown above the star*), clicking on a picture when its name is mentioned, or manipulating real objects (e.g., *Pick up the apple. Now put it in the box*) (see Figure 9).

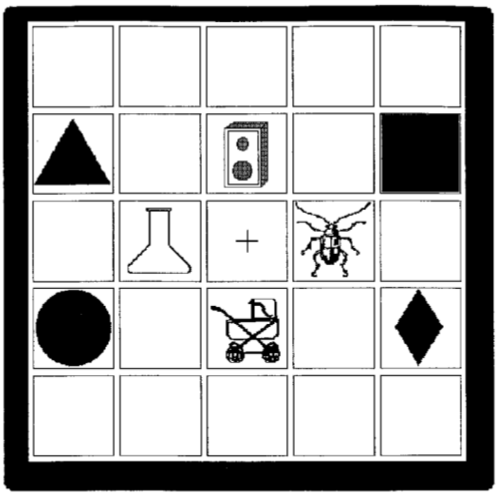


Figure 9. Example of a typical display in Visual World *Task or action-based* studies, taken from Allopenna et al. (1998). Participants were asked, e.g., *Pick up the beaker now put it above the triangle*.

Visual World studies vary as for the complexity of the visual scene presented to participants. In the simplest case, the visual scene displays the target object and one unrelated distractor object. However, to test specific hypotheses about linguistic variables, either the linguistic stimulus or the pictures on the display can themselves be systematically manipulated, creating local or temporary ambiguities. For example, in order to investigate whether 5-7 year old children would use their knowledge of the speaker's visual perspective to constrain reference, Nadig and Sedivy (2002) created 3 different visual conditions, as shown in Figure 10 below.

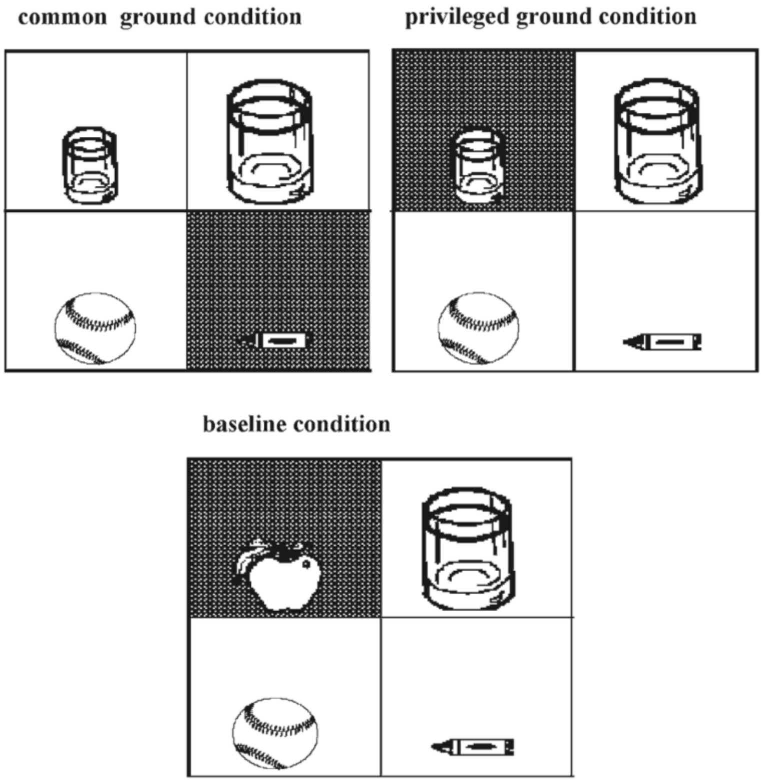


Figure 10. Sample displays for the three conditions tested in Nadig & Sedivy (2002). Shading denotes the wooden partition blocking the object from the speaker's view.

In this experiment, the child participant sat across the table from an adult speaker, each looking sat the opposite side of a vertical display case. One of the objects was blocked from the speaker's view, but not the child's (as indicated by the shaded background in Figure 10). For all three displays the speaker uttered *Pick up the glass*. While in the bottom display only one glass was present and target identification should occur right after the acoustic information, in the first two conditions a second glass was displayed, either visible or not by the speaker. This competitor objects created potential ambiguity as for the target reference, allowing for targeted predictions to be made, as for the child's taking into account the speaker's knowledge. Thus, if differences in perspective are considered through the trials, eye-movements in the privileged-ground condition should pattern similarly to the baseline. In such a design, the presence of the competitor in the top left display resulted in a globally ambiguous reference.

To avoid infelicitous stimuli (as in the common-ground condition above), many eye-tracking studies introduce *temporary* referential ambiguities. For example, in Sedivy et al. (1999)'s study (fully discussed in §2.5.2.4), the display contained two tall objects (e.g., a glass and a pitcher, see Figure 11), with a target instruction such as *Pick up the tall glass*. Here, the instruction as a whole refers to a single object (i.e., the tall glass) and is perfectly felicitous. However, a temporary ambiguity is created as *tall* can refer to two items, thus influencing eye-movements during the *tall*-window.

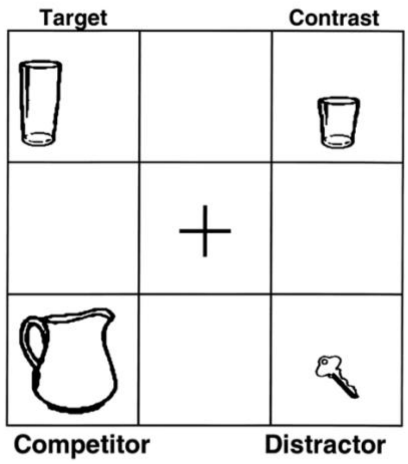


Figure 11. Example of the visual display used by Sedivy et al. (1999).

In conclusion, using this task allows not only for the analysis of the final offline responses. On the contrary, observing eye-movements opens a window on the unconscious interpretation of a sentence given a specific visual context. As a matter of fact, one great advantage of the Visual World Paradigm, compared to other psycholinguistic paradigms, is that listeners do not need to perform any meta-linguistic judgments, which might be difficult or impossible to elicit from some groups of listeners, including young children. Indeed, this paradigm solely relies on the listeners' tendency to direct their gaze towards relevant parts of the display as they are mentioned. We turn now to the technology used to run Visual World studies, i.e., eye-tracking.

3.4.2 The eye-tracking technology

The eye-tracker is a device using high-definition cameras in combination with near infrared lights. For the purposes of the research in the linguistic field, it records the position of the pupil and the corneal reflection, i.e., the reflection of a point-light projected by the near infrared light onto the participant's eyeball (e.g., Holmqvist et al., 2011, see Figure 12). As the participant's eye moves, the pupil moves but the corneal reflection remains constant. Measuring the distance between these two points makes it possible to calculate the position of the eyeball within the head (Wass, 2016).

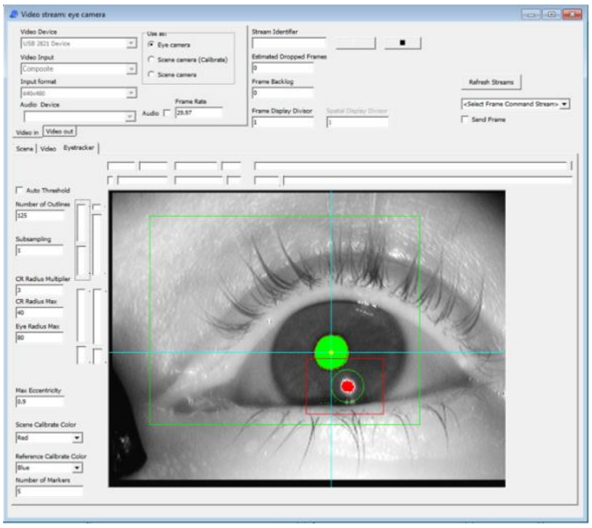


Figure 12. Example of the user-view from the MoBIGaze eye-tracking system. The pupil is highlighted in green. The little red dot underneath is the corneal reflection, projected from the eye-tracker into the participant's eye (from https://sccn.ucsd.edu/).

The camera records eye-movements in real time either mounted on a headband and placed near the eye (head-mounted eye-tracker), remotely from a desktop camera or embedded within a computer screen (screen-based eye-tracker). In the present work, we will only take into account screen-based eye-trackers.

Screen based eye-tracking involves presenting stimuli on a computer screen while measuring where the participant is looking. Within screen-based eye-tracking, two approaches are widely used, i.e., head-fixed and head-free. The head-fixed modality requires the participant to use a chinrest during tracking, to keep a constant distance between the eye and the screen. Although this technique offers the best temporal and spatial resolution, it is not widely used in developmental research, due to young children's troubles in maintaining their head still on a chinrest.

In contrast, head-free eye trackers allow participants to freely move their head in 3D space during the recordings. By setting this remote modality, the eye-tracker tracks the position of the head during the experimental session including information about the distance of the head from the screen while calculating the location of the pupil and the corneal reflection. To do so, a target sticker is usually placed in the middle of the child's forehead, allowing the eye-tracker to have a third reference during calculations.

Whenever we perform everyday tasks concerning vision, shifts of attention are accompanied by shifts in gaze. These shifts are accomplished by ballistic moments, called *saccades*, which bring the attended region into the central area of the fovea where visual acuity is greatest, resulting in *fixations* (Tanenhaus & Brown-Schmidt, 2008). The eye-tracker output provides data about both saccades and fixations.

In addition, there exist other types of eye movements such as smooth pursuit (when the eyes move continuously to maintain an object at fixation), vergence eye movements (when the eyeballs rotate towards one another in order to maintain fixation on an object as it moves closer to us), vestibular-ocular reflex movements (when the eyes make compensatory movements as the head moves, in order to keep them stably oriented) and optokinetic nystagmus movements (when the eyes are kept stably oriented over large-scale movements of the environment) (see Henderson, 2006). Although these parameters are important, most of the literature has examined fixations and saccades almost exclusively.

Figure 13 shows an example of experimental setting during an eye-tracking experiment with children.

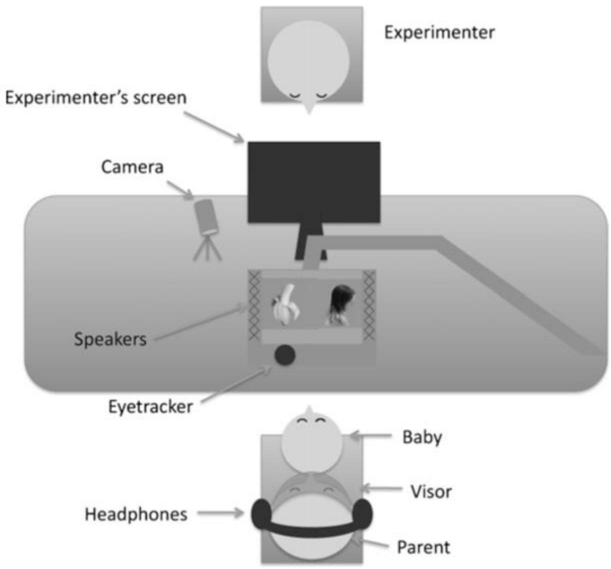


Figure 13. Example of the experimental setup used in Bergelson and Swingley (2012). The child sat on her parent's lap and was presented with images on a screen and sounds from a computer equipped with an eye tracker and speakers. The researcher sat behind a screen and was not visible to the child. The experimenter controlled the presentation of stimuli and monitored the child on a live-feed camera.

All eye-tracking experiments start with a calibration phase to establish a mapping of screen coordinates and measurements. This is achieved by asking the participant to fixate a sequence of calibration points. In each of these points, the position of the pupil is recorded, so that the eye-tracker "learns" that when a participant is looking at, say, the middle of the screen ('point X'), the distance between the pupil and the corneal reflection is 'vector Y'. This allows the system to pair information about the relative position of the pupil and corneal reflections in the eye with specific spatial locations. If, during the experiment, the distance between the two points is 'vector Y', we can conclude that the participants were looking at 'point X'. At any point after the calibration phase, it is thus possible to estimate where the participant is looking within a given scene, and gaze position is available in screen coordinates. Usually, the calibration phase is followed by a validation run, aimed at determining whether the estimated eye position is close to the known coordinates. In addition, it is a recommended common practice, especially with young children, to present additional validation points in between the trials (the so-called drift-corrections) and to recalibrate in case of failure.

To assess what the participant was looking at throughout the trial, the experimenter defines interest areas (or regions of interest) in the visual word, each associated with one of more objects. An automated coding procedure then scores each fixation as directed at one area of interest or not directed at any. As for saccades, since they are triggered by a shift in visual attention to a new location, that location can be considered the locus of attention during a saccade (Salverda & Tanenhaus, 2017).

3.4.3 The analysis of eye-tracking data

The eye-tracker records eye-movements data continuously from the very beginning of the speech stimulus, although it is sometimes useful to start recording even before the linguistic stimulus starts, in order to be able to determine whether there are any biases in gaze prior to stimulus presentation. The eye-tracking system automatically generates an output data file, logging the precise timing and location of the relevant events.

Regardless of all the measures that can possibly be reported and analyzed (e.g., total number of fixations to the target object or latencies for first fixations to the target object[26]), it is usually useful to get a temporally detailed, qualitative picture of the continuous distribution of eye movements over the objects in the visual display as the speech stimulus unfolds (Sedivy, 2010). These are generally plotted by identifying very fine time bins (e.g., 50 or 100 millisecond time "slices" containing averaged eye movement data) and, for each bin, by computing an average across all subjects and items separately for each area of interest in the screen. In this way, it is possible to get the proportion of time in that bin spent fixating each object in the visual array. Thus, a proportion of fixations plot represents, at each moment in time, the proportion of trials with a look to each picture, averaged across participants (or items). Proportion-of-fixation plots usually present data aligned to a relevant linguistic event, which typically requires a further alignment across trials. Moreover, in evaluating the data, it is important to take into account that information in the speech signal influences eye movements with a

[26] For a rigorous and detailed description of all the measures that can be recorded with the eye-tracker, the reader is referred to Salverda and Tanenhaus (2017).

delay of approximately 200-250ms (Salverda et al., 2014). Figure 14 shows an example of the fixation proportions computed for the four interest areas.

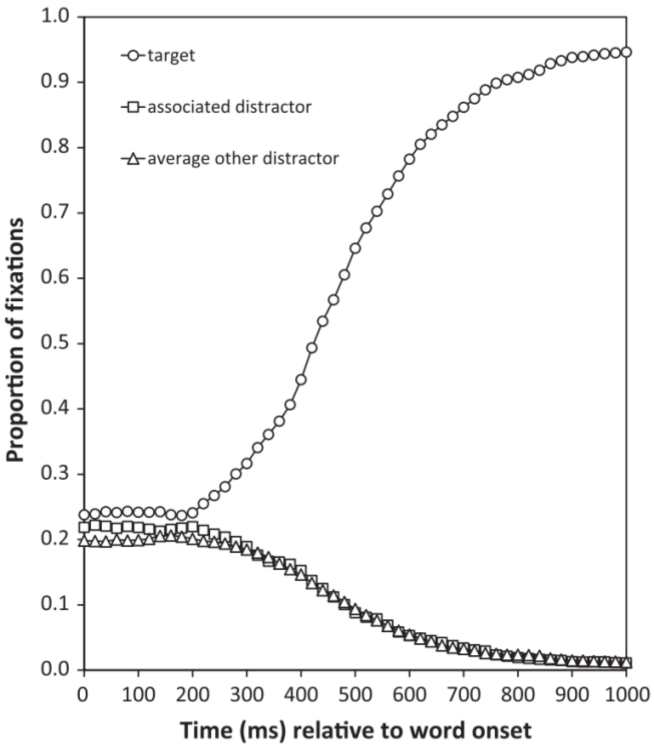


Figure 14. Proportion of fixations to the target, the associated distractor, and the averaged other distractors in Salverda et al. (2014). In this experiment, the participant saw a display with a target picture and three distractors and followed a simple spoken instruction to click on the target.

Presenting data in this qualitative manner often brings to light relevant patters in the data that may be missed by the mere observation of measures such as the total number of fixations. The resulting graphs allow for a detailed picture of the participant's shifts of attention as the speech unfolded in time.

Visual word eye-movement data can be analyzed with a range of statistical analyses. Rather than reporting separate analyses for each time bin, which could result in an increased potential for spuriously significant results, researchers often choose to create time-windows over which the eye movement data are averaged, and then submitted for statistical analysis. Traditionally, separate analyses of variance were conducted for each bin or time-window (i.e., ANOVAs), using subject and items as random factors, and with uncorrelated independent variables. Only measures on the target were specified as dependent variable. However, during the last decade, linear mixed models (Kliegl et al., 2007; Baayen et al., 2008) reduced the number of analyses to be run, making it possible to specify subjects and items as crossed random effects in one single analysis. These models returned considerable responsibility to the data analyst, concerning hypotheses and fixed and random effects specification, as well as within-subject and within-items effects (Bates et al., 2015). In addition, linear mixed models handled the problem of missing-at-random eye-movement data records and broke down the distinction between experimental and "correlational" analysis, allowing for the analysis of interactions (Salverda & Tanenhaus, 2017).

In conclusion, all the aforementioned advances in statistical inference are possible because eye-tracking yields a very high density of behavioral observations. The brief description of this technology, if compared with video-cameras used for the tasks described in §3.2 and §3.3, makes it clear that the eye-tracking provides temporal and spatial sensitivity to a much higher level of precision. Moreover, the eye-tracking technology allows for theories of representation to be integrated with theories of processing, providing a clearer picture of how and when adults, children and atypical populations integrate various kinds of information. In the next section, we turn to the rationale behind the observation and evaluation of eye movements in studying language processing.

3.5 Rationale for the use of eye-tracking to study language processing

In order to make it apparent why the use of eye-tracking can be successfully applied across a wide range of age (including children), Sevidy (2010) identified five basic assumptions behind the use of this technology.

The first assumption, as already mentioned, is that people tend to direct their gaze to things and events they are attending in their visual environment. Indeed, as we view a scene, although mostly unconsciously, we make numerous saccades shifting from one point to another in a visual display. Such saccades also occur as a result of incoming language stimuli, in response to changes in information cognitively available. Thus, linguistic studies using eye-tracking tools exploit human's natural tendency to look at pictures or words corresponding to the language they are hearing (Ambridge & Rowland, 2013).

Second, eye movement behavior is driven primarily by the goal of establishing reference. "The process of interpreting language is assumed to involve a cognitive mapping of linguistic expressions to things in the world, making language a prime candidate for one of the internal cognitive factors that drive saccadic behavior" (Sedivy, 2010: 116). This assumption is the result of what Cooper (1974) first and Tanhenhaus et al. (1995) and Eberhard et al. (1995) later (among others) observed (see §3.4). As Sevidy (2010) explains, a referent might be described as *the starred red square* to make a distinction among other similar objects in a visual context; as each individual word accumulates in the speech stream, people shift their attention. However, instead of simply shifting from the starred objects to the red objects to the set of square objects, people's gaze shows a successive narrowing of the set of potential candidates for the complex expression they are hearing. By doing so, upon hearing *square*, the square objects that do not bear the other two properties are disregarded. This does not suggest that the referential interpretation must be completed in order for a saccade to be triggered. On the contrary, saccades occurring on the basis of partial interpretations provide an insight of the underlying cognitive mechanisms involved in language interpretation.

The third assumption is that eye movements reflect incremental linguistic interpretations. Unlike traditional off-line methodologies (e.g., grammaticality judgment tasks), which usually ask participants to consciously respond to linguistic stimuli after their presentation, paradigms using eye-tracking measure the participants' unconscious and automatic response to language stimuli. The main advantage is that these types of tasks are largely immune to metalinguistic abilities. Indeed, participants cannot use their explicit knowledge about language because they do not have the time to think about the meaning or grammaticality of the sentences they are exposed to. Consequently, comprehension tasks using eye-tracking measure how participants process sentences as they unfold (Marinis, 2012). Thus, measuring eye-movements appear to be quite "cheap". Indeed, they are triggered by very partial commitment during language interpretation and provide a precise moment-by-moment window into language processing. Numerous studies have shown how subjects often begin to direct their gaze towards potential referents for a word within a few hundred milliseconds after its onset (e.g., Allopenna et al., 1998) and that they make predictive anticipatory saccades for upcoming events, even prior to the onset of the disambiguating word. For instance, Altmann & Kamide, 1999 (see Figure 8A) presented participants with visual scenes depicting, for example, a boy, a cake, and some toys while listening to sentences such as “The boy will move the cake” or “The boy will eat the cake”. They found that eye movements to the cake, which was the only edible object in the scene, started significantly earlier in the “eat” condition than in the “move” condition. Altmann and Kamide (1999) interpreted this result as evidence that the information conveyed by a verb can be used to predict an upcoming argument.

The fourth essential assumption is that eye movements reflect the output of (complex) linguistic computations. As a matter of fact, experimental evidence shows how eye movements reflect interpretive processes hinging on the representation of syntactic structures. For example, to investigate if and when listeners make use of visual information in interpreting language, Tanenhaus et al. (1995) presented participants with sentences such as “Put the apple on the towel in the box”, where the first prepositional phrase (in this case, “on the towel”) is temporarily ambiguous between denoting the destination of the apple or its current location. In the *one-referent* condition participants saw just one apple on a towel, an empty towel, a box, and a pencil, while the *two-referent* condition showed two apples: one on a towel and one on a napkin. In this second condition, a modifier was needed to inform the listener which of the two apples should have been moved. Tanenhaus et al. (1995) found that there were significantly more early looks to the empty towel in the *one-referent* than in the *two-referent* condition, indicating that "on the towel" was immediately interpreted as a modifier, not as a destination. In other words, referential support (i.e., having two apples present in the visual array, one of which is on a towel) facilitated disambiguation, enabling listeners to correctly interpret the phrase in real-time. On the contrary, when the situation is left temporarily ambiguous (i.e., when just one apple was present) adults’ visual fixation patterns reflect their initial misinterpretation. This seminal study demonstrated that “with well-defined tasks, eye movements can be used to observe under natural conditions the rapid mental processes that underlie spoken language comprehension” (Tanenhaus et al., 1995: 1634).

The fifth assumption is related to the developmental continuity of what stated so far (see Trueswell (2008) for a full discussion of related questions). To date, the literature suggests a strong developmental continuity with respect to these assumptions. More specifically, there is reason to believe that children’s eye-movement behavior, from a young age, is attention-bound (e.g., McMurray & Aslin, 2004), referentially-driven (e.g., Nadig & Sedivy, 2002), incremental at the level of the word (e.g., Thorpe & Fernald, 2006; Fernald et al., 2008) and the sentence (e.g., Trueswell et al., 1999; Tribushinina & Mak, 2016). Nevertheless, some aspects of the visual system do continue to mature throughout childhood (Sedivy, 2010).

3.6 Summary

In understanding spoken language, humans are naturally inclined to look for a visual reference corresponding to the language they are hearing and constraining the intended meaning of utterances. Thus, what we look at reveals a great deal about what is going on in our minds.

The present chapter aimed at presenting the Visual World Paradigm and its evolution from the first linguistic experiments investigating the relationship between language and vision with young children. In §3.2 and §3.3 we have seen that, to investigate when and how young listeners make use of visual information in language comprehension, the *Intermodal Preferential-Looking Paradigm* before, and the *Looking-While-Listening* task after, started to be used. Although both methods contributed to the literature on acquisition with interesting and far-reaching results, they present some intrinsic limitations, being suitable for a limited number of research questions, task designs and populations tested.

In §3.4 we have observed how the introduction of the Visual World Paradigm in eye-tracking experiments provided a sensitive and time-locked response measure, exploiting the close temporal link between gaze and cognition to study the fast and highly automatic processes involved in language processing.

We have highlighted that a particular advantage of the Visual World Paradigm is its versatility: it has successfully been used with participants of all ages, including infants and atypical populations, to investigate a wide range of psycholinguistics questions. Furthermore, it allows for fine-grained observations about language processing in real-time due to the high levels of precision and sensitivity of eye-trackers. Such a methodology is useful to measure several variables while participants are exposed to visual scenes with a wide range of complexity. In §3.5 we concluded by presenting a series of assumptions behind the use of such a methodology to study language processing.

In conclusion, the Visual World Paradigm presents some limitations in both the forms and types of questions that can be asked. For instance, this paradigm cannot evidently be used for the study of linguistic structures that are not at least partially related to the Visual World or to events and entities that are not easily depictable. Furthermore, because the Visual World Paradigm creates a restricted set of possible referents, it might introduce task-specific strategies that do not reflect "natural" language processing. However, such an effect has never been found in studies directly addressing this issue (e.g., Dahan et al., 2001; Magnuson et al., 2007). The eye-tracking technology is, to date, the most reliable tool to investigate online language processing in combination with vision.

PART II

Introduction

Part II of this dissertation reports three eye-tracking studies investigating the online processing of Italian adjectives by children and adults.

Before presenting the studies in detail, some considerations need to be made about the development of the experimental design and the data collection. This research project was developed between 2018 and 2019 and envisaged the realization of eye-tracking experiments in Italian and German. This led us to establish a cotutelle agreement with the University of Konstanz, as our initial goal was to exploit the well-equipped laboratories in Konstanz to test German children and adults, as well as Italian-German early bilinguals. Indeed, Konstanz is a South-German town with a high percentage of Italian immigrants growing bilingual children, who actively participate in various research activities at the Linguistics department. Data collection in Italy was planned to begin in early 2020, followed by the data collection in Germany in September 2020. The Covid-19 outbreak and the related situation of public health emergency in both countries have interrupted the participants' recruitment. The testing was completed at the beginning of 2021 in Italy and was canceled in Germany. This emergency has determined a revision of the original objectives and a consequent adaption of the research questions to just one language of investigation halfway through the doctoral program.

A total of 41 children and 24 adults participated in the testing. Children were recruited either in kindergartens in the areas around Verona and Brescia (North-East of Italy) or via an online subject pool. All of them were native Italian speakers from families where Italian was the dominant language. They had normal vision and hearing, and none of them reported any history of speech or language disorders. Caregivers gave complete written and informed consent and were asked to complete a questionnaire indicating whether the child knew and understood the words in the experiments. The checklist included a total of 80 nouns and 35 adjectives. The questionnaire was used to indirectly access children's vocabulary. As a criterion for inclusion to the test, we determined that children were required to know at least 75% of the nouns and 75% of the adjectives. Following this criterion, none of the children was excluded. Families received 15 euros for participation, while children received a book of their choice. Adults were recruited from Verona, Brescia, and their surrounding areas (North-East of Italy). All had normal or corrected-to-normal vision. No history of language or hearing disorder was reported. They gave written and informed consent and were paid 7 euros for participation.

Children and adults participated in the experiments in one testing session which lasted approximately 40 minutes. For adults, the order of the administered tests was randomized. For children, the order of Experiments 1 and 3 was randomized, while Experiment 2 was always conducted last. This was done because Experiment 2 was the longest and most tiring and we wanted to avoid dropouts after the completion of one test only. This resulted in a lower number of child participants in Experiment 2 (N = 28), compared to Experiment 1 (N = 38) and Experiment 2 (N = 39). Favorably, this did not affect age-range, nor the mean age in Experiment 2 (2;4-5;2, mean age = 3;7) in relation to Experiment 1 (2;4-5;3, mean age = 3;6) and Experiment 3 (2;4-5;3, mean age = 3;6).

The next chapters will report on the three experiments conducted for this dissertation. Building upon the literature overviewed in chapter 2, and in light of the considerations on the eye-tracking technology made in chapter 3, the present work investigates adjective acquisition and processing from multiple perspectives.

First, as discussed in §2.5.2.1, the interpretation of nouns and adjectives in combination has rarely been investigated using refined online technologies and never in languages with noun-adjective order, such as Italian. In an offline picture-pointing task, Ninio (2004) presented Hebrew-learning children with four pictures of two objects crossed with two properties (e.g., a big teddy, a small teddy, a big clock and a small clock), and asked children "Where is the Noun-Adj?". To account for the difficulties children showed in integrating noun and adjective meanings, Ninio (2004) formulated the *Noun-Anchor Hypothesis* (Weisleder & Fernald, 2009). According to this hypothesis, children interpret noun-adjective combinations in two-steps, by first identifying the set of objects labeled by the noun, and later identifying the subset of those object possessing the property labeled by adjective. Although various studies re-tabled this issue in different ways (i.e., Thorpe et al., 2006; Fernald et al., 2010), this hypothesis has never been tested using a paradigm similar to Ninio (2004)'s, measuring eye-movements online. Second, the acquisition of intersective, relative and absolute adjectives has been tested offline in several studies (as reported in §2.5.3) but has never been framed in terms of processing costs due to their different semantic statuses within a developmental online study.

Building upon these results, together with the findings on adults' processing of adjective classes reported in §2.6, we designed Experiment 1 (Ch. 4). We tested Italian monolingual children in an eye-tracking study with intersective (e.g., *black*), relative (e.g., *big*) and absolute (e.g., *closed*) adjectives, to answer the following research questions:

RQ1. Is there evidence of a *Noun-Anchoring* mechanism in children's online processing of Italian noun-adjective combinations? Do children focus on the interpretation of the noun-word while omitting or delaying the integration of the adjective meaning? Do children and adults differ in processing noun-adjective combinations?

RQ2. Are there differences in the way children interpret different semantic classes of adjectives in real-time? Are children sensitive to contextual cues affecting the interpretation of each adjective class? Does adults' processing vary across the different classes of adjectives?

RQ3. With respect to the first two research questions, do children show developmental differences in processing noun-adjective combinations? Do 2-, 3- and 4-year-old children differ in processing different classes of adjectives?

Second, as pointed out in §2.5.2.4, Weisleder and Fernald (2009) suggested an alternative proposal, accounting for the asymmetry between noun and adjective interpretation, the *Sequential-Integration Hypothesis*. According to this view, in interpreting noun-adjective combinations, children younger than 366 months might be disrupted by the second word in the combination, be it noun (as in Spanish) or adjective (as in English), if the first word is already informative about the referent object. They tested Spanish-learning children, by presenting them with a visual context showing pairs of objects (e.g., a red car and a blue plane) and asked them the Spanish translation of "Where is the car blue?". What the *Sequential-Integration*

Hypothesis predicts is that, despite in the *blue car – red plane* context the noun is already informative about which picture is the target, the adjective might cause disruption and children might still be looking back and forth the two pictures after the informative noun. Results showed no such effect of disruption due to the over-informative post-nominal adjective and failed to provide evidence of the hypothesized mechanism.

Experiment 2 (Ch. 5) was conducted to test the *Sequential-Integration Hypothesis* on a different language, namely Italian, and by measuring eye-movements with the eye-tracking technology. Experiment 2 addressed the following research questions:

RQ1. Do children show a *sequential-integration* pattern in interpreting color-adjectives combined with informative and non-informative nouns? Are they disrupted by uninformative postnominal adjectives?

RQ2. Do children and adults differ in the pattern of interpretation with respect to speed and accuracy?

RQ3. Are there differences between 2-, 3- and 4-year-old children in processing noun-adjective combinations across different informativeness conditions?

Third, as discussed in §2.5.2.5, Tribushinina & Mak (2016)'s Visual World study reported results against the *Noun-Anchor Hypothesis*. They tested Dutch-speaking children in an eye-tracking study where two objects on the screen (e.g., a grey butterfly and a grey stone) were paired with an adjective-noun phrase. In such combinations the pre-nominal adjective was either informative (e.g., *heavy*) or uninformative (e.g., *grey*) about the following noun (e.g., *stone*). All informative adjectives expressed properties that were prototypical about the noun and never visible on the screen (e.g., *heavy*). They found that, in the informative condition, children were able to predict the noun based on the adjective meaning and identified the target object during the adjective-window, without waiting for the noun. This finding was interpreted as evidence against a *Noun-Anchoring* mechanism in the interpretation of adjective-noun occurrences, as children processed the combination incrementally. Starting from Tribushinina & Mak (2016)'s findings, Experiment 3

(Ch. 6) addresses the interpretation of Italian adjectives in predicative use by asking yes-no questions containing a prenominal adjective followed by a noun. Simultaneously, two pictures are displayed on the screen. We investigate children's interpretation of informative vs. uninformative adjectives preceding nouns, including both inherent (e.g., *soft*) and non-inherent (e.g., *closed*) properties of the objects on the screen. Experiment 3 aims at answering the following questions:

RQ1. Are children able to predict the noun based on the pre-nominal adjective in Italian predicative constructions? Do they show a *Noun-Anchoring* mechanism in the interpretation of prenominal adjectives? Does their looking pattern differ from the adults'?

RQ2. Are there differences between the interpretation of adjectives expressing an inherent property of the object and do not have a visual representation (e.g., *soft*, for a pillow), and adjectives that are visible on the screen (e.g., *open*, for a window)? Do adults show processing differences in this respect?

RQ3. If children and adults differ with respect to the first two research questions, do 2, 3- and 4-year-old children show developmental differences in processing informative vs. uninformative adjectives? And in the interpretation of inherent vs. non-inherent properties of objects?

We turn to chapters 4, 5 and 6 where we will present the three experiments in detail.

4 Experiment 1: Processing subsective adjectives in development

4.1 Introduction

There is ample evidence showing that adults process speech incrementally and that they effortlessly integrate linguistic knowledge with information from the visual scene (e.g., Allopenna et al., 1998; Altmann & Kamide, 1999; Trueswell & Tanenhaus, 2005, among others). By contrast, the question of whether young children also process language continuously as the speech signal unfolds has not been widely investigated and only a few studies have addressed this issue.

For many years, studies on language comprehension relied primarily on offline-assessments, and language development was investigated on the basis of parent-report measures and corpora analyses. As already discussed in chapter 3, the refinement of online methodologies, such as eye-tracking, enabled a finer and detailed investigation of such developmental change. Being suitable for use with young children, the eye-tracking technology revealed a continuous development in young language learners' processing of spoken language, allowing for fine-grained exploration of speed, accuracy and emergence of adult-like performance.

The purpose of the present work is to measure children's looking behavior in speech processing of noun-adjective combinations in real time. As discussed in chapter 2, the literature on adjective acquisition has shown that property words are particularly challenging for young children for several reasons. As Gentner (1982) explained, the denotation of words like *big* and *good* can vary considerably in relation to the different nouns they modify; for this reason, understanding the meanings of adjectives depends on linguistic knowledge to a greater extent than is the case for concrete nouns. According to this view, referred to as *Relational Relativity*, adjectives, verbs, prepositions, and other relational terms are, with respect to concrete nouns, cross-linguistically more variable in the way they map from concept to word (see §2.3). In contrast, concrete and proper nouns refer to things that are perceptually coherent and easily individuated even by infants (Markman, 1989). Moreover, while nouns are used to label categories of objects sharing many correlated features, adjectives have far fewer shared correlated features because they shift semantically in different linguistic contexts (Markman, 1989). While all the things we call *apple* tend to share many perceptual attributes (e.g., size, shape, material, or function), this is not the case for the things we call *red* (Gasser and Smith, 1998). This conceptual complexity that adjectives show compared to concrete nouns could account in part for the fact that children are typically slower to learn property terms than object terms (Gentner & Boroditsky, 2001). Adjectives may also be challenging for young learners because property terms occur less frequently in speech to children than nouns and verbs and, consequently, children have less experience with this class of words (Sandhofer et al., 2000). These are factors that, among many others deeply discussed in §2.4, may help explaining children's difficulties with the word class of adjectives.

As discussed in §2.5.1, developmental studies have investigated at what age and under what conditions children can correctly identify the property indicated by the adjective. They mainly focused on when and how children correctly interpret novel adjectives as property terms and how young children interpret novel adjectives in interaction with the nouns they modify (e.g., Gelman & Markman, 1984; Hall & Graham, 1999; Klibanoff & Waxman, 1999; Mintz & Gleitman, 2002; Mintz, 2005). Other studies focused on how children extend novel adjectives to other same-category or different-category objects (e.g., Klibanoff & Waxman, 2000; Waxman & Klibanoff, 2000). However, the processes involved in learning novel adjectives and mapping a single word to a relevant property may significantly differ from the processes involved in the rapid interpretation of familiar noun-adjective phrases in continuous speech. Crucially, children's difficulty in the interpretation of noun-adjective combinations has rarely been framed in terms of processing costs and challenges at the level of phrasal integration.

4.1.1 Processing attributive adjectives

Processing attributive adjectives seems to be particularly problematic for young children. In her study on Hebrew, Ninio (2004) proposed that the difficulty children show in the interpretation of such combinations depends on the computational complexity of the integration of noun and adjective meanings. Using a picture-pointing task, Ninio (2004) asked toddlers to point to a *big teddy* in a set of four items that crossed the relevant property with the object kind (e.g., big and small fishes and teddies) (see §2.5.2.1). Although the children were fairly accurate in identifying the correct picture, a more detailed analysis of first responses and self-corrections showed that, when making a mistake, children always selected the other object (e.g., the other teddy) and not the other property (e.g., the other big object). To account for these results, Ninio proposed the ***Noun-Anchor Hypothesis*** (Weisleder & Fernald, 2009; see §2.5.2.1), according to which young children have difficulties in interpreting noun-adjective combinations because they require a two-step process: first, there is the identification of the set of objects labeled by the noun and second, the subcategorization of that set containing those objects that also possess the property labeled by the adjective. In other words, interpreting the phrase *big teddy* implies identifying the set of teddies in the discourse context, and then the subset of the teddies that are big. This two-step procedure has been automatized in adults and older children who are fast and accurate but it seems to be particularly challenging for toddlers younger than 36-months, who focus on the identification of a suitable referent for the noun and omit or delay the second step of adjectival meaning integration. Thus, young children perform an *overt* two-step procedure. By contrast, older children and adults perform the two steps as one, instantly and without awareness. Ninio argued that learning how to fuse the two-step procedure into a single step is a major developmental task along the way to the mastery of attribution that occurs at around 36 months of age.

Although Ninio's (2004) experimental findings were very clear, her research could not account for cross-linguistic variation concerning adjective position in attributive use. Indeed, children were tested in Hebrew, a language in which adjectives are positioned post-nominally. Thus, the sequential steps of interpretation she proposes are aligned with the order of nouns and adjectives that are actually heard in the tested language. For speakers of a language with prenominal adjectives, this model could involve additional processing costs. If noun-adjective combinations are interpreted in two steps, in a noun-first and adjective-second order, then for children speaking an adjective-noun language the interpretation would be even more demanding. Indeed, they would have to hold the prenominal adjective in working memory, wait until hearing the noun before beginning the interpretation of the whole combination, and then retrieving the adjective for the identification of the proper referent. Hence, this would make the same task relatively more challenging for, e.g., English-speaking children than for their Hebrew-speaking peers.

To explore this issue, Thorpe et al. (2006) conducted a replication of Ninio (2004)'s study on English, exploiting the same off-line methodology (see §2.5.2.2). Their results were in line with those found in Hebrew and the youngest children showed the same noun-biased error pattern of the original study. This same pattern was interpreted by Ninio as evidence of the two-step process that children go through, focusing on nouns and partially neglecting adjectives, resulting in the failure of identification of the correct picture. However, using a picture-pointing task only allows for the access to offline measures, i.e., accuracy, but cannot provide any information about the real-time processing of noun-adjective combinations.

Fernald et al. (2010) partially addressed this issue deploying the Looking-While-Listening paradigm (see §2.5.2.3). Participants sat in front of a screen displaying a visual array of two pictures, while listening to familiar adjective-noun phrases (e.g., *Where is the blue car?*) in a visual context where the color-adjective was either informative (e.g., a blue car paired with a red car) or uninformative (e.g., a blue car paired with a blue house). The analysis of participants' gaze pattern revealed that, while 30-month-old English-learning children failed to take advantage of the prenominal color word to identify the target even if it was informative, 36-month-olds were more successful and showed an impressive gain in the ability to integrate nouns and adjectives in combination in real time. The authors concluded that children develop the skill to process color-adjectives and nouns in combination over the third year and that it is possible that, returning to Ninio (2004)'s hypothesis, 30-month-olds waited for the noun before interpreting the property term. For this reason, since the noun could refer to either of the two objects on the screen, children looked back and forth between the pictures until remembering the adjective, which finally allowed them to identify the target.

The evidence presented above might suggest that having pre-nominal adjectives may be sub-optimal from a processing perspective. Encountering the adjective in post-nominal position (as in Hebrew or Spanish) should allow children to integrate noun and adjective meanings as they are being heard, eliminating the extra challenge of holding the adjective in memory until after hearing the noun. However, Weisleder and Fernald (2009) argued that the same problem could occur with postnominal adjectives. For instance, if Spanish-speaking children were exposed to a blue-car/blue-house condition and hear *¿Dónde está el carro azul?* ('Where is the car blue?') they might face an analogous challenge: the informative word (e.g., car), which offers enough information to identify the target, is followed by an uninformative adjective (e.g., blue) which is relevant for both pictures and might disrupt children's processing. Weisleder and Fernald (2009) referred to this possible interpretation as the ***Sequential-Integration Hypothesis*** (see §2.5.2.4). According to this view, children have difficulties with the sequential integration of any two content words, thus the integration is easier when the informative word comes last, as this arrangement is less costly on working memory. Thus, according to this hypothesis, in the blue-car/blue-house situation, English-speaking children can easily listen through the uninformative adjective and wait for the noun for the interpretation. Conversely, Spanish learners hear the informative noun first and must hold it in memory while listening to the uninformative adjective. On the other hand, according to the *Noun-Anchor Hypothesis*, the difficulty shown by English learners might be due to the asymmetry between nouns and adjectives and noun-adjective combinations are always easier when the noun comes first, guaranteeing an overall advantage to Spanish learners.

To understand which hypothesis could account for the results found in English, Weisleder and Fernald (2009) conducted the same experiment with Spanish-speaking children. More specifically, they aimed at understanding whether children's processing of adjective-noun vs. noun-adjective phrases is affected by the order in which the relevant information is heard and by cross-linguetic differences. They found that 36- to 40-month-old Spanish-learning children interpreted the speech stimuli incrementally: when the noun was informative (i.e., in the blue-car/blue-house condition), children were able to identify the target early and did not appear to suffer any disruption from the color-adjective. When the informative word was the color-adjective (i.e., in the blue-car/red-car condition), children listened through the whole combination and shifted their gaze to the target after the adjective word. Thus, comparing these results with those found for English-learners, children's processing of phrases with color adjectives seems to be facilitated when the noun is heard before the adjective. However, Ninio's hypothesis proposes that difficulties with the integration of noun and adjective meanings concern children younger than 36 months, that were not tested in Weisleder and Fernald (2009)'s study on Spanish. Thus, no result could provide evidence in favor or against the *Noun-Anchor Hypothesis.*

Tribushinina and Mak (2016) addressed the issue of incremental interpretation by young children in a Visual World eye-tracking study (see §2.5.2.5). 3-year-old Dutch speaking children were presented with 2 pictures on the screen (e.g., a grey stone and a grey butterfly) while hearing an adjective-noun phrase that was either informative (e.g., the heavy stone) or uninformative (e.g., the grey stone) on the adjective. Eye-movement analysis showed that, in the informative condition, children were able to identify the target object after hearing the adjective word, without the need to wait for the following noun. As we highlighted in §2.5.2.5, in this task children could have attended to the adjective and used conceptual knowledge of the target object, i.e., knowing that a stone is typically heavy whereas a competitor (e.g., a butterfly) is not. Adjective-noun co-occurrence statistics are also likely to have influenced the early looking behavior. Although Tribushinina and Mak (2016) concluded that their results discard the *Noun-Anchor Hypothesis*, we believe that they did not directly demonstrate adjective-noun integration and, as pointed out for Weisleder and Fernald (2009)'s results, they only tested children older than 36 months.

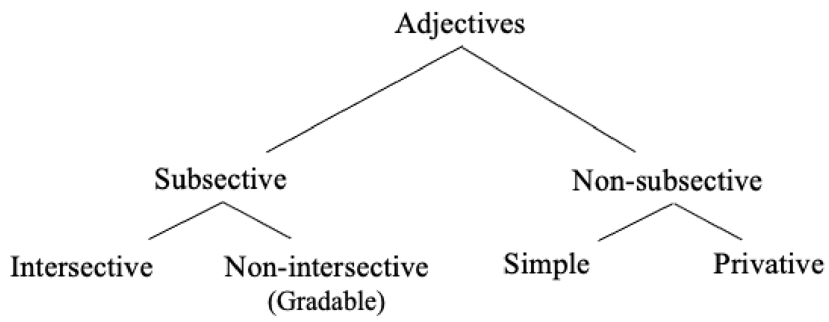
Leaving aside the noun-first / adjective-later pattern expected by the two-step process proposed by Ninio (2004), a further prediction deriving from the *Noun-Anchor Hypothesis* is that postnominal frames result in more efficient processing than prenominal frames. To test children's integration of referring expressions in pre-nominal vs. post-nominal adjective conditions, Davies et al. (2021) designed an eye-tracking study with English-speaking 3-year-olds. Children were presented with four black-and-white pictures on the screen (e.g., a big cow, a small cow, a big flower and a small tree) and eye-movements were recorded while listening to a question in two syntactic conditions. In the pre-nominal condition, children heard an adjective-noun phrase (e.g., *Where is the big cow?*), whereas in the post-nominal condition, the adjective was uttered in a relative clause (e.g., *Where is the cow that is big?*). Interestingly, they found that, like adults, children processed modified noun phrases equally quickly regardless of adjective position. We believe this finding is more robust than previous research in which experimental paradigms meant that the referential task could be passed using adjective information alone (Thorpe et al., 2006; Fernald et al., 2010) or world knowledge (Tribushinina & Mak, 2016). However, as Davies et al. (2021) underline, the use of relative adjectives only (i.e., *big* and *small*) has limited the outcome of the experiment. Moreover, once again, they only tested 3-year-olds and, consequently, nothing can be said about younger children's online processing of such combinations.

Quite intuitively, what emerges from this brief overview is that the literature on the topic is scarce, the methods are inconsistent, and further investigation is needed. The present study aims at addressing the interpretation of nouns and adjectives in combination in a Visual World paradigm using the eye-tracking technology with toddlers between two and five years of age. We are interested in children's looking behavior during the integration of noun and adjective meanings, both in terms of processing costs and taking into account semantic differences among subsective adjectives.

4.1.2 Processing different semantic classes of adjectives

As illustrated in §1.4.2.2, formal semantic accounts involving adjectives point to the difficulty of providing a characterization of adjectival meaning. For the aim of the present work, we will adopt Kamp and Partee (1995)'s classification of adjectives into two broad categories (i.e., subsective and non-subsective adjectives), focusing on subsective adjectives only (i.e., intersective and gradable adjectives). (87) reports a summary of this classification (see §1.4.2.1 for a detailed discussion).

(87)



As discussed in §1.4.2.1, Kamp and Partee's classification is modeled in set-theoretic terms and is based on the inferences that a noun-adjective combination (or intersection) triggers. Among the so-called subsective adjectives, the simplest category includes adjectives with the most stable core meanings, referred to as the class of (subsective) **intersective adjectives**. Under a classical model-theoretic view of semantic representation, the meaning of such adjectives can be identified with the set of entities that bear the property encoded by the adjective. Thus, the meaning of the word *red* corresponds to the entities that have the property of being red. When combined with nouns, it is easy to see how simple expressions can compose into more complex ones. For instance, a strawberry can be described as a "red fruit" and the meaning of this combination is simply the intersection between the set of objects labeled by the noun (e.g., fruits) and the set of properties labeled by the adjective (e.g., red). Various studies showed that color-adjectives can be context-dependent to different degrees (see, e.g., Kennedy & McNally, 2010). Nevertheless, this type of adjectives is usually interpretable independently of the context and we assume do not evoke a contextually salient comparison class in their interpretation. On the other hand, it is well known that many adjectives do not conform to the simple compositional analysis described above. Indeed, treating noun-adjective combinations as an intersection of predicates only works for those adjectives with an independent and stable core meaning. However, a variety of adjectives (such as *big* or *good*) appear to lack an invariant meaning: they depend upon the noun they modify and can only be evaluated with respect to the set denoted by the head noun. Thus, the set denoted by the noun-adjective combination is necessarily a subset of the set labeled by the noun; for this reason, they are called (subsective) gradable adjectives.

Non-intersective (or gradable adjectives, Kennedy, 1999) have context sensitive meanings (Kennedy & McNally, 2005; Kennedy, 2007; McNally, 2011; among many others) and map the object they refer to onto a scale of fully ordered degrees (Kennedy, 1999; Kennedy & McNally, 2005; Kennedy, 2007). According to how context influences their meaning and according to the structural features of the scales, a further distinction needs to be made between **relative adjectives** and **absolute adjectives**.

Relative adjectives (such as *big*, *long* or *old*) are context-sensitive in the sense that the context determines how much size is required to count as e.g., big, long or old and the standard of comparison is located around the midpoint of the scale. Moreover, relative adjectives are mapped onto "open" scales, i.e., with neither a lower nor an upper boundary. On the contrary, for gradable adjectives like *clean* or *empty*, the context determines how much deviation from e.g., total cleanness or emptiness is allowed to count as clean or empty. They evoke a scale that is "closed" at one end (e.g., *clean/dirty*) or both ends (*empty/full*). Indeed, a particular cloth can be dirty with no conceivable limit of dirtiness, but there is a standard of cleanness that cannot be overcome; however, a particular glass can be "fuller" than another, but there is a limit in both directions, because a glass cannot overcome what counts as completely full or completely empty.

Both the context and the presence/absence of boundaries on a scale determine how gradable adjectives are interpreted. Hence, relative adjectives, such as *big*, need to be interpreted according to a standard of comparison which is contextually fixed: an object counts as *big* only relative to a standard of size that can be contextually retrieved. Absolute adjectives, on the other hand, evoke a bounded scale, and that boundary serves as the standard: an object counts as clean only if it possesses the maximum standard of cleanness. If the gradable adjective has two boundaries (e.g., *empty*), both ends could constitute the standard.

4.1.2.1 Experimental studies on intersective, relative and absolute adjectives

In the past years, various studies have investigated the interpretation of these adjective classes by children and adults. However, most of them used offline methodologies (e.g., Syrett et al., 2006, Syrett, 2007, Foppolo & Panzeri, 2013) or focused only on adult processing (e.g., Sedivy et al., 1999, Aparicio et al., 2015).

Most of the studies conducted with children made use of the Scalar Judgment Task, an experimental paradigm that measures offline responses. In a typical procedure of these studies, 3- to 5-year olds saw a series of seven objects displaying different degrees of the same property. Participants were asked to judge whether the property labeled by the adjective in question was true for each of the seven objects ('Is this ADJ.?', e.g., *Is this full?*). Results from different languages confirmed that relative and absolute adjectives are interpreted in different ways, i.e., the cut-off point was identified in the middle of the scale for relative adjectives (see Syrett, 2007 for English, Foppolo & Panzeri ,2013 for Italian, Tribushinina, 2013 for Dutch, Weicker & Schultz, 2018 for German) and on one boundary for absolute adjectives (Syrett, 2007, Foppolo & Panzeri, 2013, Weicker & Schultz, 2018). Although limited in the experimental methodology, the findings reported in these studies suggested that absolute adjectives share properties with both relative and intersective adjectives. Indeed, similarly to relative adjectives, absolute adjectives are, to different degrees, context sensitive. On the other hand, similarly to intersective adjectives, objects described by absolute adjectives are always judged as either possessing the property or not and, unlike objects described by a relative adjective which are in the middle of a scale, are never vague.

Fewer studies have addressed the issue of how these semantic differences affect the online interpretation of these different types of adjectives and, at our knowledge, none has explored children's processing in this respect. Sedivy et al. (1999) made use of a Visual World paradigm with eye-tracking to investigate how the presence of a comparison class in the visual display influences the adult processing of sentences containing relative adjectives. Participants heard a sentence such as *Pick up the tall glass* while presented with four objects simultaneously. They tested two conditions, namely the Contrast and the No-Contrast condition. In the Contrast condition, the visual scene displayed the target (e.g., a tall glass), an object-competitor (e.g., a short glass), a property-competitor (i.e., another tall object) and a distractor (i.e., an object that could not be described neither by the noun nor by the adjective). In the No-contrast condition, they did not include the object-competitor, but added another distractor. The main finding was that participants' fixations converged on the target faster in the Contrast condition than in the No-Contrast condition. Crucially, participants zoomed into the target object during the adjective-window when the head-noun had not been processed yet, suggesting that the information about the contrasting object was used very quickly, already at the point in which the linguistic instruction was still compatible with both the target and the property-competitor. Although in the same paper Sedivy et al. (1999) reported a similar study on intersective adjectives, they did not provide any comparison between adjective classes.

In another eye-tracking study, Aparicio et al. (2015) investigated adults' interpretation of intersective, relative and absolute adjectives used as restrictive modifiers in pre-nominal position. In an experimental design similar to the one developed by Sedivy et al. (1999), they exploited intersective adjectives as a baseline and found that when the visual context supports a restrictive interpretation of the adjective (i.e., in the Contrast condition), target identification is faster for both relative and absolute adjectives. However, for absolute adjectives, this effect is significantly delayed. They argued that target identification for relative adjectives was faster because the contrasting object provided a contextually salient comparison class and facilitated the lexical semantic processing. However, they suggested that this effect of contrast appeared later for absolute adjectives because participants were asked to commit to a precise interpretation of predicates when presented with endpoint-oriented absolute adjectives (e.g., *clean*). They argued that committing to this precise interpretation might be costlier than committing to the imprecise interpretation of absolute adjectives involving one open boundary (e.g., *dirty*) and, hence, of relative adjectives (which have two open boundaries by definition).

Although studies on adjective interpretation have addressed various issues concerning incrementality, semantic differences and processing costs, none of them has explored these aspects from the developmental point of view and in real time.

Building upon these results, the present study takes on a new approach by investigating the early development of children's ability to interpret familiar adjectives in online sentence comprehension using eye-tracking. Children and adult participants were tested in Italian, a language with post-nominal adjectives that allowed us to test a) the *Noun-Anchor Hypothesis*, b) potential differences children show in interpreting different semantic types of adjectives (i.e., intersective vs. relative vs. absolute adjectives) in real time, c) differences between children and adults at the level of speed, accuracy and processing costs and d) the emergence of adult-like patterns of interpretation in development.

4.2 The current study

The goal of the present study is to investigate how young toddlers between two and five years of age process online noun-adjective combinations in Italian and how they distinguish between different semantic types of adjectives. To this end, we created a Visual World task with eye-recordings to compare the time course of sentence comprehension during the processing of questions containing noun-adjective combinations. Using this methodology, participants' eye-movements were recorded while looking at four pictures on the screen of two objects crossed by two properties (e.g., a black shoe, a white shoe, a black sock and a white sock), while each stimulus sentence unfolded (i.e., "Where is the Noun-Adj?", e.g., *Dov'è la scarpa nera?,* lit. 'Where is the shoe black?') (see §4.3.1 for more details). Participants were tested in three adjective-conditions, namely intersective (e.g., *black*), relative (e.g., *big*) and absolute (e.g., *closed*).

The current study builds upon the literature on adjective acquisition, noun-adjective integration, and adjective semantics and was designed to answer the following research questions:

1) Is there evidence of a *Noun-Anchoring* mechanism in children's online processing of Italian noun-adjective combinations? Do children focus on the interpretation of the noun-word while omitting or delaying the integration of the adjective meaning? Do children and adults differ in processing noun-adjective combinations?

2) Are there differences in the way children interpret different semantic classes of adjectives in real-time? Are children sensitive to contextual cues affecting the interpretation of each adjective class? Does adults' processing vary across the different classes of adjectives?
3) With respect to the first two research questions, do children show developmental differences in processing noun-adjective combinations? Do 2-, 3- and 4-year-old children differ in processing different classes of adjectives?

4.2.1 Predictions

The Visual World paradigm with eye-movement recording enables us to investigate participant's interpretation of auditory stimuli as the sentence unfolds, allowing us to access the processing of nouns and adjectives as they are being heard and interpreted. Here, we briefly outline the hypothesized outcome for each research question.

4.2.1.1 Predictions for Research Question 1

The first research question (RQ1) asks whether young children process Italian noun-adjective combinations showing a *Noun-Anchoring* mechanism, i.e., whether they focus on the noun and delay or omit the interpretation of the adjective. Furthermore, we are interested in the comparison between children's and adults' processing in this respect.

According to the *Noun-Anchor Hypothesis*, noun-adjective combinations have a complex logical structure whose interpretation requires young children to go through an overt two-step procedure, during which they initially focus solely on the noun, and delay or omit the interpretation of the (postnominal) adjective. This process, however, is effortless for fluent adults, who perform it automatically and without conscious awareness. Although the *Noun-Anchor Hypothesis* was formulated based on results from an offline task, predictions can be made about the online process of interpretation of noun-adjective combinations. What we predict is that all children will interpret the noun-word as soon as they hear it, rapidly and accurately identifying the object referent. However, upon hearing the adjective, we expect the youngest children to be slower in the interpretation, showing target preference only after sentence offset or to even leave out reference resolution, i.e., to not be able to discard the object competitor and to keep looking back and forth between the two category competitors.

In general, we expect significant differences between children and adults regarding speed, accuracy, and processing costs. We expect children to fixate the target picture above chance level but not to be as fast and as accurate as adults, thus showing more difficulties in processing noun-adjective combinations. We predict adults to be overall faster than children in object identification first (i.e., at the noun), and in shifts to the target later (i.e., at the adjective), both within the time-windows in which the disambiguating word is spoken (i.e., the noun-window and the adjective-window respectively). In other words, we expect adults to direct their gaze to the spoken word *while* it is heard, without the need to wait for its offset. What is more, since this experiment is specifically designed for young children, we do not expect adults to show any difficulty in performing the task and thus, we predict that accuracy will be at ceiling in all trials, unlike for children whom we expect to be overall less accurate. Both lower speed and lower accuracy are expected as a result of additional processing costs for children with respect to experienced adults.

4.2.1.2 Predictions for Research Question 2

The second research question (RQ2) aims at understanding whether children use different strategies in the interpretation of intersective, relative and absolute adjectives, respectively, thus showing that they know that semantic differences affect the interpretation within different contexts. In addition, we ask whether adults' processing varies across the three different classes of adjectives.

We expect children to be faster in target identification with intersective adjectives (e.g., *black*), since these adjectives are not sensitive to the context, at least not in the way relative and absolute adjectives are. In other words, once identified, e.g., the black shoe, participants do not have to look for the other shoe to verify its blackness.

By contrast, we expect this to happen with relative adjectives (e.g., *big*). More specifically, since relative adjectives are context-dependent, the identification of the target (e.g., the big teddy) is only possible after a comparison between the members of the object class identified by the noun (e.g., the big teddy and the small teddy). For this reason, we predict this comparison to take more time, resulting in a delayed gaze shift to the described object.

As for absolute adjectives (e.g., *closed*), different scenarios can emerge. On the one hand, the observation of previous developmental literature suggests that children are aware that absolute adjectives share properties with both intersective and relative adjectives. As already discussed, similarly to intersective adjectives, objects described with an absolute adjective can be judged as either possessing the property or not. For example, by looking at a picture of a shirt, participants can say if that shirt is dirty or not, without the need to compare it with the other object in the context (e.g., with the clean shirt). However, like relative adjectives, absolute adjectives are to some degree context-sensitive and can be vague (e.g., a shirt can be judged as dirty, but less dirty than another one). Any difference we will observe in looking behavior between absolute and intersective, and between absolute and relative adjectives, will suggest that children know there is a difference in their semantics, thus providing experimental evidence that this knowledge is acquired very early in development.

Finally, finding significant differences between the interpretations of the three adjective classes by adults would be unexpected though revealing. As a matter of fact, the task was designed in such a way that a (young) participant has sufficient time to look at the four objects on the screen prior to the start of the auditory stimulus. Although this window might not be enough for children, who need to focus on each picture for longer time to identify its subject, adults clearly need less time. In other words, we do not expect adults to perform a comparison between members of the noun category in the interpretation of relative (nor absolute) adjectives. However, in their Visual World study with adults, Aparicio et al. (2015) found that interpreting absolute adjectives was costlier than interpreting relative (and intersective) adjectives, and they explained this finding as the result of committing to a precise interpretation of endpoint-oriented absolute adjectives (e.g., *closed*) (§2.6.2). Thus, finding a significant difference in looking pattern would be in line with Aparicio et al. (2015)'s. In contrast to Aparicio et al., our experiment was specifically designed for young children and, therefore, this effect might not emerge due to the ease of the task.

4.2.1.3 Predictions for Research Question 3

The third research question (RQ3) investigates whether children show age differences with respect to the incremental interpretation of noun-adjective combinations and in relation to the interpretation of different semantic classes of adjectives.

We hypothesize that 3- and 4-year-old children will show stronger and earlier target preferences in response to noun-adjective combinations compared to the 2-year-olds. We expect the oldest toddlers to start the interpretation of the noun during the noun-window and to identify the target by the sentence-offset, significantly before the youngest children. Although all children possess the vocabulary knowledge required in this task, we expect processing limitations and non-adult-like performances to emerge.

In relation to the semantic classes of adjectives, we expect all children to present the pattern we predicted in §4.2.1.2, i.e., to show more difficulties with relative adjectives (e.g., *big*) than with intersective (e.g., *black*) and absolute adjectives (e.g., *closed*). In line with previous offline studies, although subsective adjectives are acquired very early in development, we expect children by the age of 5 to show difficulties in integrating all adjective with noun meanings, mainly with gradable adjectives.

4.3 Methods

4.3.1 Materials

In this Visual World eye-tracking study, all items consisted of a visual stimulus paired with and an auditory stimulus.

The visual stimuli were four digitalized colored drawings on a white background presented on the screen simultaneously. The four pictures consisted of two objects or animals crossed with two attributes in a two-by-two design, e.g., a black shoe, a white shoe, a black sock and a white sock (Figure 14). The position of the objects on the screen was pseudo-randomized and counterbalanced across trials.

The auditory stimuli were recorded by an Italian female native speaker reading with pragmatically neutral intonation (processed in Praat version 6.0.49; Boersma & Weenink, 2016). All stimuli were comparable in duration (M = 1660ms). The stimuli consisted of a question about one of the pictures, e.g., *Dov'è la scarpa nera?* ('Where is the black shoe?'), containing the carrier phrase ("Where is the"), a noun and a post-nominal adjective. Thus, the visual display contained: 1) the target object (e.g., the black shoe); 2) the noun-competitor that belonged to the same object class as the target, but could not be described by the property in the sentence (e.g., the white shoe); 3) an adjective-competitor that shared the target property, but belonged to a different object class (e.g., a black sock); and 4) a distractor that did not belong to the target object class, nor could be described by the property in the sentence (e.g., a white sock). In addition, according to the adjective type, stimuli were divided into 3 adjective-conditions, namely intersective adjectives (INT), relative adjectives (REL) and absolute adjectives (ABS) (see Figure 14). Each participant was presented with 4 trials for each adjective-condition in a randomized order.

Visual and auditory stimuli were combined to form 48 trials divided into 12 lists. All stimuli are listed in Table 6.

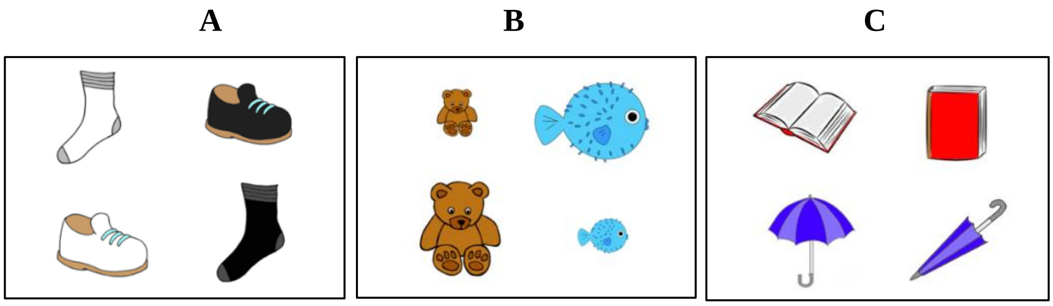


Figure 14. Example of visual stimuli: intersective (A), relative (B) and absolute (C) adjective-condition.

Table 6. List of items used in the experiment divided by adjective type.

Adj. type	Nouns	Adjectives

1	INT	*birillo, cappello*	'pin', 'hat'	*rosso, verde*	'red', 'green'
2		*calza, scarpa*	'sock', 'shoe'	*bianca, nera*	'white', 'black'
3		*cucchiaio, fiore*	'spoon', 'flower'	*giallo, viola*	'yellow', 'purple'
4		*macchina, palla*	'car', 'ball'	*blu, rosa*	'blue', 'pink'
5	REL	*albero, tavolo*	'tree', 'table'	*alto, basso*	'high', 'short'
6		*bambina, gallina*	'girl', 'hen'	*grassa, magra*	'fat', 'thin'
7		*matita, sciarpa*	'pencil', 'scarf'	*corto, lungo*	'short', 'long'
8		*pesce, orso*	'fish', 'bear'	*grande, piccolo*	'big', 'small'
9	ABS	*bicchiere, piatto*	'glass', 'plate'	*pieno, vuoto*	'full', 'empty'
10		*lampadina, torcia*	'bulb', 'flashlight'	*acceso, spento*	'bright', 'dim'
11		*libro, ombrello*	'book', 'umbrella'	*aperto, chiuso*	'open', 'closed'
12		*maglietta, mano*	'shirt', 'hand'	*pulito, sporco*	'clean', 'dirty'

4.3.2 Participants

Thirty-eight typically developing Italian children (2;4-5;3, mean age = 3;6, S.D. = 0;7) participated in the experiment; 20 were female. Three additional participants were excluded due to fussiness or inattentiveness during testing (i.e., failure to look at the four pictures on more than half of the trials). Twenty-four Italian adults (19;1-29;9, mean age = 25;4, S.D. = 2;6) served as controls; 17 were female. They had normal or corrected to normal vision by means of glasses or soft contact lenses. None of the adult participants had reported history of speech, hearing or language disorders.

The study was approved by the local Ethics Committee of the University of Verona, and was conducted in accordance with the standards specified in the 2013 Declaration of Helsinki.

4.3.3 Procedure

Participants were tested individually in a dimly lit and soundproof testing room at the Laboratory of Text, Language and Cognition (LaTeC) of the University of Verona. Participants' eye movements were recorded using an SR Research EyeLink 1000 Plus eye-tracker, which was in head-free remote mode and sampled monocularly at 500 Hz with a 16mm lens. The experiment was run on a computer connected to a 24" colour BenQ monitor for visual stimulus presentation. Speech stimuli were played over two loudspeakers on both sides of the screen. The experimental procedures were implemented in Experiment Builder and eye-movement data were extracted through Data Viewer. A small sticker on the participant's forehead tracked head movements. Calibration and validation procedures were carried out using a five-point display at the beginning of the experiment and a drift correction was repeated once every three trials.

The participants were escorted into the laboratory where the study was explained. The adult participants were seated on a chair in front of the screen and were told that they would participate in an experiment created for children and to follow the visual and auditory stimuli. They were not given a specific task but were asked to stay seated and as still as possible throughout the duration of the experimental session. The child participants sat on the caregiver's lap in front of the screen. They were told that they would play a game with a cartoon character (i.e., Peppa Pig), whose voice would tell them what to do and whose picture would appear on the screen every now and then, i.e., for calibration, validation and drift correction. Before the experiment began, a 5-point calibration and validation were performed, followed by a familiarization phase. Participants were shown 8 "warm-ups" in which a single image was labeled by a sentence played over speakers (e.g., "Look! A butterfly!") and appeared on one of the four quadrants of the screen, familiarizing the child with the four positions of the pictures on the screen. After the familiarization phase, the experimental session began. In each trial, the four pictures appeared on the screen simultaneously. The auditory stimulus started once the child fixated all four pictures on the screen or after 5 seconds. From the sentence-onset, pictures remained displayed on the screen for 3000ms, followed by an 800ms blank screen that ended the trial. An example of the trial is shown in Figure 15. A drift correction was performed every three trials, paired with one of the three filler sentences (e.g., "Are you having fun?"), recorded in a child-friendly intonation. The testing session lasted approximately 5 minutes.

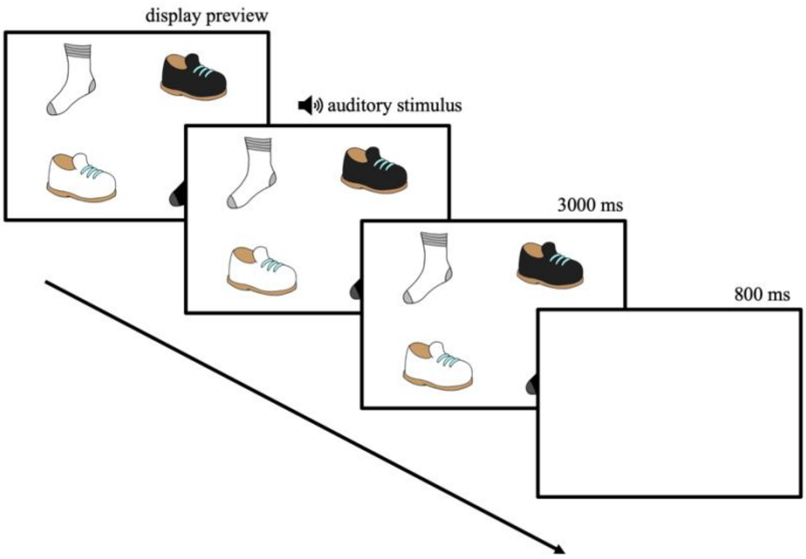


Figure 15. Experimental procedure.

4.4 Results

Eye-movement data generated by the Eyelink system were extracted through the Data Viewer software and were prepared for the statistical analysis in R (R Development Core Team, 2019). Data were analyzed using the packages lme4 and lmerTest in R (Bates et al. 2015, Kuznetsova et al. 2017, lmer function), LMERConvenienceFunctions (Tremblay & Ransijn 2015, summary function) and car (Fox & Weissberg 2011, Anova function). Before data analysis, we removed the trials in which the target picture was never fixated, neither before nor after the auditory stimulus was displayed. This led to the removal of 6,8% of the data. In addition, for each participant, we removed all trials in which none of the four pictures was fixated for more than the 50% of the duration, leading to the removal of another 6,8% of the data.

From the eye-tracking record, we determined the position of the eye in 50ms steps. For each 50ms-bin, we aggregated proportions of fixations to the four areas of interests. In other words, the analysis has been conducted only on fixations on the four pictures (target, noun-competitor, adjective-competitor and distractor). Proportion of looking time at named target picture was accessed over three time-intervals in the speech stimulus. First, the *noun-window* corresponded to the mean duration of the noun, starting from the noun onset (1000ms-1400ms). Second, the *adjective-window* corresponded to the mean duration of the adjective, starting from the average adjective onset (1450ms-1800ms). Finally, the *post noun-phrase-window* captured fixations during a 700ms window after the average offset of the adjectives (1850ms-2550ms).

The analysis of the results will be organized as follows. In §4.4.1, data will be analyzed to answer RQ1, i.e., how children process Italian noun-adjective combinations in comparison with adult controls. In §4.4.2, to answer RQ2 we will discuss the results comparing looking behavior for each adjective class, including a confrontation between the two age groups. Finally, §4.4.3 will focus on children's data and will analyze the effect of age with respect to the issues raised by RQ1 and RQ2.

4.4.1 Integration of noun and adjective meanings

The first analysis had the objective of answering the first research question (RQ1), i.e., to investigate whether children between two and five years of age process noun-adjective combinations showing a *Noun-Anchoring* mechanism, i.e., if they focus on the noun and delay or omit the interpretation of the adjective. In addition, we were interested in comparing children's and adults' behavior in this respect.

Before presenting the analysis, we comment on the Figure 16 showing the proportion of fixations to the target over the time course of the trial for children (plotted in blue) and adults (plotted in red). The vertical dotted lines are drawn according to the three time-windows identified as the areas of interests (i.e., noun-window, adjective-window and post noun-phrase-window).

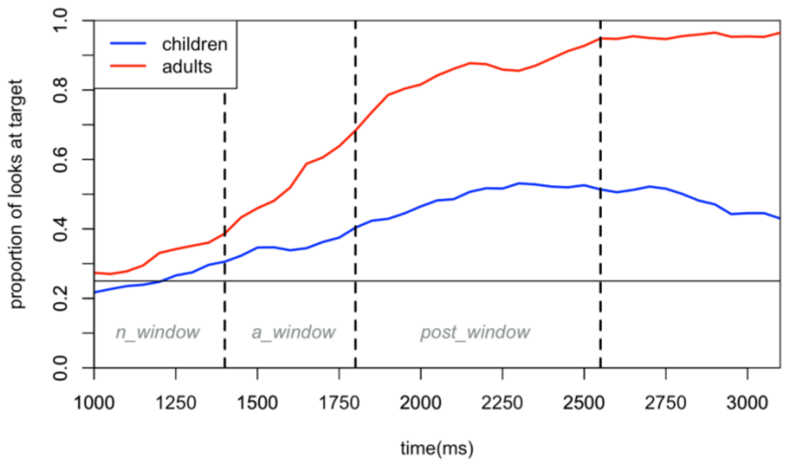


Figure 16. Proportion of looks at the target picture throughout the trial for children (in blue) and adults (in red) from the noun-onset. The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average offset of the adjective. The horizontal line indicates the chance level.

A first visual inspection of the graphs in Figure 16 offers some preliminary insights. First, looks to the target picture constantly increase as the sentence unfolds for both children and adults, though at different rates. Indeed, target preference arises earlier for adults than for children. At the noun-offset (1400ms), children's looks at target are slightly above chance level, whereas the adults' proportion target preference reaches 0.4. At the adjective offset (1800ms), the adults' proportion of looks at target are around 0.7, but the children are substantially slower, being around 0.4. In general, after hearing the whole noun-adjective combination, adults are more accurate in target identification and reach 0.95 proportion of looks at the labeled picture, whereas children do not overcome 0.5.

To answer RQ1, we performed a statistical analysis to investigate a) when participants start to interpret the noun-word and when they integrate its meaning with that of the adjective-word and b) group differences in this respect. To do so, we verified the observations reported so far in a series of linear mixed effects regression models. To investigate when looks to the named picture started to significantly increase over chance level, we performed separate analyses for each of the three time-windows. This was done because target identification was only possible *after* the onset of the adjective. During the noun-window, two pictures were potential targets, i.e., the target (e.g., the black shoe) and the object-competitor (e.g., the other shoe). Thus, chance level was 0.5 proportion of fixations. By contrast, during the adjective-window and the post noun-phrase-window, chance level decreased at 0.25.

4.4.1.1 Statistical analysis on the noun-window

To evaluate when participants start to interpret noun-adjective combinations as the sentence unfolds, we started by analyzing when eye-movements to the object-preference started to significantly diverge from fixations to the property-preference for both children and adults. Thus, we collapsed fixations to *target & noun-competitor* on the one hand (e.g., looks at the black shoe and looks at the other shoe), and fixations to the *adjective-competitor & distractor* on the other (e.g., looks at the other black object and looks at the distractor). Figure 17 plots the grand average of proportions of fixations to *target & noun-competitor* divided by *adjective-competitor & distractor* for children (in blue) and adults (in red).

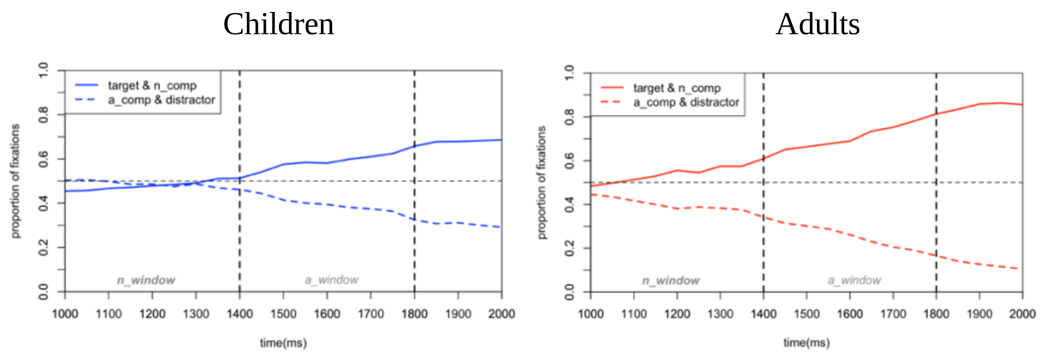


Figure 17. Proportions of fixations to target/noun-competitor divided by adjective-competitor/distractor over time for children (plotted in blue) and adults (plotted in red) from the noun-onset. The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average offset of the adjective. The horizontal line indicates the chance level[27].

A visual inspection of the graphs in Figure 17 further confirms that, while listening to the noun (i.e., in the noun-window), children's eye-movements are directed to all four pictures on the screen (i.e., proportion of fixations to *target/noun-competitor* and *adjective-competitor/distractor* is around 0.5), whereas

[27] The chance level in this graph is indicated at 0.5 as fixations to the four quadrants were aggregated into two areas of interest only. Note, however, that the statistical analysis on the aggregated fixations was conducted on the noun-window only, i.e., the time-window in which the linguistic stimulus provided information on the object category and not on its property.

adults' fixations are faster in that their looks to the *adjective-competitor/distractor* decrease below chance earlier (i.e., around 200ms after the onset of the noun). In other words, children's object-preference does not overcome chance level before the offset of the noun, whereas adults focus on the labeled objects already during the noun-window. To verify these observations, we performed a first analysis on the **noun-window** only. We calculated a linear-mixed effects regression model with fixations to target/noun-competitor as dependent variable and group as independent variable. Participants and items were added as crossed random effects (random intercepts). We found an effect of group that approached significance (ß = -0.05, SE = .03, t = -1.91, p = .06).

To investigate when in time the two groups started to significantly differ in fixations to the labeled objects we used a general additive mixed model in R (GAMMs), whose visual representation indicates when in time the effect of group becomes significant on a response variable (in this case, fixations to the labeled objects). We included a parametric coefficient for the variable group, along with a random effect for *event* (i.e., the combination of subject and item as a unique identifier), allowing for a random intercept (see e.g., Porretta et al., 2016, Zahner et al., 2019). Before calculating the model, we converted fixations to the target object to empirical logits (*elogs,* a logit transformed proportion, i.e., a ratio of the fixations to the referent object divided by the fixations directed to the competitor (Barr, 2008)). The model revealed an overall significant difference between groups (ß = .03, SE = .01, t = -3.564, p < .001), localized in the time-window: 1129 – 1800ms, i.e., from 100ms after noun onset to the end of the adjective window. This effect is shown in Figure 18 (the output of *GAMM* is only meaningful when visualized).

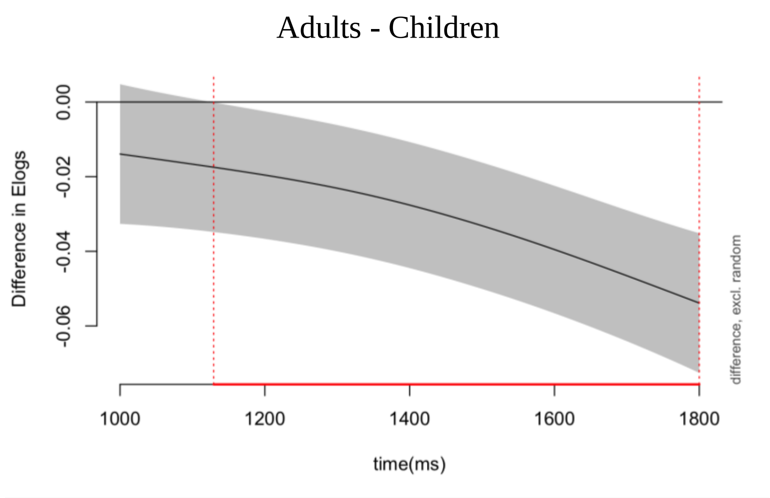


Figure 18. Difference curve in fixations to labeled objects by adults minus children in the noun- and adjective-windows. The grey band indicates the 95% confidence interval (CI) of the mean difference. Values above zero indicate more target fixations by children. Conversely, values below zero indicate more target fixations by adults. The difference is significant if the 95% CI does not include zero (1129 – 1800ms).

To sum up, the statistical analysis of looking behavior during the noun-window confirmed the observation that, while adults discard the adjective-competitor already during the noun-window, children are overall slower and less accurate, in that their weak preference for the labeled object approaches significance only at the noun offset. In other words, adults start processing the noun-word from its onset, while children from its offset.

4.4.1.2 Statistical analysis on the adjective-window

To investigate when participants start interpreting the adjective-word and integrate its meaning with the noun-word, a second analysis was performed on the adjective-window. Indeed, it was from the adjective-onset (i.e., 1400ms) that disambiguating the target was possible, by discarding the object-competitor. In line with the observations made from the graphs in Figure 17, we expect children to show difficulties in identifying the target among the two objects labeled by the noun and to keep looking at both during the interpretation of the adjective word, until the integration of noun and adjective meanings is complete. Figure 19 plots the proportion of fixations to the target (e.g., the black shoe) and the noun-competitor (i.e., the other shoe).

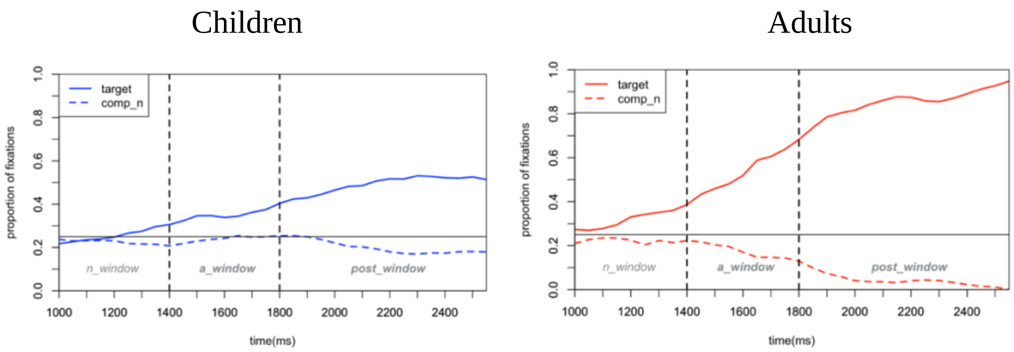


Figure 19. Proportions of fixations to the target vs. noun-competitor over time for children (plotted in blue) and adults (plotted in red) from the noun-onset. The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average offset of the adjective. The horizontal line indicates the chance level.

The graphs in Figure 19 make evident that children have difficulties in interpreting the adjective word and in combining its meaning with the noun word. Indeed, no huge difference is found between looks at the target and looks at the competitor until after the adjective-offset. Adults, by contrast, increase their looks to the target object as soon as the adjective-window begins. To verify these differences in eye-movement behavior between groups, we performed a statistical analysis on **adjective-window**. We calculated a linear-mixed effects regression model with fixations to the target as dependent variable and group as independent variable. Participants and items were added as crossed random effects (random intercepts). The model showed a significant main effect of group (ß = -0.2, SE = .03, t = -5.977, p < .001).

To evaluate when in time the effect of group started to be significant, we calculated a general additive mixed model, specifying logit-transformed fixations to the target preference as dependent variable and a parametric coefficient for the variable group. Further, we added *event* as a random effect (random intercept). The model revealed a significant difference between groups (ß = - .05, SE = .01, t = -8.844, p < .001), in the time-window 1450 – 2550ms, i.e., during the adjective-window, persisting in the post noun-phrase window (see Figure 20).

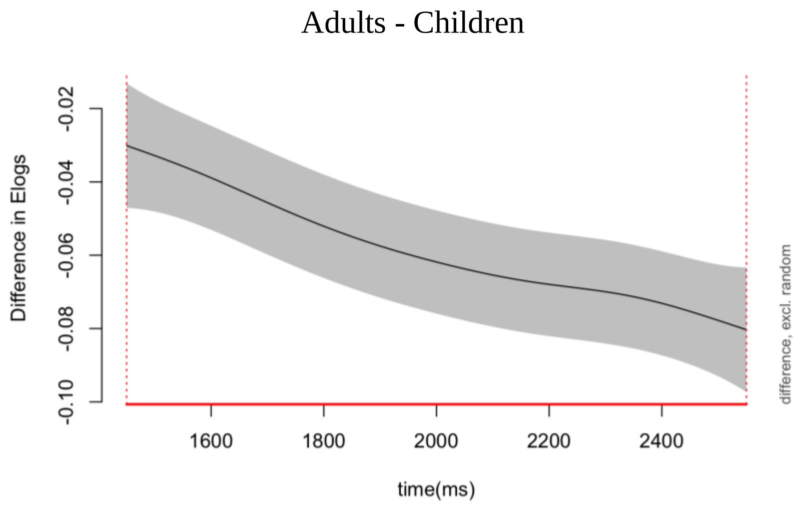


Figure 20. Difference curve in fixations to labeled objects by adults minus children from the adjective-onset. The grey band indicates the 95% confidence interval (CI) of the mean difference. Values above zero indicate more target fixations by children. Conversely, values below zero indicate more target fixations by adults. The difference is significant if the 95% CI does not include zero (1450 – 2550ms).

The statistical analysis performed on the adjective-window confirmed that adults are fast and accurate in reference resolution and that, already while hearing the adjective-word, they are able to integrate its meaning with that of the noun. Despite being slower, children started to significantly fixate more the target picture at the end of the adjective-window, showing that both nouns and adjectives are interpreted at the offset of the word, contrary to the prediction that adjectives are harder and slower to interpret (§4.2.1.1).

4.4.1.3 Statistical analysis on the competitor

The statistical analyses conducted so far confirmed the predictions that children are overall slower and less accurate in target identification. However, in line with the *Noun-Anchor Hypothesis*, we also expected to find evidence of an overt two-step process children go through in interpreting noun-adjective combinations. Figure 21 depicts the proportion of looks to each of the target, noun-competitor, and adjective-competitor for children (on the left) and adults (on the right).

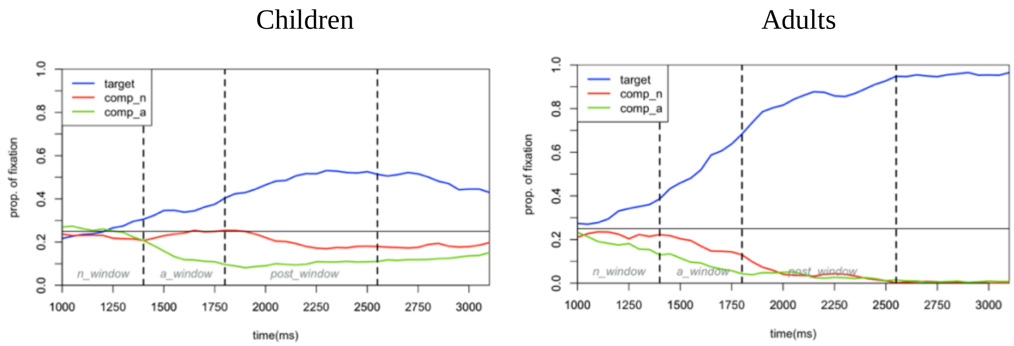


Figure 21. Proportion of looks to target, noun-competitor and adjective-competitor throughout the trial for children and adults starting from the noun-onset. The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average offset of the adjective. The horizontal line indicates the chance level.

The graphs in Figure 21 show that, unsurprisingly, during the noun-window all participants keep looking at both the target (e.g., the black shoe) and the noun-competitor (e.g., the other shoe). However, while adults rapidly discard the adjective-competitor (e.g., the other black object) as soon as they start processing the noun-word (i.e., at around 1250ms), during the noun-window children keep looking at the adjective-competitor as much as the target, without showing any preference. Only at the noun-offset, the property-competitor is rapidly discarded (as fixations drop below chance at 1500ms), while the noun-competitor is still fixated. Finally, at around 2000ms (i.e., 200ms after the adjective offset) children's looks to the object-competitor decrease below chance level.

To verify these observations on fixations to the noun-competitor, we calculated a linear mixed regression model on a broad time-window including both the adjective- and the post noun-phrase window. We specified proportion of fixations to the object-competitor as dependent variable, time-bin (in 50ms steps) as independent variable and participant and item as random intercepts. In doing so, we evaluated when looks to the object competitor started to significantly decrease below chance level after the adjective-onset. For adults, looks to the noun-competitor decreased significantly below chance already at 1600ms (ß = -0.04, SE = .02, t = -2.278, $p < .05$). For children, no significant decrease below chance was found, but the model revealed an approaching significant effect only at 2300ms (ß = -0.04, SE = .02, t = -1.852, $p = .06$).

4.4.1.4 Summary

The analyses we conducted in the present section revealed that 2- to 4-year-old children, like adults, interpret noun-adjective combinations incrementally, though at different rates. During the noun-window, children keep looking at all four pictures on the screen and do not show any preference for the labeled objects before the noun-offset. By contrast, adults immediately process the noun-word and shift their gaze towards the target and object-competitor. Moreover, during the adjective-window, adults integrate the meanings of noun and adjective as soon as the adjective word is spoken and looks to the noun-competitor immediately drop below chance. Conversely, children's looks to the target do not increase significantly before the offset of the adjective-word. However, contrary to what the *Noun-Anchor Hypothesis* predicts, children's interpretation of the adjective word is not slower than that of the noun and, above all, is not omitted.

The analyses presented so far, however, were constrained in two ways. First, they were conducted without taking into account semantic differences among adjectives. Second, possible effects of age among the group of children were not investigated. §4.4.2 and §4.4.3 respectively will address these issues.

4.4.2 Online interpretation of subsective adjectives

The second analysis had the objective of establishing whether there is a difference in the online processing of different semantic types of adjectives, i.e., to test whether children know there are differences in the way each adjective class, in combination with a noun, sets the standard of comparison. Thus, this second analysis consisted of a comparison of looks to target across time-windows, adjective-conditions, and groups.

Figure 22 shows proportions of fixations to the target picture for children and adults, divided by the different adjective-conditions (i.e., intersective, relative and absolute).

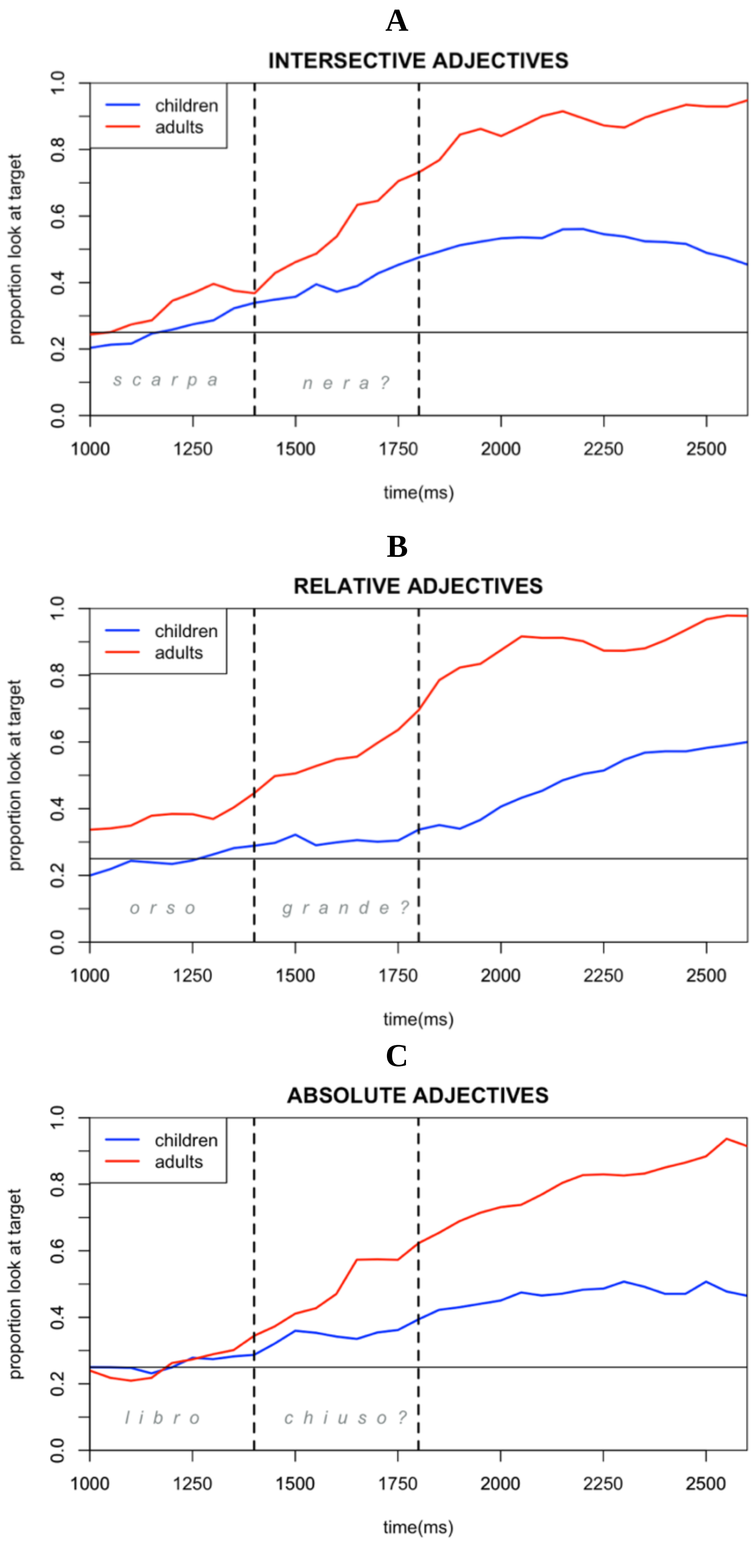


Figure 22. Proportion of looks at the target picture throughout the trial for children (in blue) and adults (in red) divided by adjective-condition; (A) shows the intersective (INT) condition, (B) the relative (REL) condition and (C) the absolute (ABS) condition. The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average offset of the adjective. The horizontal line indicates the chance level.

From a visual inspection of the graphs in Figure 22, some preliminary observations can be made. In all three adjective-conditions, adults are overall faster and more accurate in target identification. Furthermore, while adults do not seem to show different looking behaviors across adjective-conditions, children's target preference emerges at different time points across adjective-conditions. Indeed, the proportion of looks at the target picture overcomes chance level already during the adjective-window for intersective adjectives in (A) and absolute adjectives in (B), while target preference for relative adjectives in (C) emerges only 100ms after the adjective offset.

To investigate which variables significantly influenced fixations at the target object, data were analyzed using several mixed effect models. Since we were interested in the effect of adjective semantics on the interpretation of the combination, data analysis was conducted on the adjective- and post noun-phrase windows. First, we specified fixations at target as dependent variable, group and adjective-condition as fixed effects with full interaction and group and time-window as fixed effects with full interaction. Participant and item were added as crossed random effects (random intercepts). We found a significant interaction between group and adjective-condition (χ^2 = 25.1, df = 2, $p < 0.001$) and between group and time-window (χ^2 = 128.3, df = 1, $p < 0.001$). To further investigate the nature of these interactions, we split the data according to the variable group.

The graphs in Figure 23 below show the proportion of looks to the target for each adjective-condition for children and adults.

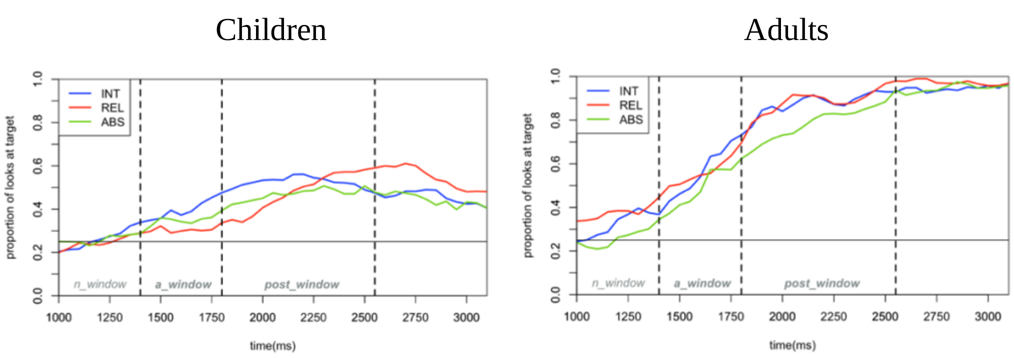


Figure 23. Proportion of looks at the target picture throughout the trial for children and adults for each adjective-condition. The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average offset of the adjective. The horizontal line indicates the chance level.

A first inspection of the two graphs further confirms that adults show a significant increase in looks to the target already in the adjective-window with respect to the noun-window, and they get over 0.8 proportion looks at target already around 1900ms (i.e., 100ms after the average adjective offset) for intersective and relative conditions, and interestingly, only around 2200 (i.e., 300ms after the average adjective offset) for the absolute adjective-condition. However, although adults are overall slower in the absolute condition with respect to the other two, accuracy is not affected as they reach 0.9 proportion looks at target at the end of the post-window (i.e., 700ms after the average adjective offset). By contrast, children vary in the time they need to identify the target across all three adjective-conditions. More specifically, children get around 0.5 proportion looks at target around 2200ms (i.e., 400ms after the average adjective offset) in the intersective adjective-condition and around 2300ms (i.e., 500ms after the average adjective offset) in the absolute adjective-condition. However, for the relative adjective-condition, children get around 0.5 proportion looks at target only around 2700ms (i.e., 900ms after the average adjective offset). Although target identification in the relative adjective-condition is evidently slower, it is interesting to notice that accuracy is not affected. Indeed, after the post-window children reach 0.6 proportion looks at target at around 2750ms in the relative adjective-condition, while they never overcome 0.5 in the other two conditions. Overall, from the graphs in Figure 23 we observed effects of adjective-condition and time-window. Thus, we sought statistical confirmation of these observations by analyzing these variables separately for the two age groups.

4.4.2.1 Children

After filtering out adults' data, we calculated a linear-mixed effects regression model with proportion of fixations at target as dependent variable and adjective-condition and time-window as independent variables with full interaction. Participant and item were added as crossed random effects (random intercepts). We found a significant main effect of time-window (ß = .12, SE = .001, t = 7.09, p < .001) and a significant main effect of adjective-condition (**INT vs. REL**, ß = - .09, SE = .04, t = -1.97, p = .05, but no significant difference between INT vs. ABS, p

= .2, nor ABS vs. REL, *p* = .4). Moreover, we found a significant interaction between adjective-condition and time-window (χ^2 = 7.48, df = 2, *p* < .05). In summary, children's fixations at target increase significantly from the adjective-window to the post noun-phrase-window. Furthermore, although we found that the proportion looks at the target picture differs significantly between the intersective and the relative adjective-conditions, no difference is found between the absolute adjective-condition and the other two adjective-conditions, signaling that processing this class of adjectives has commonalities with both. Finally, the significant interaction we found between time-windows and adjective-conditions demonstrates that, although intersective adjectives are processed faster, relative adjectives are fixated more during the post noun-phrase-window.

To investigate eye-movement differences between adjective-conditions with respect to the unfolding of the speech stimulus, we calculated three separate linear-mixed effects regression models on the adjective- and post noun-phrase-windows. We specified proportion of fixations at target as dependent variable, time-bin as independent variable and participants and items as crossed random effects (random intercepts). This model allowed us to compare the proportion looks at target in each 50ms bin with a baseline, i.e., looks at target at the adjective-onset. For intersective adjectives, we found that fixations started to significantly differ from the baseline already at around 1750ms (ß = .104, SE = .005, t = 1.90, *p* = .05) and the difference increased during the post noun-phrase-window (at 2200ms, ß = .21, SE = .005, t = 3.86, *p* < .001). For relative adjectives, a significant increase from the baseline was found only after the adjective-offset (at 2000ms, ß = .108, SE = .005, t = 1.97, *p* < .05).

4.4.2.2 Adults

The broad analysis conducted on children's results (presented in §4.4.2.1) was replicated with the adults' data in the adjective- and post noun-phrase-windows. We calculated a linear-mixed effects regression model with proportion looks at target as dependent variable and adjective-condition and time-window as fixed effects. Participant and item were added as crossed random effects (random intercepts). We found a significant main effect of time-window (ß = .3, SE = .009,

$t = 31.23$, $p < .001$) and a significant main effect of adjective-condition (**INT vs. ABS**, ß = .08, SE = .03, $t = 2.07$, $p < .05$ and **REL vs. ABS**, ß = .08, SE = .03, $t = 2.15$, $p < .05$, but no significant difference between INT vs. REL, $p = .9$). In summary, the statistical analysis confirmed that target preference increased significantly from the adjective-window to the post noun-phrase-window. Interestingly, in this broad time-window, looking behavior in the absolute adjective-condition significantly differed from the other two adjective-conditions. This result suggests that, although adults perform at ceiling in all adjective-conditions as shown by the graph in Figure 23, in the absolute adjective-condition they are slower in target identification with respect to the intersective and the relative adjective-conditions. A possible explanation for this result will be discussed in §4.5.

4.4.2.3 Summary

To summarize, the statistical analysis investigating the possible effect of the semantic classes of adjectives on looking pattern confirmed that children and adults are significantly different in all adjective-conditions with respect to both speed and accuracy of target reference resolution. More precisely, children integrate the meaning of nouns and intersective adjectives (e.g., *black*) faster than the other two adjective-conditions. Although relative adjectives (e.g., *big*) are significantly more challenging than intersective adjectives, absolute adjectives (e.g., *closed*) are in between the other two in that eye-movements do not differ significantly neither from intersective, nor from relative adjectives. Adults, by contrast, are fast and accurate in all adjective-conditions. Surprisingly, in interpreting absolute adjectives the statistical analysis revealed that they are slower than in the other two adjective-conditions.

4.4.3 Interpreting noun-adjective combinations in development

The analyses presented in §4.4.1 and §4.4.2 showed how children and adults vary in processing noun-adjective combinations with respect to speed and accuracy on the one hand, and in relation to the interpretation of different classes of adjectives on the other.

Previous literature on the topic demonstrated how, during the third year of age, children develop a fast and efficient incremental processing of adjectives in combinations with nouns. Since the age of the children in this experiment span from 28 months (2;4 years) to 63 months (5;3 years), the present section aims to determine possible effects of age in the interpretation of nouns and adjectives with respect to the first two research questions of the current study. To do so, we created three age groups. The first group (2-year-olds) included children from 2;4 years to 3;1 years (N = 12; mean age = 2;9, S.D. = 0;2, F = 7). In the second group (3-year-olds) toddlers span between 3;2 years and 3;9 years (N = 14; mean age = 3;5, S.D. = 0;2, F = 8). The last group (4-year-olds) included children between 3;11 years and 5;3 years (N = 12; mean age = 4;5, S.D. = 0;6; F = 5). Age groups were created to directly test the *Noun-Anchoring* mechanism, which establishes the age of 3 as a turning point in the automatization of the set-operations required to process noun-adjective combinations. In this respect, the *Noun-Anchor Hypothesis* predicts children in the 2-year-old group to underperform compared to the groups of older children. A comparison between age groups allows us to directly test this hypothesis. Figure 24 depicts the proportion looks at target for each of the three age groups.

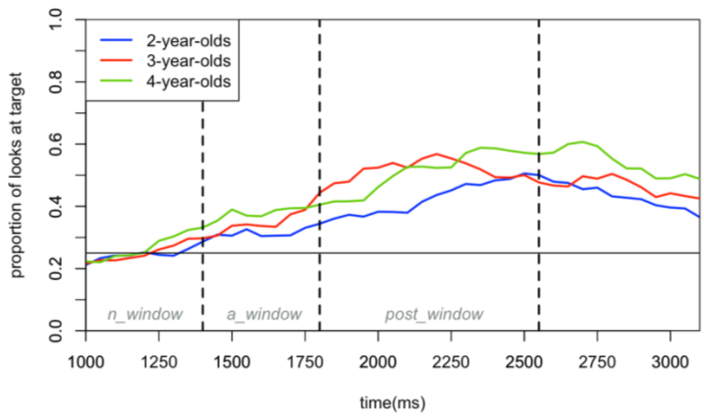


Figure 24. Proportion of looks at the target picture throughout the trial for the three age groups of children from the noun-onset. The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average offset of the adjective. The horizontal line indicates the chance level.

The graph shows that, during the noun-window, all children looked at the target at chance and that only the oldest children overcame 0.3 proportion of looks to the target object at the noun-offset. As the sentence unfolded, all groups increased their looks at the labeled picture, but it is only during the post noun-phrase-window that a visible age group difference emerges. Indeed, while 3- and 4-year-old children almost reached 0.6 proportion of looks at target at around 2200-2300ms, 2-year-olds only reached 0.5 at the adjective-offset (i.e., at 2550ms). To verify whether the observed group differences were significant, we performed a series of linear mixed-effects regression models following the analyses presented in §4.4.1 and §4.4.2.

First, we started from the **noun-window**, and we calculated a linear-mixed effects regression model specifying fixations to the labeled objects as dependent variable, age-group as independent variable and participants and items as crossed random effects (random intercepts). No significant group differences were found (χ^2 = 0.0003, df = 2, p = 0.9). Second, we analyzed fixations during the **adjective-window** by calculating a linear-mixed effects regression model specifying fixations at target as dependent variable, age-group as independent variable and participants and items as crossed random effects (random intercepts). We found a significant effect of group (**2-year-olds vs. 3-year-olds**, ß = .09, SE = .03, t = 2.624, $p < .05$; **2-year-olds vs. 4-year-olds**, ß = .09, SE = .03, t = 2.466, $p < .05$; no difference was found between 3-year-olds vs. 4-year-olds, p = .9).

Further analyses were performed to examine when, during the adjective- and post-noun-phrase windows, looks at the target started to significantly increase from the baseline (i.e., looks at the target at the noun-offset) for each age group. We calculated three linear-mixed effects regression models, specifying target preference as dependent variable and time as independent variable. Participant and item were added as crossed random effects (random intercepts). For 2-year-olds, we found a significant increase from baseline only after the adjective-offset (at 2200ms, ß = .12, SE = .05, t = 2.226, $p < .05$), further increasing during the post noun-phrase window (at 2500ms, ß = .19, SE = .05, t = 3.345, $p < .001$). No significant difference from the baseline was found during the adjective-window (at 1450ms, p = .9, at 1800ms, p = .5). For 3-year-olds, a significant increase from the baseline was found already at 1800ms (ß = .13, SE = .05, t = 2.725, $p < .01$), 200ms before the 4-year-olds (at 2000ms, ß = .10, SE = .05, t = 1.949, p = .05).

Lastly, to investigate the effect of age in the interpretation of different semantic classes of adjectives, we calculated a linear-mixed effects regression model on the adjective- and post-noun-phrase windows, specifying target preference as dependent variable and age-group and adjective-condition as fixed effects with full interaction. Participant and item were added as crossed random effects. No significant interaction was found (p = .01).

In summary, the statistical analyses reported in the present section confirmed that, during the noun-window, all children still look at all pictures on the screen and noun interpretation begins after the noun-offset. Interestingly, during the adjective-window, age differences start to emerge and while 3- and 4-year-olds' looks at target increase and reach the highest accuracy shortly after the adjective-offset, 2-year-olds' fixations to the target remain right above chance and increase significantly only half a second after the end of the sentence. As for the different adjective-conditions, the statistical analysis revealed no significant interaction between adjective-conditions and age groups, showing that the pattern we identified for children in general holds for 2-year-olds as well as for 4-year-olds. In other words, children are faster with intersective adjectives and slower with relative adjectives regardless of their age. The implication of these and previous findings will be discussed in the next section.

4.5 Discussion

The present study was developed to investigate how children interpret noun-adjective combinations in real time by analyzing their eye-movement behavior during the simultaneous presentation of visual and linguistic stimuli. In designing this Visual World task, we aimed at answering three research questions.

4.5.1 Interpreting nouns and adjectives in combination

The first research question (RQ1) asked whether, in interpreting Italian attributive adjectives in post-nominal position, young children show a *Noun-Anchoring* mechanism, i.e., if they interpret noun-adjective combinations starting from the noun-word and delay or omit the integration of the adjective meaning.

Our results revealed that children between two and five years of age are able to successfully integrate noun and adjective meanings to resolve reference when faced with 4-referent displays. In processing Italian questions of the type *Where is the Noun-Adjective?* (e.g., *Where is the black shoe?*), while looking at two objects crossed by two properties on the screen (e.g., a black shoe, a white shoe, a black sock, and a white sock), children start the interpretation of the noun by focusing their looks to the two labeled objects (e.g., the two shoes) and, after hearing the adjective, they accurately shift their gaze to the target object (e.g., the black shoe). Contrary to what the *Noun-Anchor Hypothesis* predicts, our findings show that children do not delay or omit adjective interpretation and no asymmetry between the noun and the adjective processing was found. Further, our results suggest that the presence of multiple competitors does not impede reference resolution and that, already at around 30 months, children are not distracted by either the noun or the adjective, reflecting an accurate and relatively fast integration, contrary to what has been found in previous studies with children equal in age (Ninio, 2004; Thorpe et al., 2006). Further, the integration was successful without the need to manipulate language speed (e.g., Tribushinina & Mak, 2016, Davies et al., 2021).

Finally, the comparison of children's looking pattern with adults' eye-movements revealed that children are overall slower and less accurate, as hypothesized. Our data showed that children focus on the labeled objects at the noun-offset and resolve reference at the adjective-offset, while adults are able to interpret each word of the combination during the time-window in which the word is heard, without the need to wait for its offset. Thus, although children possess the skills required, processing limitations emerge and result in a significant difference with adults' speed and accuracy. Nevertheless, contrary to what the *Noun-Anchor Hypothesis* predicts, the pattern is adult-like. Children are slower and less accurate, but they do not need more time to interpret the adjective with respect to the noun, nor the other way around.

4.5.2 The role of the semantic class in the online interpretation of adjectives

The second research question (RQ2) had the objective of investigating whether there are differences in children's interpretation of different semantic classes of adjectives.

Our results show that, by the end of the referring expression, children can integrate nouns and adjectives to resolve reference accurately in all adjective-conditions. However, as hypothesized, significant differences between adjective-conditions emerged. While the interpretation of intersective adjectives (e.g., *black*) happened shortly after the adjective-offset, integrating nouns with relative adjectives (e.g., *big*) was significantly more challenging and required more time. This is in line with the prediction that, in interpreting relative adjectives, once the noun has been presented and reference has been resolved, children look away from the target towards the noun competitor, sensibly checking their choice against the contrast object. Indeed, most semantic theories propose that establishing a standard of comparison is necessary to determine what counts as having a certain property in a given context (e.g., Kamp & Partee, 1995; Kennedy, 1999). Thus, upon hearing *the teddy big*, children first identified the two objects "teddy" on the screen, and later confronted them to identify the one that was big. As for absolute adjectives, data revealed no significant difference between this adjective-condition and both the intersective and the relative adjective-conditions. This result matches with the theoretical observation that absolute adjectives share properties with both intersective and relative adjectives (e.g., Kennedy, 1999; Kennedy & McNally, 2005; Kennedy, 2007). Thus, these findings provide evidence that children are aware of the inherent semantic differences among subsective adjectives and, consequently, more challenging adjectives (i.e., relative adjectives) require more processing time.

Interestingly, however, the analysis of adults' eye-movements showed how adults' processing of absolute adjectives is slower than the other two adjective-conditions. We propose two possible interpretations for this. First, this finding may have a theoretical explanation recalling a similar result found by Aparicio et al. (2015). In interpreting prenominal absolute adjectives, adult participants showed a delayed effect of contrast with respect to color and relative adjectives. They argued that this result might be a consequence of the precise interpretation required in processing two-closed-boundaries absolute adjectives (e.g., *open/closed*) in comparison to one-closed-boundary absolute adjectives (e.g., *clean/dirty*) and open-boundaries relative adjectives (e.g., *big/small*). Thus, while in interpreting *closed book* no comparison among books in required, to identify the *dirty* object, participants may need to check the object-competitor to make sure that the target object was the dirtier on the screen. Second, there may also be a methodological explanation for this apparent delay in adults' processing of absolute adjectives. Since the task was specifically designed for children and no difficulties were expected in adults' performance, this finding might result from a task bias concerning visual stimuli in this adjective-condition. Indeed, the visual representation of the properties labeled by absolute adjectives require the manipulation of the drawing on multiple dimensions. To put it more clearly, representing intersective and relative adjectives only required the manipulation of one single aspect of the drawing, e.g., the black/white shoes were identical except for their color and the big/small fishes were identical while differing in size. Absolute adjectives, by contrast, required the objects to be represented by two different pictures. For example, the 'open book' and the 'closed book' were drawings differing both in the predominant color (red of the cover when closed, white of the pages when open) and size (see Figure 14). Thus, a more attentive analysis of the multidimensional differences among the objects within the same category might have delayed participants' shifts to the target object. We maintain that this second explanation better accounts for our finding in this experiment. Multidimensional differences within the object-category are indeed relevant in this task; nevertheless, we believe that our results cannot be accounted for by assuming that adults have processing difficulties with absolute adjectives in general and that these limitations do not affect the interpretation of intersective nor relative adjectives. Hence, we argue that this finding is (mainly) the result of a task bias regarding specifically the absolute adjective-condition.

4.5.3 Two- to four-year-olds' processing of noun-adjective combinations

The third research question (RQ3) investigated whether children's looking behavior reflected differences in development in relation to the online processing of noun-adjective combinations and to the semantic differences among subsective adjectives.

Although contradictory results were found as to whether children interpret noun-adjective and adjective-noun combinations incrementally, all relevant studies in the previous literature showed how, at around 36 months of age, toddlers show a twist in development. At this age, toddlers become more accurate (Ninio, 2004; Thorpe et al., 2006) and, in processing adjective-noun combinations, they manage to interpret adjective-noun combinations incrementally, i.e., starting from the prenominal adjective (Thorpe et al., 2010; Tribushinina & Mak, 2016), unlike 30-month-old toddlers (Fernald et al., 2010). Our results revealed that children younger than three years of age have the most difficulties in interpreting such combinations in comparison to 3- and 4-year-olds. Their less developed cognitive resources make this task more challenging, resulting in a significantly slower process of interpretation. Nevertheless, 2-year-olds' integration of nouns and adjectives shows the same pattern as the older children's, i.e., they interpret the noun after its offset, and they integrate its meaning with that of the adjective starting from the adjective offset. This is surprising considering that utterances were presented at natural speed. Contrary to previous results with children as young as 30 months of age (Ninio, 2004; Thorpe et al., 2006; Fernald et al., 2010), our findings do not provide evidence in favor of the *Noun-Anchor Hypothesis*. Indeed, the youngest as well as the oldest children did not show an asymmetry in noun and adjective interpretation and were able to accurately integrate their meanings. Our study demonstrates that, once the meanings of nouns and adjectives are acquired, the pattern of interpretation of noun-adjective combinations is adult-like but slower and gets faster with age.

In relation to the semantic differences among subsective adjectives, our data show that the pattern we observed for children as a whole group holds for each age group. More specifically, our data revealed that the difficulties children show in integrating nouns with relative adjectives are shared among 2-year-old toddlers, as well all 3- and 4-year-olds. Since to interpret relative adjectives, a robust comparison is needed, all children show a delayed target identification with respect to intersective adjectives and absolute adjectives. Similarly, all children perform an earlier and faster integration of nouns and intersective adjectives, confirming our predictions.

4.5.4 Conclusion

In conclusion, previous studies of children's ability to interpret noun-adjective combinations in resolving reference have relied on either end-point offline data or have analyzed online behavior in response to simple displays, investigating one single adjective class. Our experiment has taken a comprehensive, rigorous approach by analyzing high-resolution online eye-tracking data in response to stimuli that demand full integration of nouns and subsective adjectives in children between two and five years of age. Findings from the current study provide evidence of a continuity in children's development of sophisticated, adult-like processing skills. Crucially, we have evidence that children as young as 30 months do show a remarkable slowness in the interpretation of nouns and adjectives in combination with respect to experienced adults, but that the pattern of interpretation is the same. In contrast to previous findings, the current experiment revealed that children younger than three years of age do not delay or leave out adjective interpretation and that they process noun and adjective meanings equally quickly. A closer look at the interpretation of different semantic classes of adjectives allowed us to observe different looking patterns across adjective-conditions, revealing that young toddlers are aware of the different ways in which each adjective class is interpreted within different contexts. Crucially, examining the online processing with the eye-tracking technology has shed light on details about noun-adjective integration and patterns of interpretation elicited by adjective semantics that never emerged in previous literature.

5 Experiment 2: Processing informative nouns and adjectives

5.1 Introduction

Experiment 1 (ch.4) failed at finding evidence of the *Noun-Anchor Hypothesis,* i.e., that, in processing noun-adjective combinations, children younger than 36 months identify the labeled object but delay or fail the integration of the adjective. In investigating the interpretation of Italian noun-adjective combinations by children as young as 28-month-olds, Experiment 1 showed that, despite being slower and less accurate than adults, children were able to identify the target reference when presented with a 4-picture display. To be able to do so, participants needed to interpret both words of the combinations (i.e., the noun and the adjective) and to integrate their meanings. This design ascertained that children (2;4 – 5;3 years) can integrate noun and adjective information by the end of a referring expression but did not allow for a comparison between conditions in which adjective interpretation is required and conditions in which it is not. Experiment 2 addresses this issue. We designed a Visual World study reducing the complexity of Experiment 1, aiming to compare experimental conditions in which the noun word is either informative or non-informative about the target object on the screen. In other words, we present children with conditions in which the interpretation of the postnominal adjective is necessary to identify the target picture, in comparison to conditions in which the noun is informative enough for reference resolution. Furthermore, adjectives in Experiment 2 are limited to colors (intersective), which, in Experiment 1, turned out to be easier to interpret with respect to relative and absolute adjectives.

Experiment 2 builds upon the findings from Experiment 1 and previous studies by Weisleder and Fernald (2009) and Fernald et al. (2010) on the ***Noun-Anchor Hypothesis***. These studies exploited the Looking-While-Listening paradigm to ask whether Spanish- and English-speaking children make use of an informative word (i.e., the noun in Spanish and the adjective in English) to resolve reference in a two-picture scenario. In their study with English-children, Fernald et al. (2010) found that, upon hearing *blue car* in a blue-car/red-car visual scene, 36-month-olds were able to exploit the adjective meaning to identify the target picture on the screen, without the need to listen through the noun. By contrast, 30-month-olds failed to take advantage of the prenominal adjective, showing a huge difficulty in processing such combinations. This finding led the authors to conclude that their findings confirmed the *Noun-Anchor Hypothesis.*

In commenting Fernald et al.'s results, Weisleder and Fernald (2009) proposed that having prenominal adjectives might be suboptimal for the processing of such combinations. Indeed, what the *Noun-Anchor Hypothesis* predicts is that postnominal adjectives eliminate the extra challenge of keeping the adjective in memory and having to recover it after the interpretation of the noun. However, Weisleder and Fernald (2009) argued that the same problem could occur with postnominal adjectives. The Spanish condition with a blue car and a blue plane (e.g., *¿Dónde está el carro azul?,* lit. "Where is the car blue?") could be seen as analogous to the situation experienced by English-learners in the red-car/blue-car condition. In this condition, an informative word (e.g., the adjective *blue*) is followed by an uninformative word (e.g., the noun *car*). By contrast, in the Spanish condition an informative word (e.g., the noun *carro*) is followed by an uninformative word (e.g., the adjective *azul*) (see §2.5.2.4 for a detailed review). Since in Spanish the noun word offers enough information to identify the target reference, children's attention might be disrupted by the adjective, which labels a property relevant to both images (i.e., the car and the plane). Weisleder and Fernald (2009) refer to this interpretation as the ***Sequential Integration Hypothesis***, suggesting that children have difficulty with the integration of any two content words. Consequently, the integration is easier when the informative word comes last, as this arrangement is less demanding on working memory. Thus, English-learning children are facilitated when the noun is informative (i.e., in the blue-car/blue-house condition), while Spanish-speaking children more easily process the condition in which only the adjective is informative (i.e., the blue-car/red-house condition). Contrary to the *Sequential Integration Hypothesis*, Weisleder and Fernald (2009) found that, when the noun was informative, 42-month-old Spanish-learners where able to identify the target reference already at the noun, without suffering any disruption from the following adjective, informative or uninformative. On trials in which the noun was uninformative, children listened though the noun and shifted their gaze to the target object only after hearing the adjective. The authors concluded that, in line with the *Noun-Anchor Hypothesis*, children's interpretation of noun-adjective combinations is facilitated when the noun is heard before the adjective.

These experiments present some evident limitations. First, as already mentioned for these and other studies in the literature investigating the developmental interpretation of nouns and adjectives in combination, children younger than 36 months of age need to be included in the experiment. Involving younger toddlers is necessary not only to find evidence against or in favor of the *Noun-Anchor Hypothesis*, which proposes that the challenging process of interpretation concerns especially children below three years of age (Ninio, 2004), but also to observe any other pattern of comprehension and acquisition of such structures. Various offline and online studies have shown how, around the third year of age, children start to develop and automatize the processing of such combinations (e.g., Klibanoff & Waxman, 2000; Ninio, 2004; Fernald et al., 2010). Thus, this protracted process needs to be observed before the emergence of the target-like pattern of interpretation.

5.2 The current study

Previous studies in the literature investigating adjective acquisition have exploited a series of different experimental methods, including both online and offline tasks. However, we believe it is important that older research questions are revisited using newer experimental methodologies. As pointed out before, eye-tracking can tell us about children's language processing, which plays an important role for acquisition.

The results of Experiment 1 were particularly interesting in that children's online interpretation of attributive adjectives in a four-referent-display revealed the great difficulty the youngest 2-year-olds have compared to older children and adults, whereas offline data from previous studies (e.g., Ninio, 2004; Thorpe et al., 2006) did not.

The current experiment focuses on 2-, 3- and 4-year-olds' online interpretation of nouns combined with color-adjectives in Italian, aiming to investigate children's processing pattern. To do so, we reduced the complexity of Experiment 1 by decreasing the number of pictures on the screen and by limiting adjectives to colors, which were the easiest and fastest to interpret in the first experiment. We designed a Visual World task with eye-movement recordings during the interpretation of Italian questions of the type *Dov'è il dinosauro verde?* ("Where is the green dinosaur?"). While listening to the question, participants look at two objects on the screen. Three visual conditions are included in a fully crossed design and vary according to the informativeness of the noun or the adjective with respect to the target referent.

We aimed to answer the following research questions:

1) Do children show a *sequential-integration* pattern in interpreting color-adjectives combined with informative and non-informative nouns? Are they disrupted by uninformative postnominal adjectives?
2) Do children and adults differ in the pattern of interpretation with respect to speed and accuracy?
3) Are there differences between 2-, 3- and 4-year-old children in processing noun-adjective combinations across different informativeness conditions?

5.3 Methods

5.3.1 Materials

The experimental design was the same used in Weisleder and Fernald (2009) and Fernald et al. (2010), adapted to Italian and to the eye-tracking technology.

We realized a Visual World task used 48 colored stimulus images on a white background created from child-friendly drawings of familiar objects. Pictures were presented in pairs that were matched for saliency and grammatical gender. The visual display showed two pictures, i.e., a target (e.g., a dinosaur) and a competitor (e.g., a hat) (see Table 7). Side of presentation of each object on the screen was pseudo-randomized and counterbalanced across trials.

Participants viewed displays while listening to utterances recorded by an Italian native speaker with neutral accent and neutral intonation. All stimuli were processed in Praat (version 6.0.49; Boersma & Weenink, 2016). The average utterance duration was 1689ms. Stimuli consisted of questions of the form "Where is the Noun-Color?" in Italian, e.g., *Dov'è il dinosauro verde?* ("Where is the green dinosaur?"), containing the carrier phrase ("Where is the"), a noun and a post-nominal color-adjective. The visual array varied to result in three experimental conditions, as shown in Table 7.

Table 7. Examples of the stimuli for each condition.

Cond.	**Visual Condition**	**Visual stimuli**	**Speech stimuli**
All Different (AD)	Different Objects/ Different Colors	green dino – red hat  	*Dov'è il dinosauro verde?* ("Where's the green dinosaur?")
Same Color (SC)	Different Objects/ Same Colors	green dino – green hat  	*Dov'è il dinosauro verde?* ("Where's the green dinosaur?")
Same Object (SO)	Same Objects/ Different Colors	green dino – red dino  	*Dov'è il dinosauro verde?* ("Where's the green dinosaur?")

Upon hearing the question, e.g., *Dov'è il dinosauro verde?*, lit. "Where is the dinosaur green?", the screen showed the target object (i.e., the green dinosaur) paired with a competitor. In the *All-Different* condition, participant saw an object-color competitor (e.g., a red hat); in the *Same-Color* condition an object competitor (e.g., a green hat); in the *Same-Object* condition a color-competitor (e.g., a red dinosaur). While in the *All-Different* and in the *Same-Color* conditions target identification was possible already at the noun, which was alone informative, in the *Same-Object* condition, the color-word was necessary to identify the correct reference of the noun-phrase.

Visual and auditory stimuli were combined to form 288 trials, divided into 12 lists of 24 trials each. All stimuli are listed in Table 8.

Table 8. List of items used in the experiment.

	Nouns		Adjectives			
1	*accappatoio, ciuccio*	'bathrobe', 'dummy'	*azzurro*	'lightblue'	*rosa*	'pink'
2	*bandiera, bottiglia*	'flag', 'bottle'	*blu*	'blue'	*verde*	'green'
3	*borsa, campana*	'purse', 'bell'	*grigia*	'grey'	*rossa*	'red'
4	*bottone, dado*	'button', 'die'	*azzurro*	'lightblue'	*rosso*	'red'
5	*camicia, collana*	'blouse', 'necklace'	*bianca*	'white'	*rosa*	'pink'
6	*camion, uccello*	'truck', 'bird'	*azzurro*	'lightblue'	*bianco*	'white'
7	*cane, martello*	'dog', 'hammer'	*grigio*	'grey'	*marrone*	'brown'
8	*cappello, dinosauro*	'hat', 'dinosaur'	*rosso*	'red'	*verde*	'green'
9	*castello, elicottero*	'helicopter', 'fish'	*azzurro*	'lightblue'	*rosso*	'red'
10	*chitarra, pistola*	'guitar', 'gun'	*azzurra*	'lightblue'	*viola*	'purple'
11	*conchiglia, coperta*	'shell', 'blanket'	*blu*	'blue'	*rosa*	'pink'
12	*coniglio, orso*	'bunny', 'bear'	*bianco*	'white'	*marrone*	'brown'
13	*corona, tromba*	'crown', 'trumpet'	*gialla*	'yellow'	*grigia*	'grey'
14	*drago, trattore*	'dragon', 'tractor'	*verde*	'green'	*viola*	'purple'
15	*forchetta, tazza*	'fork', 'cup'	*gialla*	'yellow'	*viola*	'purple'
16	*fungo, treno*	'mushroom', 'train'	*marrone*	'brown'	*rosso*	'red'
17	*guanto, pappagallo*	'glove', 'parrot'	*blu*	'blue'	*rosso*	'red'
18	*mucca, porta*	'cow', 'door'	*bianca*	'white'	*marrone*	'brown'
19	*palloncino, tappeto*	'balloon', 'carpet'	*verde*	'green'	*viola*	'purple'
20	*panchina, tartaruga*	'bench', 'turtle'	*marrone*	'brown'	*verde*	'green'
21	*papera, pentola*	'duck', 'pot'	*bianca*	'white'	*gialla*	'yellow'
22	*pennello, pettine*	'brush', 'comb'	*grigio*	'grey'	*rosa*	'pink'
23	*pipa, scala*	'pipe', 'ladder'	*gialla*	'yellow'	*marrone*	'brown'
24	*secchio, zaino*	'bucket', 'bag'	*blu*	'blue'	*giallo*	'yellow'

5.3.2 Participants

Twenty-eight Italian-learning children (2;4-5;2, mean age = 3;7, S.D. = 0;8) took part in the experiment; 14 were female. Twenty-two Italian monolingual adults (19;1-29;9, mean age = 25;7, S.D. = 2;8) participated as controls. 17 were female.

The study was approved by the local Ethics Committee of the University of Verona, and was conducted in accordance with the standards specified in the 2013 Declaration of Helsinki.

5.3.3 Procedure

The test was conducted in a dimly lit and soundproof testing room at the Laboratory of Text, Language and Cognition (LaTeC) of the University of Verona. Participants' eye movements were recorded using an SR Research EyeLink 1000 Plus eye-tracker, set in head-free remote mode and sampled monocularly at 500 Hz with a 16mm lens. The experiment was run on a computer connected to a 24" colour BenQ monitor for visual stimulus presentation. Speech stimuli were played over two loudspeakers, positioned on both sides of the screen. The experimental procedures were implemented in Experiment Builder and eye-movement data were extracted through Data Viewer.

Participants were welcomed into the laboratory where the study was explained. Adult participants were told that they would take part in an experiment designed for young children. They were seated on a chair in front of the screen and were instructed to follow the visual and auditory stimuli. They were not given specific instructions but were invited to stay seated and as still as possible throughout the duration of the experimental session. The children were seated in front of the experimental apparatus on their parent's lap and they were told that they would play a finding game with the cartoon character Peppa Pig, and that their job was to look at the screen and follow the cartoon's voice. To tackle head movements, a small sticker was placed on the participant's forehead.

Calibration and validation procedures were carried out using a five-point display before the beginning of the experiment and a drift correction was repeated once every four trials. After the calibration and validation procedures were successfully performed, a first familiarization phase was carried out by showing 4 "warm-up" pictures on the screen presented one after the other, while a sentence labeled the object (e.g., "Look, a tiger!"). This phase was useful to familiarize the child with the two positions of the pictures on the screen. After the familiarization phase the experimental session began. In each trial, two pictures appeared on the screen simultaneously. The auditory stimulus started after 2s from the presentation of the pictures. From the sentence-onset, pictures remained displayed on the screen for 3000ms, followed by an 800ms blank screen that ended the trial. An example of the trial is shown in Figure 25. A drift correction was performed every four trials, paired with one of the three filler sentences (e.g., "Look!"), recorded in a child-friendly intonation. The testing session lasted approximately 10 minutes.

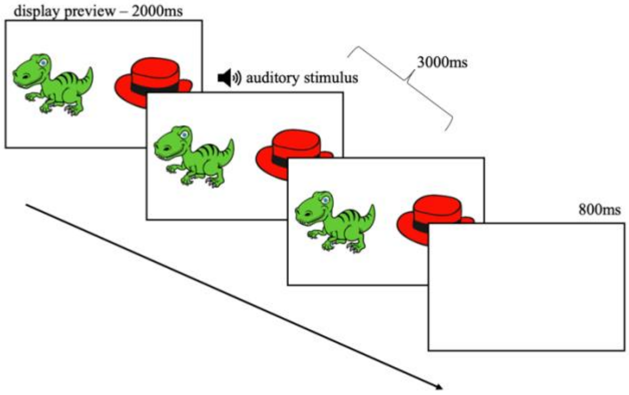


Figure 25. Experimental procedure.

5.3.4 Predictions

Exploiting the eye-tracking technology, the Visual World paradigm allows us to access the interpretation of nouns and adjectives in real time and to precisely address our research questions in the presence of a revised experimental design with new linguistic and visual stimuli.

5.3.4.1 Predictions for Research Question 1

The first research question (RQ1) asks whether children are able to interpret nouns and adjectives in different visual conditions in which words are either informative or non-informative. Predictions are made for each informativeness-condition outlined in §5.3.1.

Upon hearing, e.g., e.g., *Dov'è il dinosauro verde?* (lit. "Where is the dinosaur green?"), in the *AllDifferent* condition participants see the target object (e.g., a green dinosaur) paired with a different object in a different color (e.g., a red hat). Here, the noun is informative about the referent, while the adjective adds unnecessary, though informative information about the color. What we expect is that children easily and rapidly shift their gaze towards the target object upon hearing the noun word, without the need to go back to the distractor picture after hearing the adjective, despite it being informative.

In *SameColor* condition, participants are presented with the referent (e.g., a green dinosaur) and a second object matching in color with the target picture (e.g., a green hat). In this case, the noun is informative and suffices for the correct identification of the target. However, since the two objects share the same color, the adjective is uninformative. According to the *Sequential Integration Hypothesis* formulated by Weisleder and Fernald (2009), this condition might cause disruption at the adjective level because it can be referred to both pictures on the screen. Thus, children might look back and forth between the two until remembering the informative noun that was kept in memory, resulting in a slower target identification. Since Weisleder and Fernald (2009) did not find this pattern of interpretation in the Spanish experiment, we expect to find the same results in Italian. Although the first and the second condition differ as for the informativeness of the adjective, we predict no significant difference between the processing of the noun-adjective combinations in the first and in the second visual condition. We expect that children will shift their looks to the target picture at the informative noun, as in the first condition, and will not be disrupted by the following adjective.

The *SameObject* condition, by contrast, requires the interpretation and the integration of the adjective meaning for reference resolution. In this condition, participants see two objects matching in category (e.g., two dinosaurs) but differing in color (e.g., one green and one red). Since the noun is uninformative with respect to the target, we expect children to fixate both pictures at chance until they hear the informative adjective. In addition, since the visual scene only presents two pictures and adjectives are reduced to colors, we predict that children will be as accurate as in the other two conditions. This visual condition allows us to directly test the *Noun-Anchor Hypothesis* in that it is the only condition that requires the integration of noun and adjective meanings to resolve reference. Whereas in the previous two conditions the information from the noun was necessary though sufficient, in this condition participants need the information provided by the adjective. What the

Noun-Anchor Hypothesis predicts is that younger children delay or ignore the interpretation of the adjective combined with a noun which would result in a slower and less accurate pattern of interpretation with respect to the previous two conditions. This effect, however, can only be identified by splitting children in age groups and analyzing their looking behavior. This issue will be addressed by Research Question 3.

5.3.4.2 Predictions for Research Question 2

The second research question (RQ2) investigates whether children's interpretation of nouns and adjectives differs from the adults' when presented with a two-picture scenario in different informativeness conditions. Since this experiment was designed to reduce the complexity of Experiment 1 as for adjective semantics and number of objects in the visual display, we do not expect to find the huge group difference we observed in the previous experiment. In other words, we predict no significant difference between the children's and the adults' interpretations to emerge in terms of both speed and accuracy. However, unlike the first and the second conditions, the *SameObject* condition requires the adjective to be processed and interpreted to resolve reference. Thus, according to the *Noun-Anchor Hypothesis*, in this last condition children would show a significantly slower and less accurate pattern of interpretation in comparison to adults, being attributable to the youngest children.

5.3.4.3 Predictions for Research Question 3

The third research question (RQ3) addresses the effect of age on children's processing of noun-adjective combinations in different informativeness conditions. Although in Experiment 1 we found a significant difference between 2-year-olds on the one hand, and 3- and 4-year-olds on the other, we do not expect to find statistically significant differences in the present experiment. Having reduced the complexity of the previous experiment, we hypothesize that younger children will process the linguistic stimuli as fast and as accurate as the older children. However, in line with what the *Noun-Anchor Hypothesis* predicts, the youngest children should show a different looking pattern in the condition in which only the adjective is informative. This condition allows us to observe a possible asymmetry if compared to the conditions in which the noun is informative, in that processing the adjective word is necessary to identify the target. Thus, only in this last condition, 2-year-olds are expected to show a slower pattern of interpretation of the sentence and a lower accuracy when confronted to 3- and 4-year-olds.

5.4 Results

The eye-movement data recorded by the Eyelink system were extracted though the Data Viewer software and prepared for the statistical analysis with R (R Development Core Team, 2019). Data were analyzed using the packages lme4 and lmerTest (Bates et al. 2015, Kuznetsova et al. 2017, lmer function), LMERConvenienceFunctions (Tremblay & Ransijn 2015, summary function), car (Fox & Weissberg 2011, Anova function), itsadug and mgcv (Wood, 2006; Baayen et al., 2017; Wood, 2017; Baayen et al., 2018, bam and logit functions). Prior to analysis, we excluded trials in which neither picture was fixated, before or after the auditory stimulus was displayed. This led to the removal of 3% of the data. Furthermore, for each participant, we excluded all trials in which none of the pictures was fixated for more than the 50% of its duration, leading to the removal of another 9% of the data.

From the eye-tracking data, we determined fixations in 50ms steps. For each 50ms-bin, we aggregated proportions of fixations to the two areas of interests, i.e., we conducted the statistical analysis only on fixations on the two pictures (i.e., the target and the distractor). We accessed the proportion of looking time at named target picture over three time-intervals in the speech stimulus. First, the *noun-window* corresponded to the mean duration of the noun, starting from the mean noun-onset (2500ms-2900ms). Second, the *adjective-window* corresponded to the mean duration of the adjective, starting from the average adjective onset (2950ms-3300ms). Finally, the *post noun-phrase-window* captured fixations during a 700ms window after the average offset of the adjective (3350ms-4050ms).

The analysis of the results is organized as follows. In §5.4.1 the data will be analyzed to answer RQ1 and RQ2, aiming at investigating children's processing of nouns and adjectives in different visual conditions, and in comparison to the adult controls. §5.4.2 focuses on children only, with the objective of answering RQ3 on possible developmental differences in the computation of informative and uninformative nouns and adjectives.

5.4.1 Children and adults' processing of nouns and adjectives

The first statistical analysis aimed at investigating children's interpretation of nouns and adjectives in three different informativeness conditions in comparison to adults.

Before presenting the analysis, we comment on the graphs below, depicting the proportion of looks at the target picture for children (on the left) and adults (on the right). The graph in Figure 26 depicts children's proportion of looks to the target over the time course of the trial. The vertical dotted lines are drawn according to the three time-windows identified as the areas of interests (i.e., noun-window, adjective-window and post noun-phrase-window).

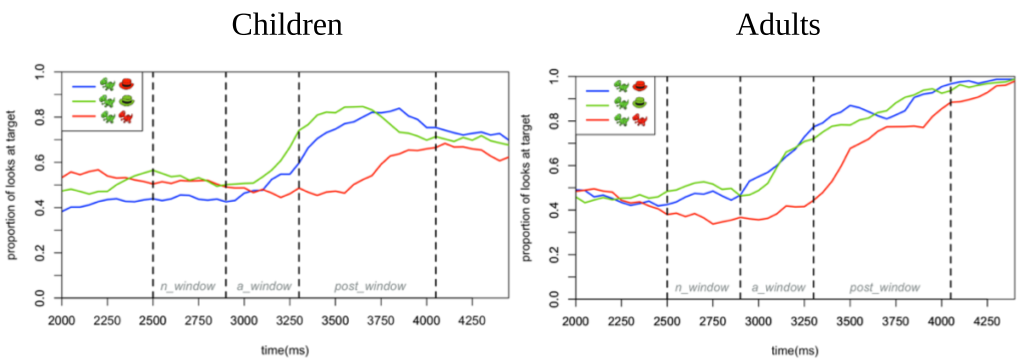


Figure 26. Proportion of looks at the target picture throughout the trial for children (on the left) and adults (on the right). Looks at target are marked in blue in the AD condition; in green in the SC condition; in red in the SO condition. The first vertical line indicates the average onset of the noun; the second vertical line indicates the average onset of the adjective; the third vertical line indicates the average offset of the adjective.

From a first visual inspection of the graphs in Figure 26, some preliminary observations can be made. First, in the *All-Different* condition (in blue) and in the *Same-Color* condition (in green), children look at both pictures at chance (around 0.5 proportion of fixations at target) until 200ms after the noun-offset, when their looks to the mentioned object start to increase. In both conditions, children reach 0.8 proportion of fixations, though faster in the *Same-Color* condition (around 3600ms) than in the *All-Different* condition (around 3800ms). Adults, by contrast, start to fixate the target object earlier in both conditions (i.e., at around 2900ms), becoming more and more accurate as the sentence unfolds and reaching ceiling at around 4000ms.

Second, while adults in the *Same-Object* condition (in red) increase their looks to the target already during the adjective window and rapidly resolve reference reaching 0.85 proportion of fixations at the target, children keep looking back and forth between the two pictures on the screen while listening to the uninformative noun and start to increasingly fixate the target only around 3550ms (i.e., 350ms after the adjective offset). Interestingly, the *Same-Object* condition seems to result in a lower accuracy, in that children never overcome 0.6 proportion of fixations to the target.

To answer RQ1 and RQ2, we performed a statistical analysis aiming at identifying possible significant differences among informativeness conditions and possible group effects. To do so, we verified the observations we made from the graphs in Figure 26 in a series of linear mixed effects regression models for each of the three time-windows, investigating the effects of condition and group on looks at the target picture.

5.4.1.1 Statistical analysis on the noun-window

For the first analysis, we analyzed the data within the **noun-window** only. During this time-window, participants are listening to an informative word in the *Same-Color* and *All-Different* conditions, while the noun is uninformative in the *Same-Object* condition.

We calculated a linear-mixed effects regression model with proportion of fixations to the target object as dependent variable and group and condition as fixed effects with full interaction. Participants and items were added as crossed random effects (random intercepts). The model showed a significant main effect of condition ($p < .001$), but no significant effect of group ($p = .3$). Further, we found a significant interaction ($\chi^2 = 56.7$, df = 2, $p < 0.001$).

To investigate the nature of this interaction, we split the data according to the variable group and we calculated two separate linear-mixed effects regression models with fixations at the target picture as dependent variable and condition as independent variable. Participants and items were added as crossed random effects (random intercepts). For children, the model revealed a significant main effect of condition (**AD vs. SC**, ß = - .08, SE = .02, *t* = -5.17, *p* < .001 and **AD vs. SO**, ß = 0.07, SE = .01, *t* = 4.49, *p* < .001, but no significant difference between SO vs. SC, *p* = .5). For adults, we found a significant main effect of condition (**SO vs. AD**, ß = - .10, SE = .02, *t* = -5.67, *p* < .001 and **SO vs. SC**, ß = - .13, SE = .02, *t* = -5.51, *p* < .001, but no significant difference between SC vs. AD, *p* = .07).

In summary, the statistical analysis on the noun-window revealed no significant difference between groups, but significant processing differences across conditions. On the one hand, adults are understandably significantly slower in the *Same-Object* condition with respect to the other two visual conditions, signaling that they are interpreting the informative noun in the *Same-Color* and in the *All-Different* condition, while remaining around chance level in the *Same-Object* condition. By contrast, children show a different pattern. They fixate the target object less in the *All-Different* condition in comparison to the *Same-Object* and the *Same-Color* conditions. In other words, during the noun-window, while in the *Same-Object* and the *Same-Color* conditions children look at both pictures around chance, in the *All-Different* condition they fixate the distractor object more often. Since we balanced our stimuli in both their position on the screen and in their status as target or distractor, we believe this finding might be related to the low number of participants, resulting in more frequent fixations to the distractor in comparison to the target.

5.4.1.2 Statistical analysis on the adjective-window

The second analysis was performed on the **adjective-window**. During this window participants are listening to an informative, though unnecessary adjective in the *All-Different* condition, and to an uninformative adjective in the *Same-Color* condition. By contrast, in the *Same-Object* condition, the information provided by the adjective is informative and necessary for target identification.

We calculated a linear-mixed effects regression model with proportion of fixations to the target as dependent variable and group and condition as independent variables with full interaction. Participants and items were added as crossed random effects (random intercepts). We found a significant main effect of condition ($p <$ .001) and of group ($p < .001$). Moreover, we found a significant interaction (χ^2= 65.1, df = 2, $p < 0.001$).

To better understand the nature of this interaction, we split the data according to the variable group and we calculated two separate linear-mixed effects regression models with fixations at the target object as dependent variable and condition as independent variable; participants and items were added as crossed random effects (random intercepts). For children, we found a significant main effect of condition (**SC vs. AD**, ß = - .07, SE = .01, t = -4.26, $p < .001$; **SC vs. SO**, ß = - 0.11, SE = .01, t = -6.67, $p < .001$; **SO vs. AD**, ß = - .04, SE = .02, t = -2.36, $p <$ .05). For adults, we also found a significant main effect of condition (**SO vs. AD**, ß = - .22, SE = .01, t = -12.73, $p < .001$ and **SO vs. SC**, ß = - .19, SE = .02, t = -11.14, $p < .001$, but no significant difference between SC vs. AD, $p = .09$).

To sum up, the statistical analysis on the adjective-window indicated that the looking behavior of the two groups differed across conditions. Children look at the target picture more on the *Same-Color* and the *All-Different* conditions than in the *Same-Object* condition, as expected. Interestingly, however, they seem to be facilitated in the *Same-Color* condition, i.e., when the over-informative adjective follows the informative noun, children are faster in target identification. Adults keep increasing their looks at the target picture during the adjective window in all three conditions, although their fixations at the referent picture in the *Same-Object* condition are still significantly less frequent than in the *Same-Color* and *All-Different* conditions, as the disambiguation point coincides with the postnominal adjective.

5.4.1.3 Statistical analysis on the post noun-phrase window

A further analysis was conducted on the **post noun-phrase window**, i.e., a 700ms time-window after the average offset of the adjective. This analysis aimed at investigating participants' looking behavior at the end of the sentence for each informativeness condition. Thus, we calculated a linear-mixed effects regression model with fixations to the target object as dependent variable and group and condition as fixed effects with full interaction. Participants and items were added as crossed random effects (random intercepts). The model showed a significant effect of condition ($p < .001$) and an effect of group approaching significance ($p = .05$). Furthermore, we found a significant interaction between condition and group ($\chi^2 = 45.2$, df = 2, $p < .001$).

To further investigate this interaction, we analyzed the data separated for each group and we calculated two separate linear-mixed effects regression models with fixations at the target object as dependent variable and condition as independent variable; participants and items were added as crossed random effects (random intercepts). Both models showed a significant effect of condition, for children (**SO vs. AD**, ß = .23, SE = .01, $t = 20.90$, $p < .001$ and **SO vs. SC**, ß = .23, SE = .01, $t = 20.78$, $p < .001$, but no significant difference between SC vs. AD, $p = .7$) and for adults (**SO vs. AD**, ß = .14, SE = .01, $t = 13.68$, $p < .001$ and **SO vs. SC**, ß = .14, SE = .01, $t = 12.91$, $p < .001$, but no significant difference between SC vs. AD, $p = .3$).

In summary, the statistical analysis on the post noun-phrase window revealed that, after the end of the sentence, children and adults do not strongly differ in their looking pattern. However, children are overall less accurate. Moreover, while no difference was found between the *All-Different* and the *Same-Color* conditions in both groups, a statistically significant difference was confirmed for both groups between the *Same-Object* condition and the other two conditions.

From the analysis on each time-window we found statistical confirmation of the differences between groups and conditions we observed in the graph in Figure 26. In the next section (§5.4.1.4), data will be analyzed within a broad time-window and the differences between groups will be addressed for each condition separately. Thus, we aim at investigating possible statistically significant differences in speed and accuracy of interpretation across groups.

5.4.1.4 Statistical analysis on the broad time-window

The last statistical analysis was performed on a broad time-window aiming at identifying when, as the sentence unfolds, the two groups start to significantly differ in their fixations to the target picture. To do so, we considered each visual condition separately and we analyzed the effect of time-bin (in50ms steps) starting from the disambiguation point.

First, we considered the *All-Different* condition. The graph in Figure 27 shows the proportion of looks at the target picture for children (in red) and adults (in blue).

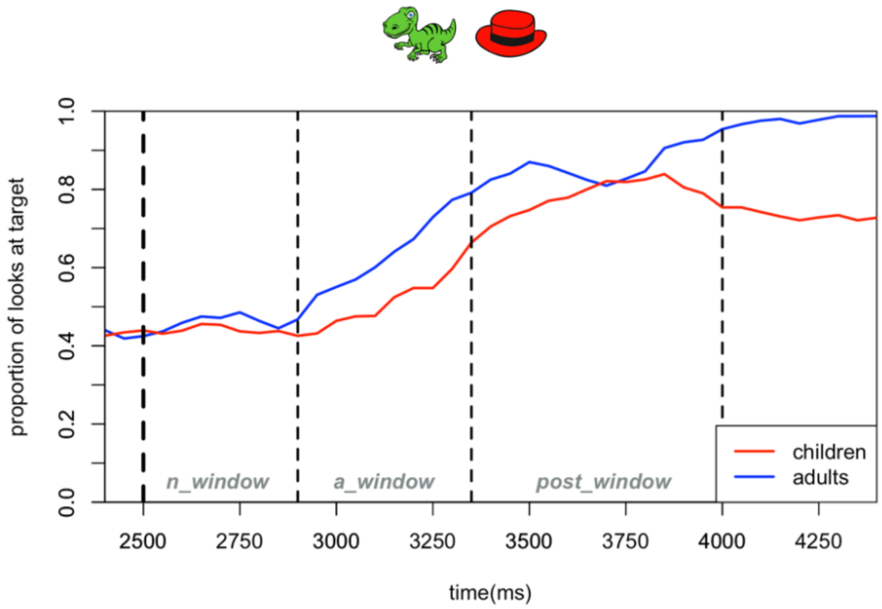


Figure 27. Proportion of looks at the target picture throughout the trial for the *All-Different* condition for children (in red) and adults (in blue). Vertical dashed lines represent mean onset and offset times. The bold vertical line indicates the disambiguation point. Bold text annotations indicate the broad time window considered for the analysis.

The graph in Figure 27 makes clear that, in a condition in which the noun is informative, adults start to process the informative word and shift their looks to the target before the end of the noun-window. Children, by contrast, show a weak delay as the curve starts to rise during the adjective window. In general, accuracy is very high for both groups. Adults reach 0.9 proportion of target fixations already around 3500ms, while children are slower but get to 0.8 around 3900ms.

To statistically verify the observed differences between groups, we used general additive mixed modelling in R (GAMMs), whose visual representation indicates when in time an effect becomes significant on a response variable. In addition, we included a parametric coefficient for the variable group, along with a random effect for *event* (i.e., the combination of subject and item as a unique identifier), allowing for a random intercept (see e.g., Porretta et al., 2016, Zahner et al., 2019).

We specified fixations to the target object as dependent variable and after a conversion to empirical logits (*elogs*, a logit transformed proportion, i.e., a ratio of the fixations to the referent object divided by the fixations directed to the competitor (Barr, 2008)). The model revealed an overall significant difference between groups (ß = -.5, SE = .01, t = -2.893, p < .01), localized in two time-windows: 2945 – 3560ms and 3853 – 4000ms, corresponding to the adjective-window and most of the post noun-phrase window. These localized effects of group on fixations to the target picture are shown in Figure 28 (the output of *GAMM* is only meaningful when visualized).

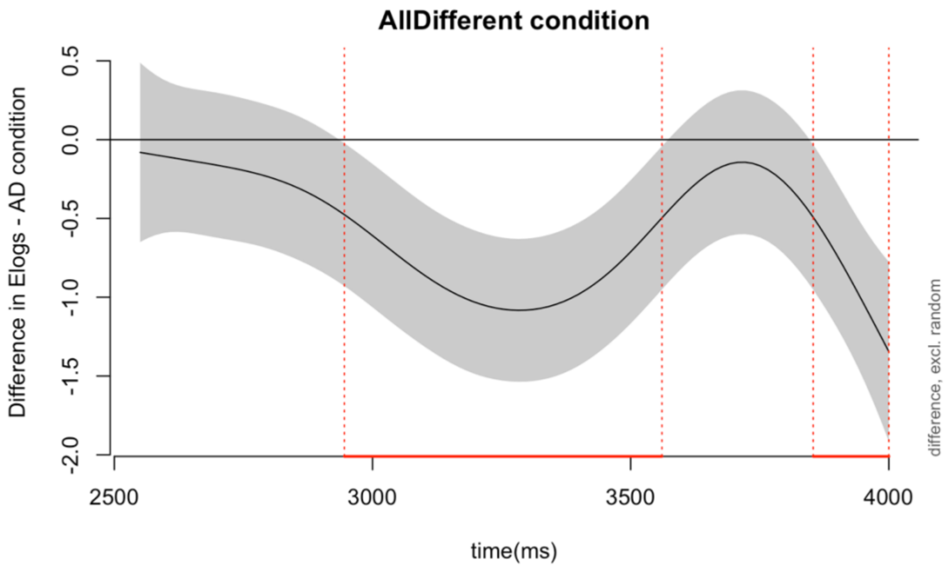


Figure 28. Difference curve in target fixations by adults minus children in the *All-Different* condition from the disambiguation point. The grey band indicates the 95% confidence interval (CI) of the mean difference. Values above zero indicate more target fixations by children. Conversely, values below zero indicate more target fixations by adults. The difference is significant if the 95% CI does not include zero (2945 – 3560ms and 3853 – 4000ms).

This result indicates that children differ significantly from adults from the adjective onset and until 400ms after the adjective offset, i.e., during the computation of the adjective- word. However, the two groups do not differ while listening to the noun, signaling an asymmetry between the interpretation of the two words. A further difference is found at the end of the post noun-phrase window, i.e., when adults reach the ceiling proportion of fixations while children get bored and lose attention.

The same analysis was conducted for the *Same-Color* condition. The graph in Figure 29 shows the proportion of looks at the target picture for children (in red) and adults (in blue).

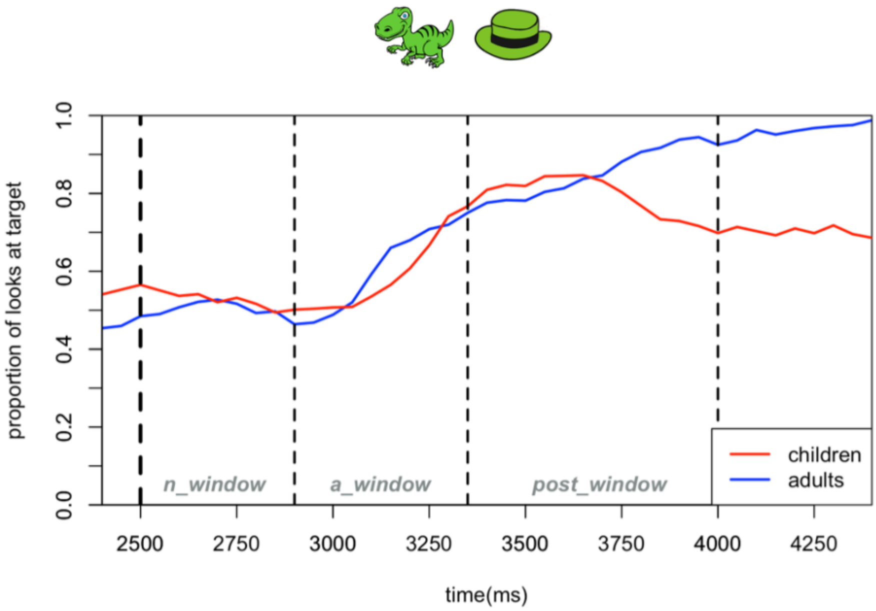


Figure 29. Proportion of looks at the target picture throughout the trial for the *Same-Color* condition for children (in red) and adults (in blue). Vertical dashed lines represent mean onset and offset times. The bold vertical line indicates the disambiguation point. Bold text annotations indicate the broad time window considered for the analysis.

The graph in Figure 29 shows how children's looking pattern does not seem to differ from the adults'. By the end of the noun-window, both groups start to increase their looks at the target picture and, while adults keep looking at the referent and reach 0.9 proportion of looks by 4000ms, children do not go over 0.8, reached at around 3600ms. Moreover, children do not seem to suffer from any disruption caused by the uninformative adjective. Rather, having an informative noun followed by an (uninformative) adjective that could be referred to both pictures seem to facilitate their performance, resulting in an adult-like pattern of interpretation.

This observation was confirmed by the statistical analysis. We calculated a general additive mixed model, specifying logit-transformed fixations to the target preference as dependent variable and a parametric coefficient for the variable group. Further, we added *event* as a random effect (random intercept). No significant effect of the parametric coefficient was found (p = .2), indicating that irrespective of time, the effect of group is not significant. Nevertheless, the model showed a localized difference in the time-window 3736 – 4000ms (see Figure 30).

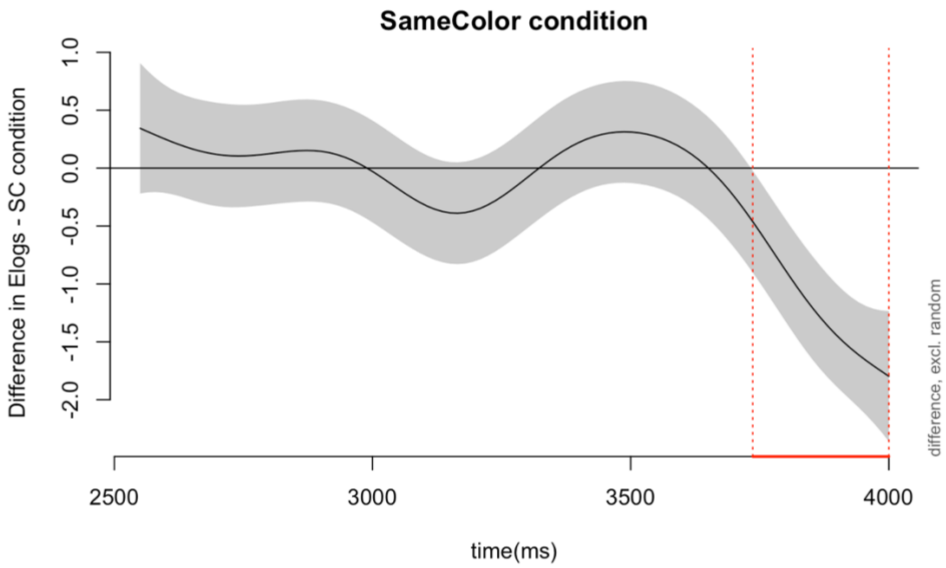


Figure 30. Difference curve in target fixations by adults minus children in the *Same-Color* condition from the disambiguation point. The grey band indicates the 95% confidence interval (CI) of the mean difference. Values above zero indicate more target fixations by children. Conversely, values below zero indicate more target fixations by adults. The difference is significant if the 95% CI does not include zero (3736 – 4000ms).

As predicted, the results from the analysis confirmed no differences between children and adults in the *Same-Color* condition. Once again, the model revealed a significant difference between groups at the end of the post noun-phrase window, attributable to the adults' engagement in the task on the one side, and to children's loss of attention on the other.

Lastly, we analyzed children's eye-movements in the *Same-Object* condition. Differently from the other two conditions, the broad time window here included the adjective- and the post noun-phrase window only, as the disambiguation point corresponded to the adjective onset. The graph in Figure 31 depicts the proportion of fixations at the target picture for children and adults.

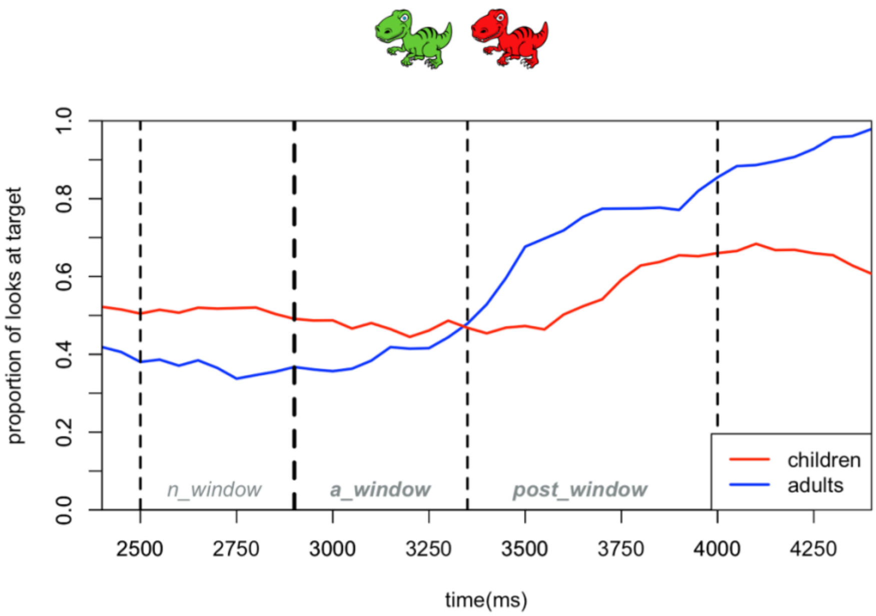


Figure 31. Proportion of looks at the target picture throughout the trial for the *Same-Object* condition for children (in red) and adults (in blue). Vertical dashed lines represent mean onset and offset times. The bold vertical line indicates the disambiguation point. Bold text annotations indicate the broad time window considered for the analysis.

A first inspection of the graphs in Figure 31 allows to make some preliminary observations concerning both accuracy and pattern of interpretation.

First, adults' accuracy is similar to the accuracy we found in the *Same-Color* and the *All-Different* conditions, overcoming 0.9 proportion of fixations at target around 4250ms. This delay in comparison to the first two conditions is due to the fact that, in this condition, the disambiguation point comes later in the sentence, corresponding to the adjective-onset. Surprisingly, children do not seem to be able to reach 0.7 proportion of fixations at target, opposite to the other two conditions in which their proportion of fixations at the referent object got to 0.8.

Second, adults in the *Same-Object* condition seem to show an unexpected looking pattern, at least until the adjective-offset, in that the proportion of looks at the target picture is below 0.4 and overcomes 0.5 only at the end of the adjective-window. Children, by contrast, stay around 0.5 (i.e., chance level) while listening to the uninformative noun and until they hear the informative adjective, and increase their fixations to the target object after the offset of the adjective. In the *Same-Object* condition, the noun-window corresponds to an uninformative word, therefore adults have no reason to look at the distractor object more than at the target picture. We believe this is the result of a low number of participants to the experiment in that stimuli have been balanced as for their position on the screen and as for their status as targets or distractors. Moreover, a more attentive analysis of the graph during the adjective-window makes evident that, while children keep looking back and forth the two pictures until around 3600ms, the adults' curve increases already during the adjective window, demonstrating that, although they were focused on the distractor, they rapidly processed the informative adjective and shifted towards the target object.

To verify these observations, we calculated a general additive mixed model, specifying logit-transformed fixations to the target preference as dependent variable and a parametric coefficient for the variable group. In addition, we added *event* as a random effect (random intercept). The model showed a significant effect of the parametric coefficient (ß = -.5, SE = .24, t = -2.302, $p < .01$), indicating that children and adults differed significantly in localized time-windows, i.e., 2950 – 3140ms (i.e., at the beginning of the adjective-window) and 3395 – 4000ms (i.e., during most of the post noun-phrase window), as shown in Figure 32.

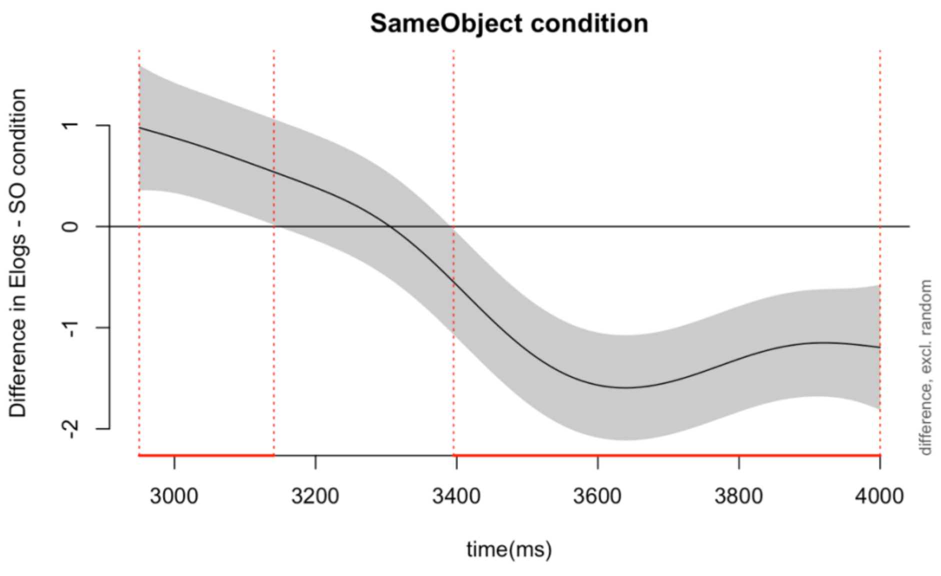


Figure 32. Difference curve in target fixations by adults minus children in the *Same-Object* condition from the disambiguation point. The grey band indicates the 95% confidence interval (CI) of the mean difference. Values above zero indicate more target fixations by children. Conversely, values below zero indicate more target fixations by adults. The difference is significant if the 95% CI does not include zero (3736 – 4000ms).

To sum up, the statistical analysis on eye-movement behavior confirmed that children and adults do differ in processing the informative adjective in the *Same-Object* condition and that this difference is largely observed after the adjective-offset. While in the first two conditions the informative noun was enough to resolve reference, independently of the pictures on the screen, in the *Same-Object* condition the noun did not provide any relevant information. Thus, we believe the difficulty children show in this condition is due to the informativeness difference between the *Same-Object* condition on the one hand, and the *Same-Color* and *All-Different* conditions on the other, that is, that this is the only condition in which adjective processing was *necessary* for reference resolution. Consequently, children are either slower or not able to compute adjective interpretation, resulting in a significant group difference that is specific of the *Same-Object* condition and does not concern neither the *Same-Color* nor the *All-Different* conditions to this extent.

In conclusion, the statistical analysis corroborated the difference between children and adults in the *Same-Object* condition, while mainly confirming our predications for the *Same-Color* and the *All-Different* conditions. Nevertheless, a more detailed investigation focusing on the group of children allows us to evaluate possible effects of development in reference resolution in the present task. We address this issue in the next §5.4.2.

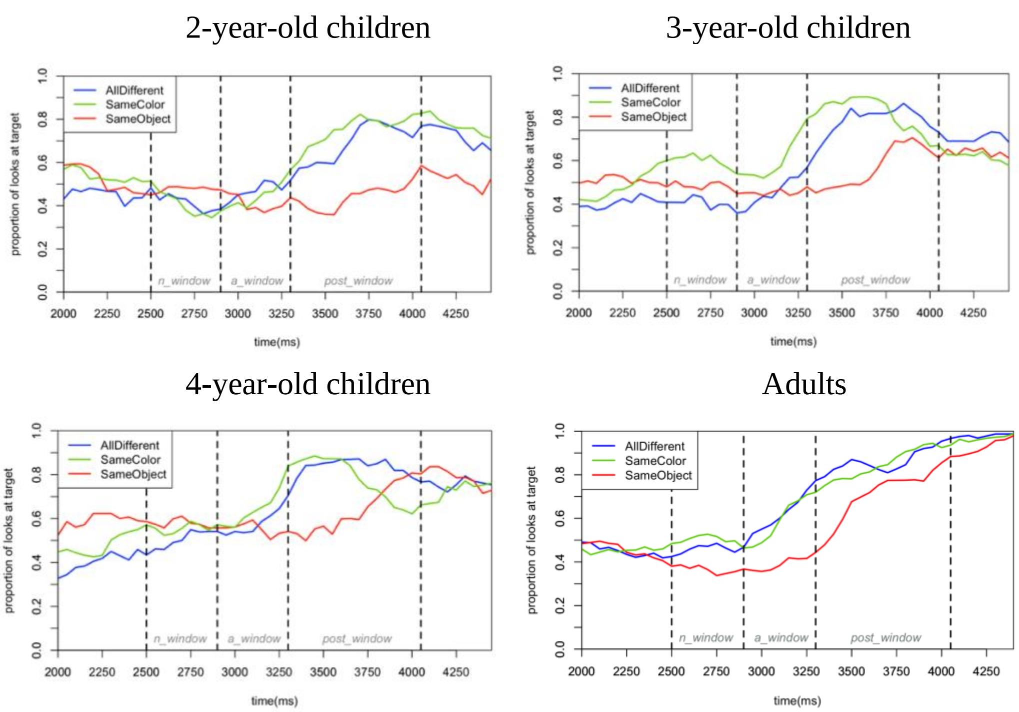
5.4.2 Children's processing of nouns and adjectives in development

The third research question (RQ3) asked whether children's processing of noun-adjective combinations in different informativeness conditions is dependent on age. The results from Experiment 1 revealed that 2-year-olds were significantly slower and less accurate with respect to 3-and 4-year-olds when asked to identify the target object in a four-picture scenario. These findings were interpreted as the result of processing difficulties due to the complexity of Experiment 1. Hence, with the aim of limiting age differences, in Experiment 2 we lowered processing costs by using only (intersective) color adjectives and by limiting the number of pictures in the visual scene.

To answer RQ3, we performed a new statistical analysis mirroring the general analysis conducted for RQ1 and RQ2. Once again, we aimed at identifying possible significant differences among informativeness conditions, together with possible age group effects.

In Experiment 2, children's age span from 28 months (2;4 years) to 63 months (5;3 years). As for Experiment 1, children were split into three age groups. This was done to investigate our research questions in development and to discuss our findings in comparison to Weisleder and Fernald (2009) and Fernald et al. (2010)'s, whose design was adapted to Italian for the present experiment[28]. The first group (2-year-olds) included children from 2;4 years to 3;2 years (N = 9; mean age = 2;10, S.D. = 0;2, F = 4). In the second group (3-year-olds), toddlers span between 3;3 years and 3;9 years (N = 10; mean age = 3;6, S.D. = 0;2, F = 6). The last group (4-year-olds) included children between 3;11 years and 5;2 years (N = 9; mean age = 4;4, S.D. = 0;6; F = 4).

The graphs in Figure 33 depict the proportion of looks at the referent pictures in the three informativeness condition for 2-, 3-, 4-year-old toddlers and adults.



[28] Weisleder and Fernald (2009) tested children between 3;2 and 3;8 months of age (mean age = 3;6). Fernald et al. (2010) compared two age-groups, i.e., 30-month-olds ranging between 2;4 and 2;8 (mean age = 2;6), and 36-month-olds between 2;11 and 3;3 (mean age = 3;0).

Figure 33. Proportion of looks at the target picture throughout the trial for 2-, 3-, 4-year-old children and for adults. Looks at target are marked in blue in the AD condition; in green in the SC condition; in red in the SO condition. The first vertical line indicates the average onset of the noun; the second vertical line indicates the average onset of the adjective; the third vertical line indicates the average offset of the adjective.

From a first visual inspection of the graphs in Figure 33, some preliminary observations can be made. First, from the steepness of the curves we can clearly observe that children get faster with age. While 2-year-olds's proportion of fixations at the target in the *All-Different* condition (in blue) and the *Same-Color* condition (in green) is just above 0.5 by the end of the adjective-window, 3- and 4-year-olds are above 0.7 in both conditions. As for accuracy, no evident difference can be observed between groups in the *All-Different* and in the *Same-Color* conditions. The hugest group difference, however, can be noticed in the *Same-Object* condition (in red). While 4-year-old children are almost as accurate as in the other two conditions (they get around 0.8 target fixation proportion by the end of the post noun-phrase window), 3-year-olds do not overcome 0.7, while the youngest 2-year-old children do not get to 0.6 and stay around the baseline for the whole trial, failing referent identification.

5.4.2.1 Statistical analysis on time-windows

As for the general analysis presented in §5.4.1, we started by analyzing the data within the **noun-window**. During this time-window, participants are listening to an informative word in the *Same-Color* and *All-Different* conditions, while the noun is uninformative in the *Same-Object* condition. To investigate possible differences among conditions and age-groups, we calculated a linear-mixed effects regression model with fixations to the target object as dependent variable and age-group and condition as independent variables with full interaction. Participants and items were added as crossed random effects. The model showed a significant main effect of condition ($p < .05$), but no significant effect of group ($p = .1$). Further, we found a significant interaction ($\chi^2 = 59.2$, df = 6, $p < .001$).

A second, parallel analysis was conducted on the **adjective-window**, by calculating a linear-mixed effects regression model with fixations to the target object as dependent variable and age-group and condition as fixed effects with full interaction. Participants and items were added as crossed random effects. The model revealed a significant effect of condition ($p < .05$), and a significant main effect of age-group ($p < .001$). Moreover, we found a significant interaction between condition and age-group ($\chi^2 = 105.3$, df = 6, $p < .001$).

A last analysis was performed on the **post noun-phrase window**. We calculated a linear-mixed effects regression model with fixations to the target object as dependent variable and age-group and condition as independent variables with full interaction. Participants and items were added as crossed random effects. The model showed a significant main effect of condition ($p < .001$) and of age-group ($p < .001$). In addition, we found a significant interaction ($\chi^2 = 127.6$, df = 6, $p < .001$).

To further investigate the interactions we found in all time-windows, in the next section (§5.4.2.2) we split the data according to the variable condition and we analyze the effect of age-group for each informativeness condition separately.

5.4.2.2 The *All-Different* condition

We started with the *All-Different* condition. The graph in Figure 34 shows the proportion of looks at the target picture for 2-year-olds (in red), 3-year-olds (in orange), 4-year-olds (in yellow) and adults (in blue).

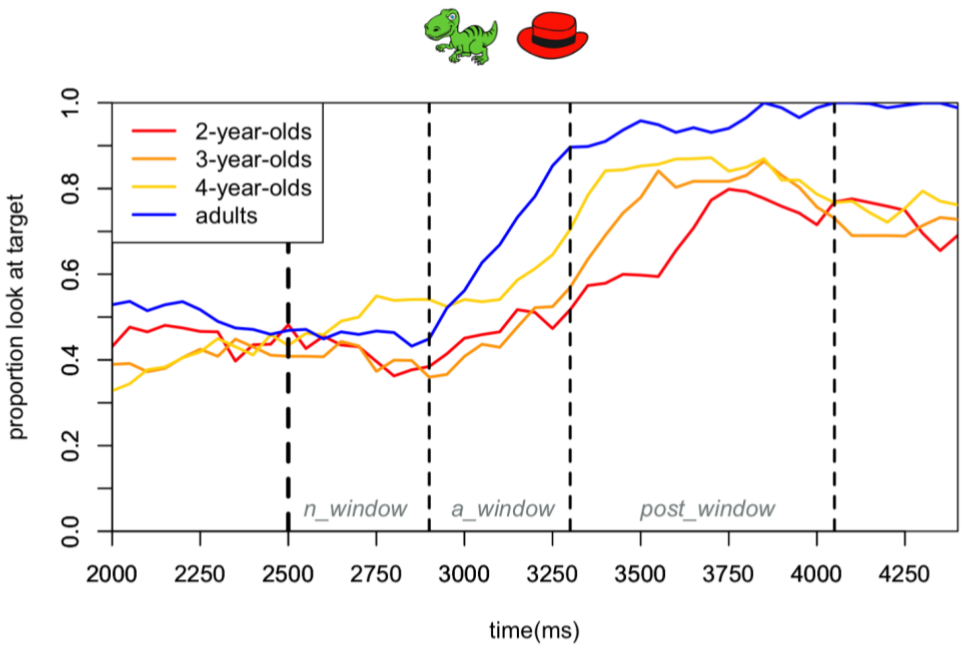


Figure 34. Proportion of looks at the target picture throughout the trial for the *All-Different* condition, for each age-group of children and for adults. Vertical dashed lines represent mean onset and offset times. The bold vertical line indicates the disambiguation point.

The graph in Figure 34 shows that in the *All-Different* condition 4-year-olds are faster and slightly more accurate than the other two groups of children. 2- and 3-year-olds present the same looking pattern until the end of the adjective-window, when 3-year-old children rapidly resolve reference and reach 0.8 proportion of fixations to the target. 2-year-olds, by contrast, are slower and get over 0.7 only around 3750ms (i.e., 450ms after the offset of the adjective). Adults stay around chance level until the end of the noun-window, when they rapidly shift to the target and get to 0.9 proportion of fixations to the referent object at the adjective-offset.

To verify these observations, we statistically analyzed the effect of age-group on fixations to the target picture using a linear-mixed effects regression model for each time-window. We specified fixations to the target object as dependent variable and age-group as fixed factor. Participants and items were added as crossed random effects. The models showed no significant effect of age-group in the noun-window (p = .5). The effect of group was significant in the adjective-window (2-year-olds vs. 3-year-olds, p = .7; **2-year-olds vs. 4- year-olds**, ß = 0.14, SE = .06, t = 2.243, p < .05; **2-year-olds vs. adults**, ß = 0.27, SE = .05, t = 4.589, p < .001; **3-year-olds vs. 4-year-olds**, ß = 0.12, SE = .06, t = 2.002, p = .05, **3-year-olds vs. adults**, ß = 0.25, SE = .05, t = 4.438, p < .001; **4-year-olds vs. adults**, ß = 0.12, SE = .05, t = 2.279, p < .05), and in the post noun-phrase window (2-year-olds vs. 3-year-olds, p = .06; **2-year-olds vs. 4- year-olds**, ß = 0.12, SE = .05, t = 2.150, p < .05; **2-year-olds vs. adults**, ß = 0.28, SE = .05, t = 5.172, p < .001; **3-year-olds vs. 4-year-olds**, ß = 0.4, SE = .06, t = 0.625, p = .05, **3-year-olds vs. adults**, ß = 0.16, SE = .06, t = 2.668; p < .05; **4-year-olds vs. adults**, ß = 0.12, SE = .06, t = 1.948, p = .05).

To sum up, in the *All-Different* condition, while no group-difference emerged in the noun-window, during the adjective and the post noun-phrase windows 4-year-old children outperformed 2- and 3-year-olds and were found to be almost target-like. However, adults were always significantly faster than all groups of children.

5.4.2.3 The *Same-Color* condition

The graph in Figure 35 shows the proportion of looks at the target picture for 2-year-olds (in red), 3-year-olds (in orange), 4-year-olds (in yellow) and adults (in blue) in the *Same-Color* condition.

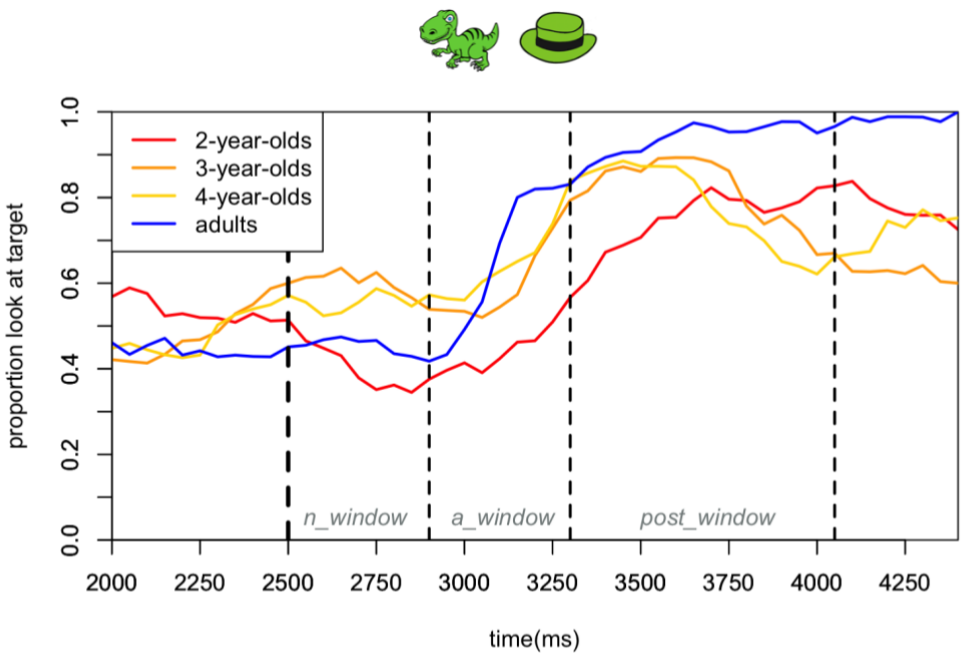


Figure 35. Proportion of looks at the target picture throughout the trial for the *Same-Color* condition, for each age-group of children and for adults. Vertical dashed lines represent mean onset and offset times. The bold vertical line indicates the disambiguation point.

From the observation of the graph in Figure 35, we can clearly see that 2-year-olds are overall slower than 3- and 4-year-old children. While the oldest children get to 0.8 proportions of fixations to the target picture by the end of the adjective-window, the youngest toddlers are still around 0.6. However, their fixations to the target keep increasing and they get over 0.8 at around 3750ms (i.e., 450ms after the end of the sentence). As for the *All-Different* conditions, adults' fixations around chance level quickly shifted to the target picture at the end of the noun-window and got over 0.8 proportion of fixations by the end of the adjective-window.

These observations were verified with a linear-mixed effects regression model for each time-window. We specified fixations to the target object as dependent variable and age-group as fixed factor. Participants and items were added as crossed random effects. The model showed a significant effect of age-group in the noun-window (**2-year-olds vs. 3-year-olds**, ß = .20, SE = .06, *t* = 2.931, *p* < .01; **2-year-old vs. 4-year-olds**, ß = 0.18, SE = .06, *t* = 2.611, *p* < .05; 2-year-old vs. adults**,** *p* = .3; 3-year-old vs. 4-year-olds, *p* = .7; **3-year-old vs. adults,** ß = -.14, SE = .05, *t* = -2.474, *p* < .05; **4-year-old vs. adults,** ß = -.12, SE = .05, *t* = -2.096, *p* < .05), in the adjective-window (**2-year-olds vs. 3-year-olds**, ß = .16, SE = .06, *t* = 2.430, *p* < .05; **2-year-old vs. 4-year-olds**, ß = 0.22, SE = .06, *t* = 3.337, *p* < .01; **2-year-old vs. adults**, ß = 0.23, SE = .06, *t* = 3.773, *p* < .001; 3-year-old vs. 4-year-olds, *p* = .3; 3-year-old vs. adults, *p* = .2; 4-year-old vs. adults, *p* = .9), and in the post noun-phrase window (2-year-olds vs. 3-year-olds, *p* = .1; **2-year-old vs. 4-year-olds**, ß = 0.14, SE = .06, *t* = 2.152, *p* < .05; **2-year-old vs. adults**, ß = 0.26, SE = .06, *t* = 4.260, *p* < .001; **3-year-old vs. 4-year-olds**, ß = 0.04, SE = .06, *t* = 0.625, *p* = .05; **3-year-old vs. adults,** ß = .16, SE = .06, *t* = 2.668, *p* < .05; **4-year-old vs. adults,** ß = .12, SE = .06, *t* = 1.948, *p* = .05).

In summary, the effect of age-group was found in all time-windows for the *Same-Color*. Differently from our findings in the *All-Different* condition (see §5.4.2.2), it seems that 3-year-old children are facilitated by a condition in which the adjective is uninformative. This resulted in a pattern of interpretation analogous to that of 4-year-olds. Both groups were found to be significantly faster in target identification with respect to 2-year-olds. This difference, however, is reduced as the sentence unfolds and does not result in a higher accuracy. Adults, by contrast, are always faster and more accurate.

5.4.2.4 The *Same-Object* condition

Lastly, the graph in Figure 36 shows the proportion of looks at the target picture for 2-year-olds (in red), 3-year-olds (in orange), 4-year-olds (in yellow) and adults (in blue), in the *Same-Object* condition.

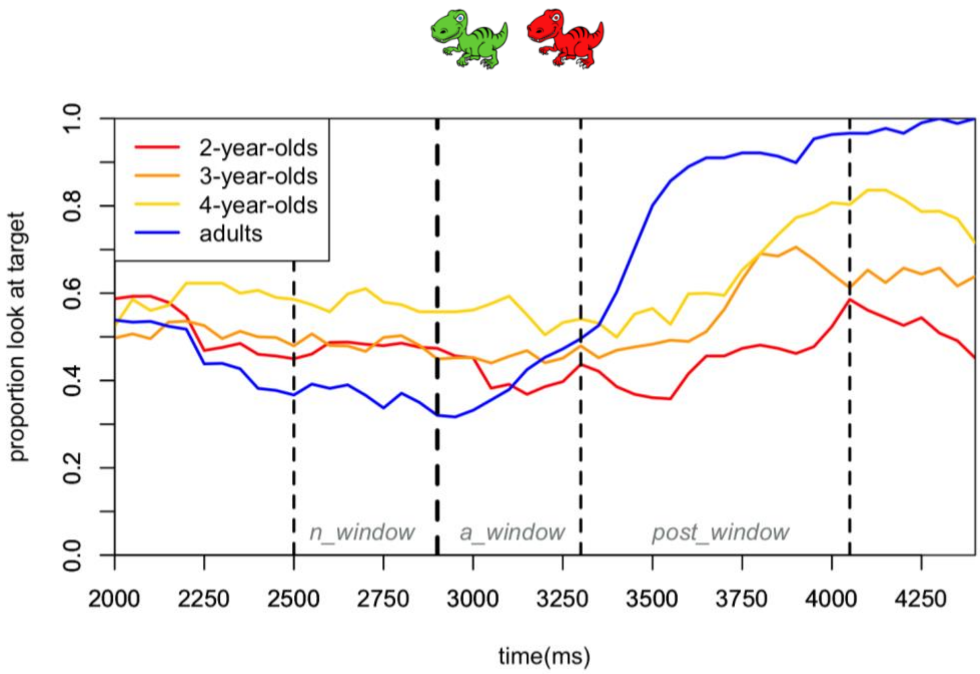


Figure 36. Proportion of looks at the target picture throughout the trial for the *Same-Object* condition, for each age-group of children and for adults. Vertical dashed lines represent mean onset and offset times. The bold vertical line indicates the disambiguation point.

A visual inspection of the graph in Figure 36 shows that all three groups' proportion of fixations to the target picture are around 0.5 while listening to the uninformative noun and until the end of the informative adjective. By the adjective-offset, however, 4-year-olds rapidly shift their gaze towards the referent picture and get to 0.8 proportion of fixations to the target around 700ms after sentence-offset. 3-year-old children seem to be as fast as older children, but accuracy is lower as they do not overcome 0.7 proportion of fixations. Interestingly, 2-year-old toddlers seem to stay around 0.5 proportion of fixations throughout the duration of the trial and never get to or above 0.6.

Using a linear-mixed effects regression model, we verified these observations for each time-window. We specified fixations to the target object as dependent variable and age-group as fixed factor. Participants and items were added as crossed random effects. The model showed no significant effect of age-group during the adjective-window (p = .1). During the post noun-phrase window a significant effect of age-group emerges (2-year-olds vs. 3-year-olds, p = .1; **2-year-old vs. 4-year-olds**, ß = .14, SE = .06, t = 2.152, p < .05; **2-year-old vs. adults,** ß

= .26, SE = .06, *t* = 4.260, *p* < .001; **3-year-old vs. 4-year-olds**, ß = .04, SE = .06, *t* = 0.625, *p* = .05; **3-year-old vs. adults,** ß = .16, SE = .06, *t* = 2.668, *p* < .05; **4-year-old vs. adults,** ß =.12, SE = .06, *t* = 1.948, *p* = .05).

To summarize, all groups of children keep looking at both pictures on the screen at chance until the offset of the adjective, when 4-year-olds shift to the target object and rapidly resolve reference, being almost target-like. Interestingly, 2-year-olds remain at chance level and fail target identification. 3-year-olds are in between the other two groups. Adults, by contrast, increase their fixations to the target during the adjective-window and result to be overall faster and more accurate than all groups of children. We will discuss the implication of this finding in §5.5.

5.5 Discussion

Experiment 2 investigated children's interpretation of nouns and adjectives in different informativeness conditions. We designed a Visual World eye-tracking study manipulating the linguistic stimulus in combination with objects of the screen. This was done by reducing the complexity of Experiment 1, i.e., by lowering the processing overload associated with both the different semantic classes of adjectives and the 4-picture visual display. Here, in Experiment 2, we presented participants with only two objects on the screen, and we asked Italian monolingual participants a question containing a noun and a color-adjective (e.g., *Dov'è il dinosauro verde?*, lit. "Where is the dinosaur green?"). Thus, the visual conditions in the present experiment varied according to the informativeness of the noun or the adjective with respect to the visual scenario, aiming at answering three research questions.

5.5.1 Children's interpretation of informative nouns and adjectives

The first research question (RQ1) asked whether children show a *Noun-Anchoring* mechanism in interpreting noun-adjective combinations by looking at two pictures in different informativeness conditions.

The *All-Different* condition presented participants with e.g., a green dinosaur and a red hat. Thus, the two objects differed in both category and color. Consequently, in this condition both the noun (i.e., "dinosaur") and the adjective (i.e., “green”) are informative, as they are uniquely referrable to the target object. In line with our predictions, we found that children rapidly shifted their gaze towards the target object after hearing the informative noun, without the need to go back to the distractor after the over-informative adjective.

In the *Same-Color* condition, participants saw, e.g., a green dinosaur and a green hat. Hence, while the noun was informative (as in the previous condition) the adjective here was uninformative, as it could be referred to both objects. Through the analysis of fixations, we found that children processed the informative noun after hearing it and kept looking at the target picture throughout the trial, as we hypothesized. Contrary to the predictions made by the *Sequential Integration Hypothesis*, but in line with the results found by Weisleder and Fernald (2009) for Spanish, children did not suffer from any disruption due to the uninformative adjective. Rather, in this condition children identified the target object shortly after the noun-offset and, surprisingly, more rapidly than in the *All-Different* condition. When the color-adjective was uninformative, children were facilitated in reference resolution. The reason of this facilitation may be due the fact that the two objects in the *Same-Color* condition only differ in one dimension (i.e., the category), while in the *All-Different* condition the two pictures varied in both the category and the color. This second more complex visual scenario might have negatively affected processing at first, as children needed more time to explore the screen. Nevertheless, they were able to rapidly recover and equal the accuracy of the *Same-Color* condition.

The *Same-Object* condition asked participants to identify the target object while looking at, e.g., a green dinosaur and a red dinosaur. Differently from the two previous conditions, the noun here was not informative and the interpretation of the adjective was required to resolve reference. Our results revealed that, contrary to our predictions, children were not able to replicate their performance in the other two conditions. Indeed, they kept looking at both pictures at chance until the offset of the informative adjective, when they slowly started to increase their fixations to the target object but failed to reach the accuracy of the *All-Different* and the *Same-Color* conditions.

These results match the predictions made by the *Noun-Anchor Hypothesis*. Indeed, by comparing the *Same-Object* to the previous two, we found a computational asymmetry at the expense of the adjective. Thus, when adjective interpretation is essential to complete the task, children are less accurate than in the conditions in which it is not, showing processing difficulties exclusively linked to adjective interpretation. This result contrasts with Weisleder and Fernald (2009), who found that 3;6-year-old Spanish speaking toddlers were able to shift their gaze rapidly and accurately towards the target object even in the condition in which the noun was not informative. Although the children's mean age in our experiment (i.e., 3;7 years) was very similar to that of the children in Weisleder and Fernald's (i.e., 3;6 years), children's age range was different (2;4 – 5;2 in our experiment, 3;2 – 3;8 in Weisleder and Fernald's), negatively influencing our overall results. Nevertheless, testing a wider age-range of children allowed for possible age differences in this respect to be detected. We will see that this was the case in our Experiment (§5.5.3).

5.5.2 Adults' interpretation of informative nouns and adjectives

The second research question (RQ2) investigated possible differences between children and adults in the online processing of referring expressions with adjectives. This was done to make sure that children's data were the result of processing difficulties and were not linked to task effects.

Results from adults matched our predictions. In the *All-Different* and in the *Same-Color* condition (i.e., when the noun was informative about the referential scene), adults were able to rapidly identify the target upon hearing the noun and increased their fixations to the referent picture throughout the rest of the trial. Furthermore, in the *Same-Object* condition, adults listened through the uninformative noun and shifted their gaze towards the target picture after the adjective was heard. In contrast with children's data, accuracy was at ceiling in all three conditions, as predicted.

5.5.3 Interpreting nouns and adjectives in development

The third research question (RQ3) explored the effect of age on children's interpretation of noun-adjective combinations across different informativeness conditions. As predicted, in the *All-Different* condition all groups of children successfully resolved reference. However, 4-year-olds were found to be faster than 3-year-olds, who were faster than 2-year-olds. Adults were overall more rapid and more accurate, as expected.

A significant effect of age-group was found in the *Same-Color* condition. In line with the *All-Different* condition, 4-year-olds outperformed 2-year-olds, but, here, 3-year-olds were found to be as fast as 4-year-olds. As we observed for the group of children in general in §5.5.1, the *Same-Color* condition facilitated target identification in comparison to the *All-Different* condition. This could be attributable to the shared color-property of target and distractor of the *Same-Color* condition, in contrast to the different-object/different-color opposition in the *All-Different* condition. From the analysis of age-groups, however, we found that this facilitation only regards 3-year-olds but does not affect 2- and 4-year-olds.

The most revealing result, though, came from the *Same-Object* condition. As expected, no significant group differences were found during the adjective-window, which, in this condition, corresponded to the first informative word of the combination. After the end of the sentence, however, an interesting pattern emerged. 4-year-olds were able to resolve reference as accurately as in the other two conditions. 2-year-olds, by contrast, failed the task and kept looking at both pictures around chance. 3-year-olds' performance was in between the other two groups, corroborating the hypothesis that they are improving their processing skills but are still not as good as children one year older. Thus, by comparing the results from the three visual conditions it is evident that the youngest children indeed show an asymmetry between noun and adjective processing, in line with the *Noun-Anchor Hypothesis*. These findings provide evidence of an evolution in development mainly concerning adjective interpretation, signaling a processing difficulty which affects the youngest children, decreases with age, and disappears during the fourth year.

5.5.4 Conclusion

In conclusion, the present study added to previous literature by exploring different age groups of children's interpretations of Italian nouns and adjectives in different visual conditions. Our results revealed a surprising effect of noun and adjective informativeness, showing how children as young as 2;4 years rapidly and accurately process informative nouns. When the computation of the adjective was necessary to resolve reference, however, a developmental pattern emerged. 4-year-old children are able to interpret adjectives as fast and as accurately as nouns, while 2-year-olds show an asymmetry, failing adjective interpretation and, consequently, task resolution. The present study adds to the previous literature on the topic by providing evidence of a developing pattern of adjective processing in a Visual World paradigm where only familiar color adjectives were used.

6 Experiment 3: Interpreting adjectives in prenominal predicative use

6.1 Introduction

A key controversy in the study of language comprehension is how and at what point in a speech stream language users integrate different types of information. Before the seminal study by Tanenhaus et al. (1995), most psycholinguistic investigations of sentence processing focused on the examination of reading times and eye-movements during reading. These studies relied on the assumption that longer reading times of words in a sentence, as well as eye movement patterns across text, can be used to quantify moment-by-moment increases and decreases in processing difficulty. In such reading tasks, contextual manipulations were inevitably limited and typically involved changes in the preceding sentence or discourse (see Frazier, 1995; MacDonald et al., 1994; for review).

The Visual World paradigm has proven to be a very useful tool to investigate cognitive processing and provided ample evidence that adults interpret language fast and incrementally, that is, that words are processed as they are being heard and immediately integrated with the information from the visual scene (e.g., Tanenhaus et al., 1995; Hagoort et al., 2009; Rayner & Clifton, 2009). Furthermore, adults can predict how discourse will unfold on the basis of semantic information associated with a word currently being processed (e.g., Altmann & Kamide, 1999; Sedivy et al., 1999; Kamide et al., 2003, Knoeferle et al., 2005; Kukona et al., 2011).

The first Visual World study addressing this issue with adults was developed by Altmann and Kamide (1999). They presented participants with semirealistic scenes depicting, e.g., a boy, a cake and some toys, while hearing a sentence such as *The boy will move the cake* or *The boy will eat the cake*. They found that eye movements to the cake (the only edible object) started earlier in the "eat" than in the "move" condition. Altmann and Kamide (1999) interpreted these results as evidence that information conveyed by the verb meaning can be used to anticipate the upcoming word, demonstrating the importance of eye movement behavior in prediction during language processing. Kamide et al. (2003) found that also information conveyed by the grammatical subject and by the verb can jointly drive anticipatory eye-movements. They found increased fixations to a motorbike when participants heard *The man will ride...* but increased fixations to a carousel when participants heard *The girl will ride....*

Sedivy and colleagues (1999) were among the first to show that, in response to modified nouns, e.g., *The tall glass*, adults routinely resolve reference during the prenominal adjective before the noun has been produced. They do so by exploiting the relationship between the linguistic input, the nonlinguistic context, and their knowledge of referential principles. This effect in adults was documented by early looks to a target member of a contrast set, e.g., a tall glass, alongside a short glass in the presence of a singleton object that was also tall, e.g., a tall jug.

In sum, language-mediated eye-movements reflect not only linguistic processing nor only processing of the visual scene, but continuously updated mental representations based on information derived from sentence structure, word meanings and the visual input (Huettig et al., 2011).

Little is known about how children process language in real time and whether they make use of the context during interpretation. The paucity of child research is partly due to the fact that most techniques for studying language processing have relied upon reading skills of reading, an ability that young children do not have or are only beginning to acquire. In recent years, however, the use of eye-tracking has increasingly used to study language comprehension, allowing for very young infants and toddlers to be tested. To the best of our knowledge, few studies explicitly investigated children's predictive understanding, only a couple of which focused on prenominal adjectives using the Visual World paradigm (e.g., Tribushinina & Mak, 2016; Davies et al., 2021; see Ch. 3 for an overview of the Visual World paradigm).

Tribushinina and Mak (2016) investigated preschoolers' ability to integrate meaning from adjectives and nouns, and to predict the noun based on adjective meaning. They tested whether Dutch 3-year-olds (3;0 – 3;5) could integrate properties of adjectives with relevant objects by measuring whether children looked at the target referent upon hearing the prenominal adjective. When looking at, e.g., a stone and a butterfly, Tribushinina and Mak (2016) indeed found that toddlers showed an early preference for the target object (e.g., a stone): looks to the target increased on hearing the adjective when it was informative (e.g., *heavy*). Unsurprisingly, instead, looks to the target increased only on hearing the noun when the adjective was not informative (e.g., *new*). Tribushinina and Mak (2016) concluded that children younger than four years of age can process adjectives as they hear them and even predict the noun based on the adjective meaning and, contrarily to results of Ninio (2004), adjectives are not *always* processed after the head-noun. Their conclusions, however, present some limitations. First, language stimuli were recorded with a 3-second break between the onset of the adjective and the onset of the noun, leaving children the time to consciously process and compute the adjective meaning before hearing the adjective. Hence, in this respect, Tribushinina and Mak (2016)'s result does not mirror natural language processing. Second, their study aimed to understand if the conceptual knowledge associated with the adjective can help the incremental processing of adjective-noun combinations. Thus, informativity in their study only depended on world knowledge, e.g., knowing that a stone is typically heavy whereas a competitor (e.g., a butterfly) is not. The items used, as well as the kind of informativity they consider, do not allow for a full comprehension of adjective-noun combination processing, as they only reflect a limited type of adjectival modification. Thus, Tribushinina and Mak (2016)'s findings cannot be generalized to adjective-noun combinations, as the successful predictive processing they found could be limited to the cases in which the adjective expressed an inherent property of the object (e.g., being *heavy* for a stone). Comparing the online interpretation of adjectives expressing inherent (e.g., *heavy* for a stone) or non-inherent properties of the referent object (e.g., being *blue* for a sofa) would allow for a more complete understanding of incremental processing. In this latter case, the visual screen would need to be exploited for task resolution (e.g., to identify the blue object) and could result in a different process and timing of interpretation with respect to the condition in which the property is not visible on the screen (e.g., to identify the heavy object). This issue is addressed in Experiment 3.

A further attempt to explore the online interpretation of prenominal adjectives was the study conducted by Davies and colleagues (2021), who investigated the effect of prenominal contrastive adjectives of size (i.e., *big* and *small*), by comparing anticipatory fixations at "the *big* cow" in the presence or absence of a contrastive set. They tested 36 English monolingual children (3;6 – 4;0) but failed to find evidence of incremental processing, as participants were able to resolve reference only by the end of the utterance, i.e., after hearing the noun, when the stimulus sentence was presented at a natural speed. To allow participants more time for processing, Davies et al. (2021) conducted a second experiment in which they manipulated the auditory stimuli by inserting 500ms of silence between the offset of the adjective and the onset of the noun. As expected, this manipulation allowed children to contrastively infer successfully, as they showed a stronger preference for the target during the adjective in "contrast displays" than in "non-contrast displays". The authors concluded that 3-year-olds need ample time to demonstrate these sophisticated online language processing skills but argued that their finding is more robust than conclusions from previous research with the same age-group. Previous studies made use of referential tasks that could be passed using world knowledge (Tribushinina & Mak, 2016) or adjective information alone, where the integration with the noun was not required (e.g., Fernald et al., 2010, using the *Looking-while-Listening* Paradigm).

The current study takes on a new approach aiming to enrich the scarce literature on the topic and surpass the limitations of previous research.

6.2 The current study

As already mentioned, there is abundant evidence in the literature demonstrating that online language processing in adults is fast and incremental, that is, words are processed as they are heard, and their meaning is integrated with the world knowledge and with the information provided by the visual scene. In addition, adults can predict how a sentence will unfold based on the semantic meaning of the words they are processing. However, whether children also possess these abilities is still open to debate.

Predictive processing has often been addressed by testing the interpretation of prenominal adjectives. This investigation, however, has so far been confined to languages with prenominal adjectives, as English or Dutch. To test the online interpretation of adjectives in prenominal position in Italian, we designed an experiment with predicative yes/no questions which, in Italian, present the adjective-noun order, e.g., *È morbido il cuscino?* (lit., 'Is soft the pillow?', "Is the pillow soft?"). While looking at two pictures on the screen (e.g., a white pillow and a white bone), participants are presented with questions containing adjectives that are either informative (e.g., *soft*) or uninfomative (e.g., *white*) about the referent object.

The aim of the present study is to investigate whether children between two and five years of age can process language incrementally, i.e., if they are able to process adjectives as they are being heard and to predict the following noun based on the adjective lexical meaning. We target this issue in an eye-tracking experiment, comparing eye-movement behavior of children and adults.

The following research questions will be addressed:

1) Are children able to predict the noun based on the pre-nominal adjective in Italian predicative constructions? Do they show a *Noun-Anchoring* mechanism in the interpretation of prenominal adjectives? Does their looking pattern differ from the adults'?
2) Are there differences between the interpretation of adjectives expressing an inherent property of the object and do not have a visual representation (e.g., *soft*, for a pillow), and adjectives that are visible on the screen (e.g., *open*, for a window)? Do adults show processing differences in this respect?
3) If children and adults differ with respect to the first two research questions, do 2-, 3- and 4-year-old children show developmental differences in processing informative vs. uninformative adjectives? And in the interpretation of inherent vs. non-inherent properties of objects?

6.2.1 Predictions

The present study contributes to the literature on adjective acquisition by investigating how children, in comparison to adults, interpret adjective-noun combinations in Italian by means of the unmarked pre-nominal adjective in predicative yes/no questions, aiming at answering new research questions in this respect.

6.2.1.1 Predictions for Research Question 1

The first research question (RQ1) asks whether children interpret adjective-noun phrases incrementally and whether they can predict the noun based on the preceding adjective meaning. Since the literature on the topic is not straightforward in this respect, different alternative predictions should be considered.

First, based on Ninio's (2004) proposal, young children interpret noun-adjective phrases in two steps, i.e., noun-first and adjective-later, and have difficulties integrating the two, resulting in delayed or omitted adjective interpretation. Our experiment presents children with two pictures on the screen (e.g., a white pillow and a white bone) and asks participants yes/no questions with prenominal adjectives that are either informative or uninformative about the following noun (see Table 9). What the *Noun-Anchor Hypothesis* predicts in this respect is that, in a condition in which the prenominal adjective is informative about the referent object, children should wait to respond until after hearing the noun. For instance, upon hearing *È morbido il cuscino?* ('Is soft the pillow?'[29]), children should keep fixating both pictures around chance level until they hear *pillow*, ignoring the informative adjective *soft*. The same prediction applies to the condition in which the adjective is not informative. For example, while hearing *È bianco il cuscino?* ('Is white the pillow?'), children should keep looking back and forth the two (white) objects until they hear the informative noun and shift to the labeled picture only by the end of the stimulus question. Thus, the *Noun-Anchor Hypothesis* predicts no difference between the informative- and the uninformative-adjective conditions.

---

[29] To ease comprehension of the Italian stimulus sentences in the present experiment, from now on we report the example sentences either using the word-to-word English translation with the prenominal adjective, e.g., **Is soft the pillow*? ("Is the pillow soft?") or the Italian version, e.g., *È morbido il cuscino?*

According to Tribushinina and Mak (2016)'s results, however, children do have the ability to interpret language incrementally. Hence, if this is the case, we expect the informative- and the uninformative-adjective conditions to reveal different looking patterns. Specifically, when the adjective is informative (e.g., *soft*), children should start fixating the target picture above chance level before listening through the noun, opposite to the conditions in which the adjective is not informative (e.g., *white*), in which children are expected to resolve reference by the end of the referring expression.

As for adults, ample evidence in the literature has shown that they interpret language incrementally. Thus, we expect them to process the first informative word they hear, be it the adjective or the noun and, as a result, to shift their gaze to the target object shortly after the disambiguation point.

6.2.1.2 Predictions for Research Question 2

The second research question (RQ2) investigates possible processing differences linked to the prototypicality of the adjective. In the present study, we control for the visual representation of the property labeled by the adjective by presenting participants with either inherent prototypical properties of the objects on the screen (e.g., *soft* for a pillow) or non-prototypical properties (e.g., *closed* for a window). This manipulation does not only affect the linguistic stimulus, but also its visual representation. Indeed, the properties labeled by prototypical adjectives (e.g., *soft*) are not visually represented on the screen, while non-prototypical properties (e.g., *closed*) are, and can only be interpreted by means of the visual scene. Consequently, looking at, e.g., a pillow and a bone on the screen is not necessary to determine whether *soft* can uniquely be referred to the object pillow. By contrast, exploring the visual scene is required to verify whether the mentioned window is *closed* or not. Since interpreting prototypical properties of objects requires knowledge of the world, while non-prototypical properties are linked to the context, we expect to find differences in their computation. One possibility is that, in line with Tribushinina and Mak (2016)'s results, relying on conceptual knowledge may facilitate adjective interpretation, as children would only need to retrieve salient properties of familiar objects, eased by adjective-noun co-occurrence statistics in child-directed speech. A second possibility is that the visual representation of a property on the screen may facilitate adjective interpretation, as it requires the exploration of the visual scene without the need to recover previous knowledge, which might be costlier and take more time. A further possibility, however, is that the interpretation of prototypical and non-prototypical adjectives does not entail different processing pattern and that, being familiar with both types of properties, children are able to interpret them equally quickly. This third possibility is what we predict for the group of adults. Since familiar and salient pictures are presented, we expect them to easily identify the referent object, regardless of its visual representation on the screen.

6.2.1.3 Predictions for Research Question 3

The third research question (RQ3) tests whether the potential difference children and adults show with respect to RQ1 and RQ2 is due to an emerging processing ability. Hence, we ask whether children show age differences with respect to informative and uninformative adjectives interpretation and in relation to the integration of noun meaning with inherent and non-inherent properties of objects[30].

In line with our findings in Experiment 1 and 2, we predict the youngest children (i.e., 2-year-olds) to show more processing difficulties than the older children and to delay target identification, even in conditions in which the adjective is informative. By contrast, the oldest children (i.e., 4-year-olds) are predicted to be adult-like both in speed and in accuracy of interpretation. Hence, 3-year-olds are expected to lay in between, corroborating the developmental pattern emerged in the previous studies.

The same age differences are expected for the interpretations of prototypical and non-prototypical adjectives, in line with the predictions for the RQ2. We expect older children to process prototypical adjectives (e.g., *soft*) as rapidly and as

[30] Henceforth, we will refer to the dichotomy between inherent and non-inherent properties with the term *prototypicality*. Prototypical adjectives describe an inherent property of the objects that is not visible on the screen (e.g., *soft* for a pillow). Non-prototypical adjectives label properties that are visible on the screen (e.g., *open* for a window) (see Section 6.3.1).

accurately as non-prototypical ones (e.g., *closed*). By contrast, we predict the youngest children to take advantage of the visual representation of the labeled property, having less developed linguistic skills in comparison to older children, and, as a result, to be faster and more accurate with non-prototypical properties.

6.3 Methods

6.3.1 Materials

Experiment 3 was inspired by Tribushinina & Mak (2016), but the experimental design was modified and adapted to Italian. The items consisted of a sentence paired with two colored pictures presented on the screen simultaneously. The visual stimuli were based on 32 digitalized colored drawings on a white background, consisting of an object or an animal (see Table 9). The position of the objects on the screen was pseudo-randomized and counterbalanced across trials. The auditory stimuli were recorded by an Italian native speaker with pragmatically neutral intonation and processed in Praat (version 6.0.49; Boersma & Weenink, 2016). The duration of each item was comparable in length (M = 1507ms). The language stimuli consisted of a yes/no-question in Italian, e.g., *È morbido il cuscino?* ("Is the pillow soft?"), starting with the third person singular of the verb 'to be' ("Is"), followed by an unmarked pre-nominal adjective, a determiner, and a noun. Half of the nouns were feminine, and half were masculine.

The visual scene presented a target object (e.g., a white pillow) and a competitor (e.g., a white bone). Prenominal adjectives in the stimulus questions were either informative (e.g., *soft*) or uninformative (e.g., *white*) about the following noun (see Table 9).

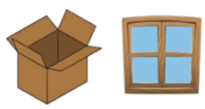
Table 9. Examples of linguistic and visual stimuli for the informative and the uninformative conditions.

Informativeness	
Informative	Uninformative

È morbido il cuscino? "Is the pillow soft?"	*È bianco il cuscino?* "Is the pillow white?"

Furthermore, adjectives were either prototypical or non-prototypical. Prototypical adjectives were adjectives labeling properties that are prototypically linked to the referent objects (e.g., the softness of pillows) and not depicted in the visual scene. Non-prototypical adjectives, by contrast, labeled properties that can arbitrarily be referred to the referent objects (e.g., the openness of windows) and are visible on the screen. In contrast to the interpretation of prototypical adjectives, in this latter condition adjective interpretation is only inferable from the visual cues (see Table 10). Thus, all stimuli totaled 2 informativeness-conditions crossed by 2 prototypicality-conditions.

Table 10. Examples of linguistic and visual stimuli for the prototypical and the non-prototypical conditions.

Prototypicality	
Prototypical	Non-prototypical
È morbido il cuscino? "Is the pillow soft?"	*È chiusa la finestra?* "Is the window closed?" 

To keep participants engaged and avoid loss of attention, for each informativeness-condition, the yes/no question could either elicit a 'yes' (true) or a ‘no’ (false) answer, resulting in a total of 4 informativeness-conditions, summarized in Table 11[31].

Table 11. Stimulus conditions with examples.

Prototypicality	**Truth**	*Informative*	*Uninformative*
Prototypical	True	*È morbido il cuscino?* “Is the pillow soft?”	*È bianco il cuscino?* “Is the pillow white?”
	False	*È morbido l’osso?* “Is the bone soft?”	*È viola l’osso?* “Is the bone purple?”
Non-prototypical	True	*È aperta la finestra?* “Is the window open?”	*È aperta la scatola?* “Is the box open?”
	False	*È chiusa la finestra?* “Is the window closed?”	*È chiusa la scatola?* “Is the box closed?”

Visual and auditory stimuli were combined to form 64 trials, divided into 4 lists of 16 trials each. All stimuli are listed in Table 12.

Table 12. List of items used in the experiment, divided by prototypicality and informativeness condition.

	Prot.	Nouns		Adjectives			
				Informative		Uninformative	
1	PROT	*cuscino, osso*	‘pillow’, ‘bone’	morbido	‘soft’	*bianco*	‘white’
2		*coniglio, ragno*	‘bunny’, ‘spider’	*spaventoso*	‘scary’	*peloso*	‘hairy’
3		*serpente, rospo*	‘snake’, ‘toad’	*lungo*	‘long’	*verde*	‘green’
4		*topo, elefante*	‘mice’, ‘elephant’	*piccolo*	‘small’	*grigio*	‘grey’
5		*caramella, casa*	‘candy’, ‘house’	*dolce*	‘sweet’	*gialla*	‘yellow’
6		*moto, lumaca*	‘bike’, ‘snail’	*veloce*	‘fast’	*disegnata*	‘drawn’
7		*piuma, valigia*	‘feather’, ‘bag’	*leggera*	‘light’	*rossa*	‘red’
8		*torre, forchetta*	‘flower’, ‘fork’	*alta*	‘tall’	*grigia*	‘grey’
9	NON-PROT	*divano, tamburo*	‘sofa’, ‘drum’	*viola*	‘purple’	*azzurro*	‘blue’
10		*gatto, bambino*	‘cat’, ‘boy’	*sveglio*	‘awake’	*addormentato*	‘asleep’
11		*guanto, cappello*	‘glove’, ‘hat	*pulito*	‘clean’	*sporco*	‘dirty’
12		*pennarello, orologio*	‘crayon, ‘clock’	*blu*	‘blue’	*giallo*	‘yellow’
13		*candela, televisione*	‘candle’, ‘tv’	*spenta*	‘off’	*accesa*	‘on’
14		*chiave, sedia*	‘key’, ‘chair’	*verde*	‘green’	*marrone*	‘brown’
15		*cintura, scopa*	‘belt’, ‘broom’	*grigia*	‘grey’	*rossa*	‘red’
16		*finestra, scatola*	‘window’, ‘box’	*chiusa*	‘closed’	*aperta*	‘open’

[31] The uninformative-false condition was added to balance the yes/no answers to the questions. Since this condition does not provide any insight about adjective processing, it won’t be included in the analysis.

6.3.2 Participants

The participants were thirty-nine typically developing Italian children (2;4-5;3, mean age = 3;6, S.D. = 0;75; 20 female). Two additional participants were excluded due to inattentiveness during the trial or fussiness (i.e., they failed at looking at either picture for more than half of the trials). There were also twenty-four Italian adult participants (19;1-29;9, mean age = 25;4, S.D. = 2;6; 17 female), who participated as controls.

The study was approved by the local Ethics Committee of the University of Verona and was conducted in accordance with the standards specified in the 2013 Declaration of Helsinki.

6.3.3 Apparatus and procedure

The experiment was run on a SR Research Eyelink 1000 Plus eye-tracker, sampling monocularly at 500 Hz with a 16mm lens, set in a head-free remote mode. The experiment was implemented in Experiment Builder. The visual items were presented on a 24" colour BenQ monitor, while two loudspeakers on both sides of the screen played the language stimuli. Both children and adult participants were tested individually, in a dimly lit and soundproof testing room at the Laboratory of Text, Language and Cognition (LaTeC) of the University of Verona. Before starting, a small sticker was placed on the participant's forehead, in order to track head movements. Calibration and validation procedures were carried out using a five-point display at the beginning of the experiment and a drift correction was repeated once every four trials.

Adult participants were not given a task. They were told that the experiment was created for children and that they had to sit and remain as still as possible throughout the duration of the experiment. Children were sitting on their caretaker's lap in front of the screen and were invited to play a game with a cartoon character, Peppa Pig, whose voice would ask them questions about the drawings appearing on the screen and keep their attention asking filler questions (e.g., "Do you like this game?") after each drift correction. Before the experiment began, a 5-point calibration and validation were performed, followed by a familiarization phase.

Participants were shown 4 “warm-ups” in which a single image was labeled by a sentence played over speakers (e.g., “Look! A penguin!”) and appeared on one of the two halves of the screen, familiarizing the child with the two sides of the screen in which the drawings will appear. After the familiarization phase, the experimental session began. In each trial, two pictures appeared on the screen simultaneously. The auditory stimulus started after 2000ms. From the onset of the sentence, pictures remained displayed on the screen for 3000ms, followed by an 800ms blank screen that ended the trial. Answers to yes/no questions were not recorded. An example of the trial is shown in Figure 37. A drift correction was performed every four trials, paired with one of the three filler sentences (e.g., “Are you having fun?”), recorded in a child-friendly intonation. The testing session lasted approximately 5-10 minutes.

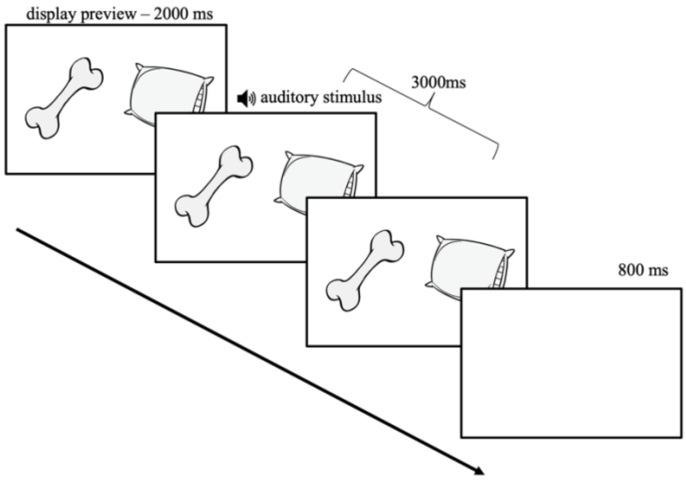


Figure 37. Experimental procedure.

6.4 Results

The eye-movements data recorded by the Eyelink system were extracted though the SR Research software Data Viewer and prepared for data analysis in R (R Development Core Team, 2019). Data were analyzed using the packages lme4 and lmerTest in R (Bates et al. 2015, Kuznetsova et al. 2017, lmer function), LMERConvenienceFunctions (Tremblay & Ransijn 2015, summary function), car (Fox & Weissberg 2011, Anova function), itsadug and mgcv (Wood, 2006; Baayen et al., 2017; Wood, 2017; Baayen er al., 2018, bam and logit functions). Before data analysis, we removed the trials in which neither picture was fixated, before or after the auditory stimulus was displayed. This led to the removal of 2,9% of the data. Furthermore, we removed those trials in which the participant fixated the screen outside the two areas of interest for more than half the duration of the trial. This led to the removal of another 6,1% of the data.

From the eye-tracking record, we determined the position of the eye in 50ms steps. For each 50ms-bin, we aggregated proportions of fixations to the two areas of interests, i.e., target and distractor. The proportion of looking time was accessed over three time-windows as the sentence unfolds. The *adjective-window* captured fixations starting at the mean adjective-onset (i.e., 2200ms) and ending at the average noun onset (i.e., 2700ms). The *noun-window* corresponded to the mean duration of the noun, starting from the average noun-onset, and ending at the average sentence offset (i.e., 2750-3200ms). Finally, the *post-stimulus window* accessed looking behavior for 800ms from the average noun-offset (i.e., 3200-4000ms).

The analysis of the results is organized as follows. §6.4.1 will analyze the data to answer RQ1, i.e., to investigate the role of informative prenominal adjectives in children's and adults' interpretation of adjective-noun phrases. §6.4.2 will explore how adjectives labeling prototypical vs. non-prototypical properties are processed by children and adults in comparison. In §6.4.3 the group of children will be investigated, aiming to identify a possible developmental pattern with respect to RQ1 and RQ2.

6.4.1 The role of the informativeness of the adjective

We started data analysis by exploring the role of informative prenominal adjectives in predicting the meaning of the following noun (RQ1). First, we performed a statistical analysis on the two informative adjective conditions, i.e., the informative-true and informative-false conditions (e.g., 'Is *soft* the pillow?' (yes) vs. 'Is *soft* the bone?' (no)). In both these conditions the adjective could only be referred to one of the pictures on the screen and, therefore, was informative about the following noun. Thus, we aimed to establish a baseline, expecting to find the same eye-movement behavior in the two conditions within groups. In other words, we expect the ability to predict the noun would emerge independently of whether the answer to the stimulus question is ‘yes’ or ‘no’ and thus, no difference in looking pattern was expected.

The second part of the analysis considered two different informativeness conditions, i.e., the informative-true and the uninformative-true conditions (e.g., ‘Is *soft* the pillow?’ (yes) vs. ‘Is *white* the pillow?’ (yes)). While the answer to the stimulus question was ‘yes’ for both conditions, at the adjective level only the informative-true condition allowed for target identification, while the uninformative-true condition provided an information that was relevant for both pictures. Thus, this second analysis aimed at establishing whether children can predict the noun based on adjective lexical meaning, by comparing two conditions in which the adjective was either informative or not.

6.4.1.1 The informative conditions

The first part of the analysis considered the informative-true (IT) and the informative-false (IF) conditions, i.e., the conditions in which participants saw, e.g., a pillow and a bone, and heard ‘Is *soft* the pillow?’ (IT) or ‘Is *soft* the bone?’ (IF).

Figure 38 depicts the proportion of fixations to the picture possessing the labeled property (e.g., the pillow) over the time course of the trial for children (on the left) and adults (on the right).

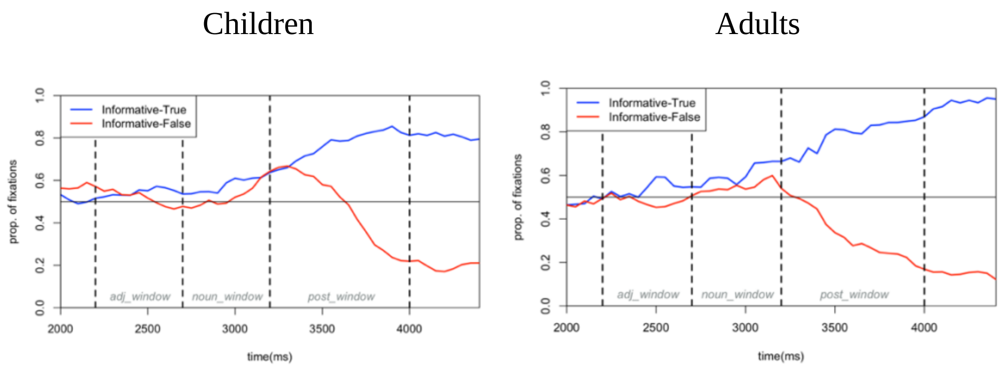


Figure 38. Proportion of fixations to the picture possessing the labeled property (e.g., the pillow) over time in the informative-true condition (plotted in blue) and in the informative-false condition (plotted in red). The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average onset of the noun, the third vertical line indicates the average offset of the sentence. The horizontal line indicates chance level.

A visual inspection of Figure 38 shows that, in both conditions, during the adjective-window (e.g., while hearing *soft*) children keep looking at both pictures around chance level and start to increase their fixations to the picture with the labeled property (e.g., the pillow) while listening to the noun (i.e., around 2900ms), getting above 0.6 proportion of fixations to the picture around the noun-offset. However, fixations to the, e.g., pillow, in the informative-false condition start to drop around 200ms after the sentence-offset and get below chance level at around 3600ms, i.e., 400ms after noun-offset. Adults, by contrast, start to fixate the picture with the labeled property above chance already during the adjective-window in the informative-true condition, and at the noun-onset in the informative-false condition. This difference adults show between conditions during the adjective-window is surprising, as the same word was being heard in the presence of the same visual context. Thus, we expect this difference to be non-significant. Nevertheless, as expected, fixations to the, e.g., pillow, in the informative-false condition rapidly drop already at 3200ms, i.e., at the noun-offset, showing that participants were indeed engaged in sentence processing.

These observations were further confirmed by the graphs in Figure 39.

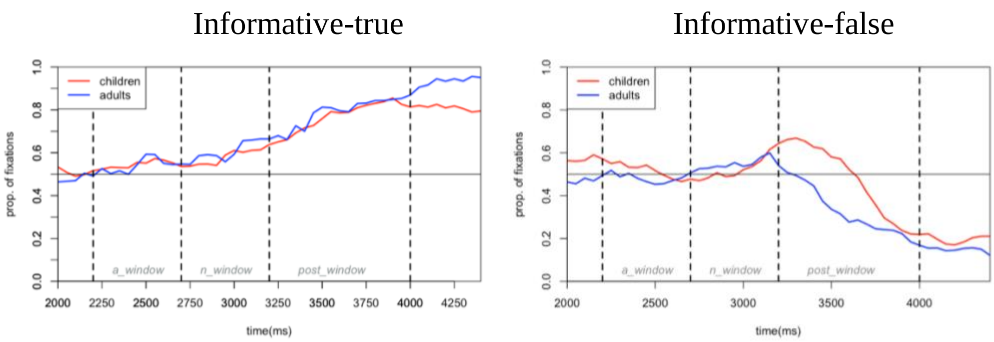


Figure 39. Proportion of fixations to the picture possessing the labeled property (e.g., the pillow) over time in the informative-true condition (on the left) and in the informative-false condition (on the right) for children (plotted in red) and adults (plotted in blue). The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average onset of the noun, the third vertical line indicates the average offset of the sentence. The horizontal line indicates chance level.

Figure 39 shows that no difference in looking pattern emerges between children and adults in the informative-true condition, neither for both speed nor for accuracy. By contrast, in the informative-false condition an evident group difference is observed after the offset of the noun. Indeed, while adults rapidly process the noun (e.g., *bone*) and immediately drop their fixations to the picture with the labeled property (e.g., the pillow), children are slower and shift to the labeled object visibly later.

To verify these observations, data were analyzed using a mixed effect regression model for each time-window. We started by analyzing eye-movements during the **adjective-window**. We specified proportion of fixations to the picture possessing the labeled property as dependent variable; group and informativeness-condition as fixed effects with full interaction; participant and item as crossed random effects (random intercepts). The model revealed no significant main effect of condition (p = .3), nor group (p = .4). No interaction was found (p = .4).

A second analysis was performed on the **noun-window**, specifying proportion of fixations to the picture possessing the labeled property as dependent variable and group and informativeness-condition as fixed effects with full interaction. Participant and item were added as crossed random effects (random intercepts). The model revealed a significant main effect of condition (ß = .09, SE = .03, t = -2.681, p < .01), but no significant effect of group (p = .7), nor interaction (p = .5).

The third analysis was performed on the **post-stimulus window**. We specified fixations to the picture possessing the labeled property as dependent variable; group and informativeness-condition as fixed effects with full interaction; participant and item as crossed random effects (random intercepts). The model showed a significant main effect of condition (ß = .49, SE = .02, t = 19.52, p < .001) and of group (ß = .14, SE = .03, t = 3.925, p < .001). Further, the interaction between condition and group was significant (χ^2 = 32.9, df = 1, p < 0.001). To further investigate the nature of the interaction, we split the data according to the variable condition and we calculated a linear mixed effects regression model specifying fixations to the picture possessing the labeled property as dependent variable and group as fixed effect; participant and item as crossed random effects (random intercepts). The model showed no effect of group in the informative-true condition (p = .6), but a significant main effect of group in the informative-false condition (ß = .13, SE = .04, t = 3.005, p < .01).

In summary, the analysis of the data revealed that, in the conditions in which the adjective is informative, children and adults do not differ in eye-movement behavior. Nevertheless, a group difference was found in the informative-false condition, i.e., when the property labeled by the informative adjective (e.g., *soft*) did not match the object labeled by the noun (e.g., *bone*). In this case, children were slower in identifying the referent object and, thus, to direct their gaze to the labeled object (e.g., the picture of the bone), while dropping their fixations to the object matching the relevant property (e.g., the picture of the pillow). Although no comparison between an informative and an uninformative adjective condition was performed, children's delay in switching to the labeled object suggests that they were indeed processing the prenominal informative adjective. This comparison will be the objective of the next section.

6.4.1.2 Informative vs. uninformative adjectives

The second part of the statistical analysis was performed on the informative-true (IT) and uninformative-true (UT) conditions, i.e., we compared conditions in which, upon looking at, e.g., a white pillow and a white bone, participants were asked a question about the target picture (e.g., the pillow), containing either an informative (e.g., *soft*) or an uninformative (e.g., *white*) prenominal adjective. Figure 40 depicts the proportion of fixations to the target picture for children (on the left) and adults (on the right).

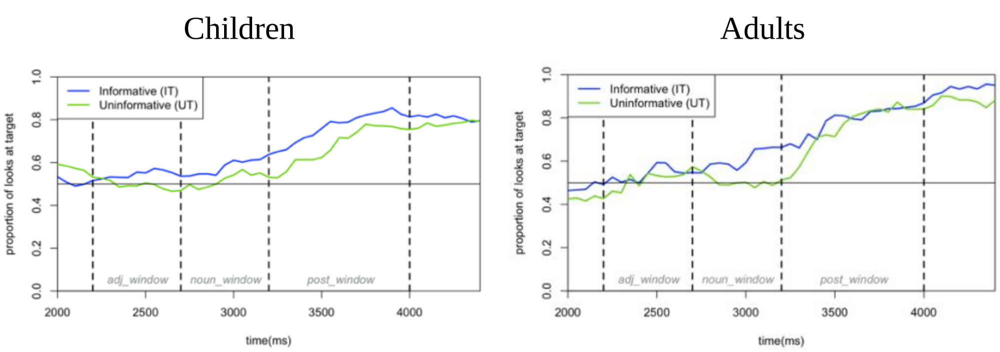


Figure 40. Proportions of fixations to the target picture (e.g., pillow) over time in the informative-true condition (plotted in blue) and in the uninformative-true condition (plotted in green). The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average onset of the noun, the third vertical line indicates the average offset of the sentence. The horizontal line indicates chance level.

A first visual inspection of the graphs indicates that having an informative adjective makes target identification faster for both groups. Indeed, as observed in the previous section (§6.4.1.1), in the informative-adjective condition (IT, in blue) both groups start to increase their fixations to the target picture (e.g., the pillow) right after the adjective-offset (e.g., *soft*). When the adjective is not informative about the target referent (e.g., *white*), i.e., in the UT condition (in green), participants need to listen through the noun (e.g., *pillow*) to identify the object on the screen, as the adjective labels a property relevant to both pictures (e.g., the target pillow and the distractor bone).

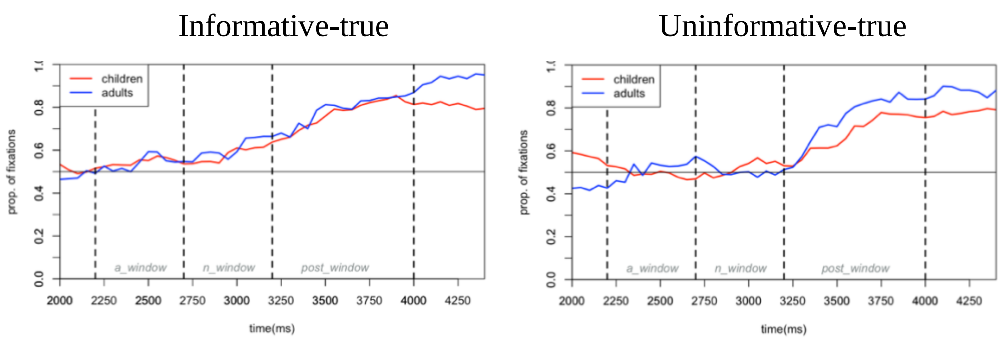


Figure 41. Proportion of fixations to the target picture over time in the informative-true condition (on the left) and in the uninformative-true condition (on the right) for children (plotted in red) and adults (plotted in blue). The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average onset of the noun, the third vertical line indicates the average offset of the sentence. The horizontal line indicates chance level.

Figure 41 shows the proportion of fixation to the target picture separately for the informative and uninformative conditions. The statistical analysis in §6.4.1.1 showed no effect of group in the informative condition. As for the uninformative condition, no group difference seems to emerge, at least until sentence-offset, when adults seem to interpret the informative noun faster than children.

To statistically verify these observations, data were analyzed using a series of mixed effect regression models for each time-window. First, we analyzed eye-movements within the **adjective-window**. During this time-window participants heard an adjective that was either informative (e.g., *soft*) or uninformative (e.g., *white*) about the referent object. Thus, if they can make use of the adjective information to predict the following noun, an effect of condition should emerge. We calculated a mixed effect regression model specifying fixations to the target picture as dependent variable; group and informativeness-condition as fixed effects with full interaction; participant and item as crossed random effects (random intercepts). The model showed a significant effect of informativeness-condition (ß = - .09, SE = .03, $t = -2.586$, $p < .01$), but no effect of group ($p < .9$). No interaction was found ($p = .5$)

Second, fixations were analyzed within the **noun-window**. In this time-window participants heard an over-informative noun in the IT condition, while new, relevant information is provided after the uninformative adjective in the UT condition. We specified fixations to the target picture as dependent variable. Group and informativeness-condition were added as fixed effects with full interaction, and participant and item as crossed random effects (random intercepts). The model revealed a significant effect of informativeness-condition (ß = - .14, SE = .03, $t = -3.893$, $p < .001$), but, again, no effect of group ($p = .5$). No interaction was found ($p = .4$)

Lastly, we analyzed the **post-stimulus window**. We specified fixations to the target picture as dependent variable; group and informativeness-condition as fixed effects with full interaction; participant and item as crossed random effects (random intercepts). The model revealed no significant main effect of informativeness-condition ($p = .5$) nor group ($p = .6$). However, we found a significant interaction ($\chi^2 = 8.694$, df = 1, $p < .01$). Since no group effect was found in the informative-true condition (p = .6) (see §6.4.1.1), the nature of this interaction was further investigated by calculating a mixed effect regression model on the uninformative condition only, specifying fixations to the target picture as dependent variable and group as fixed effect; participant and item as crossed random effects (random intercepts). The model showed a significant main effect of group (ß = - .09, SE = .04, $t = 2.078$, $p < .05$).

To investigate when in time, as the stimulus unfolded, fixations to the target picture in the two informativeness conditions started to diverge, we calculated a general additive mixed model. We included a parametric coefficient for the variable condition, and a random effect (random intercept) for *event* (i.e., the combination of subject and item as a unique identifier) and we specified fixation to the target object as dependent variable, after a conversion to empirical logits. For children, the model showed a significant effect of condition (ß = -0.5, SE = .23, t = -2.205, p < .05), localized in the time window 2780 – 3805ms, i.e., from 80ms after the noun-onset until 600ms after sentence offset. For adults, no significant effect of the parametric coefficient was found (p = .1), indicating that irrespective of time, the effect of condition is overall not significant. Nevertheless, the model showed a localized difference in the time-window 2815 – 3363ms, i.e., from the 100ms after noun-onset to 100ms after sentence offset. The output of the two models in reported in Figure 42.

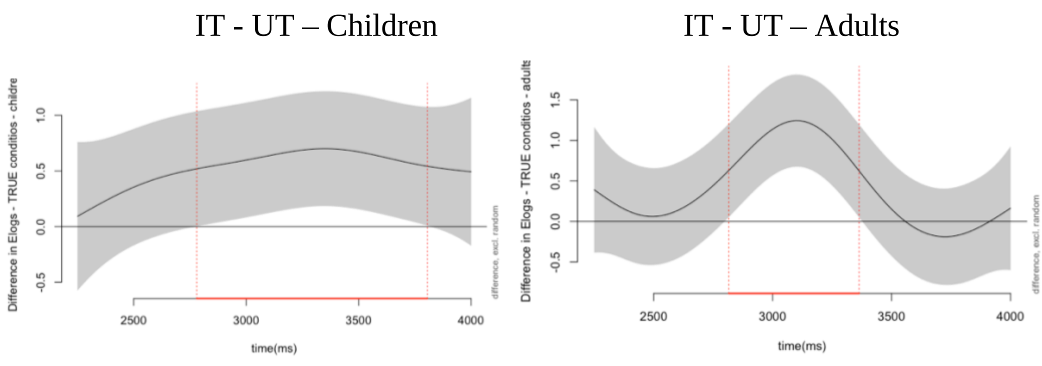


Figure 42. Difference curve in fixations to the target by children (on the left) and adults (on the right) in the informative-true condition minus the uninformative-true condition. The grey band indicates the 95% confidence interval (CI) of the mean difference. Values above zero indicate more fixations to the target in the informative-true condition. Conversely, values below zero indicate more target fixations in the uninformative-true condition. The difference is significant if the 95% CI does not include zero (2780 – 3805ms for children; 2815 – 3363ms for adults).

To sum up, the statistical analysis on the informative vs. uninformative conditions revealed that both groups do make use of the prenominal informative adjective to predict the following noun. Interestingly, no significant difference between groups is found, signaling that children are as fast as the adult controls in predicting the noun based on the adjective meaning. A weakly significant effect of group was found after the end of the stimulus question, but only in the uninformative condition, showing that, when the adjective is informative, children are facilitated and faster in target identification than in conditions in which only the noun is informative, in line with what we found in the comparison between the two informative conditions in §6.4.1.1. Lastly, although a significant difference between conditions was found in each time-window, the analysis on fixations at the target picture over the time-course of the trial revealed no effect of condition for adults, showing that the effect of informativeness of the adjective was stronger for the children than for the adults. We will get back to this result in §6.4.3 and discuss its implications in §6.5.3.

6.4.2 Processing prototypical vs. non-prototypical adjectives

In the present section, we aim to answer RQ2, i.e., to investigate a possible effect of the prototypicality of the adjective on sentence processing and, consequently, eye-movements. In the current study, we manipulated prenominal adjectives as for their status as inherent or non-inherent properties of the referred object. This manipulation also involved the visual representation of the objects on the screen. Indeed, while prototypical adjectives described properties that were *not* visible on the screen (e.g., *soft* for a pillow, *fast* for a motorcycle), non-prototypical adjectives labeled properties that were salient on the visual context (e.g., *closed* for a window, *blue* for a sofa). Hence, while interpreting non-prototypical adjectives required participants to explore the visual scene and to refer to it to answer the stimulus question, the interpretation of prototypical adjectives could have been based on world knowledge only. If the prototypicality vs. non-prototypicality of the adjective influences looking behavior, this effect can only be detected if the adjective is informative about the noun. For this reason, in the present section we will only consider the informative-true and the informative-false conditions.

Figure 43 reports the proportion of looks at target in the informative-true condition for children and adults in the prototypical vs. non-prototypical adjective conditions.

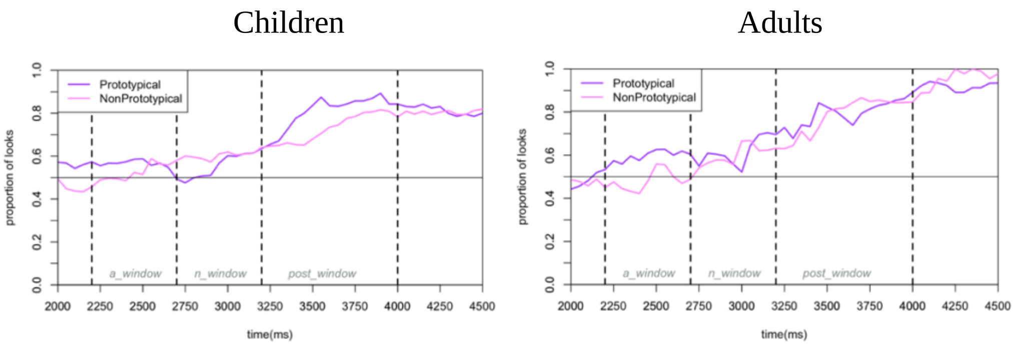


Figure 43. Proportions of fixations to the target picture over time in the informative-true condition for children and adults in the prototypical adjective condition (plotted in purple) and in the non-prototypical adjective condition (plotted in pink). The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average onset of the noun, the third vertical line indicates the average offset of the sentence. The horizontal line indicates chance level.

Figure 43 indicates no huge difference between the prototypical and the non-prototypical adjective conditions for adults, while children seem to take advantage of the prototypicality of the adjective, as they are surprisingly more rapid in target identification when the adjective is prototypical in comparison to when it is not. To verify whether that was indeed the case, we performed a linear mixed effects regression model including the three time-windows, specifying looks at target as dependent variable and group and prototypicality-condition as fixed effects. Participant and items were added as random effects. The model revealed no significant main effect of group (p = .1), nor prototypicality-condition (p = .9).

Figure 44 depicts the proportion of looks at the picture with the labeled property in the informative-false condition for children and adults in the prototypical vs. non-prototypical adjective conditions.

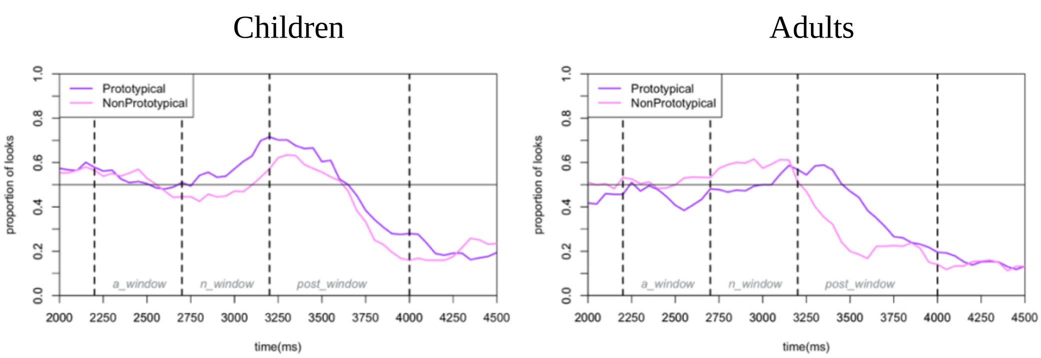


Figure 44. Proportions of fixations to the picture possessing the labeled property over time in the informative-false condition for children and adults in the prototypical adjective condition (plotted in purple) and in the non-prototypical adjective condition (plotted in pink). The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average onset of the noun, the third vertical line indicates the average offset of the sentence. The horizontal line indicates chance level.

A visual inspection of Figure 44 suggests that children are facilitated in processing prototypical adjectives, as they process prototypical informative adjectives faster. This difference, however, decreases with time. Adults, by contrast, interpret the non-prototypical adjective faster and, in this condition, are more rapid in processing the following noun and, thus, drop fixations to the picture possessing the labeled property. To verify these observations, we performed linear mixed effects regression model, specifying looks at the object with the labeled property as dependent variable and group and prototypicality-condition as fixed effects. Participant and items were added as random effects. The model revealed no significant main effect of group (p = .08), nor prototypicality-condition (p = .5).

In conclusion, the statistical analysis on the prototypicality of the adjectives showed that the visual representation of a property on the screen does not facilitate nor hamper processing. In other words, relying on world knowledge did not take more time than evaluating the visual context, neither to adults, nor to children.

6.4.3 Developing the ability to predict nouns based on adjective meaning

The third research question (RQ3) asked whether, within the group of children, our data reveal a developmental pattern in the processing of informative vs. uninformative prenominal adjectives and in the interpretation of prototypical vs. non-prototypical properties of objects. Opposite to what we found in Experiments 1 and 2, no significant difference between children and adults has emerged for the informativeness of the adjectives and the prototypicality of the properties. This is contrary to our initial predictions for RQ1 and RQ2. Our findings showed that children process predicative yes/no questions with prenominal adjectives as fast and as accurately as adults and are able to make use of the information from the prenominal adjective, when available. In other words, on average, children's looking pattern did not differ from the adults'.

The absence of a group effect in the general analyses we conducted to answer RQ1 and RQ2 does not allow us to statistically investigate the effect of age within the group of children only. Nevertheless, we are allowed to explore the graphs of each age-group's looking pattern and to compare it to the adults'. We do so for one main reason, i.e., we believe data from our adult control group might have been biased by our experimental conditions. By the end of the experiment, most adult participants claimed to have been influenced by the trials in the informative-false condition (i.e., "Is *soft* the bone?") as they perceived the presence of a possible "deception", when the informative adjective was followed by, and referred to, the distractor (e.g., the bone) and not the target (e.g., the pillow). Thus, they admitted having consciously manipulated their looking behavior and, most of the times, waited for the noun while ignoring the adjective, even in conditions in which it was informative.

As for RQ1, although a task bias might have influenced our adult data, we did find a significant effect of condition between the informative and the uninformative adjective conditions, signaling that, when informative, the adjective was interpreted before the noun. However, no adult-advantage was found, which is surprising considering our results in Experiments 1 and 2. No group effect was found by Tribushinina and Mak (2016) either.

In the current experiment, children's age ranged from 28 months (2;4 years) to 63 months (5;3 years). To explore our research question on development, we created three age groups. The first group (2-year-olds) included children from 2;4 years to 3;1 years (N = 13; mean age = 2;9, S.D. = 0;2, F = 6). In the second group (3-year-olds), toddlers span between 3;2 years and 3;9 years (N = 14; mean age = 3;5, S.D. = 0;2, F = 8). The last group (4-year-olds) included children between 3;11 years and 5;2 years (N = 12; mean age = 4;5, S.D. = 0;6; F = 6).

Informative-true Informative-false

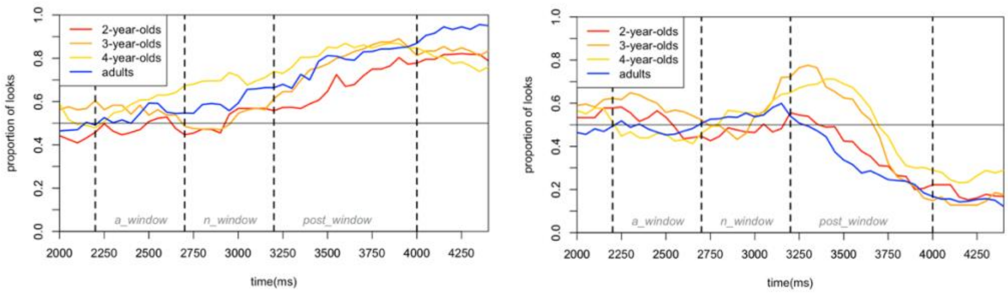


Figure 45. Proportion of fixations to the picture possessing the labeled property (e.g., the pillow) over time in the informative-true condition (on the left) and in the informative-false condition (on the right) for 2-year-olds (in red), 3-year-olds (in orange), 4-year-olds (in yellow) and adults (in blue). The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average onset of the noun, the third vertical line indicates the average offset of the sentence. The horizontal line indicates chance level.

A visual inspection of Figure 45 reveals the following. First, in the informative-true condition, we observe that, among children, 2- and 3-year-olds are the slowest groups in interpreting the stimulus question, as they start to increase their fixations to the target picture above chance around 2850ms, i.e., 150ms after the offset of the adjective. By contrast, 4-year-olds are faster, as they show an increase in target fixations already during the noun-window. As for accuracy, while 3- and 4-year-olds reach 0.9 proportions of looks at the target, 2-year-olds do not exceed 0.8. What is surprising, however, is the adults' looking pattern. Although the adjective was informative about the referent noun, adults significantly increase their fixations to the target only during the noun-window, concurrently with 2- and 3-year-olds (i.e., around 2900ms). Hence, while 4-year-olds resolve reference at the adjective, adults seem to be more cautious and wait for the noun validation.

A similar pattern emerges in the informative-false condition. Indeed, while 3- and 4-year-olds direct their gaze towards the picture possessing the property labeled by the adjective almost getting to 0.8 proportion of fixations, and shift to the labeled object only around 3750ms, adults, as well as 2-year-olds, do not seem to interpret the informative adjective, but rather they wait until the noun to focus on the referent picture. We believe Figure 46 shows the effect of the task bias we hypothesize.

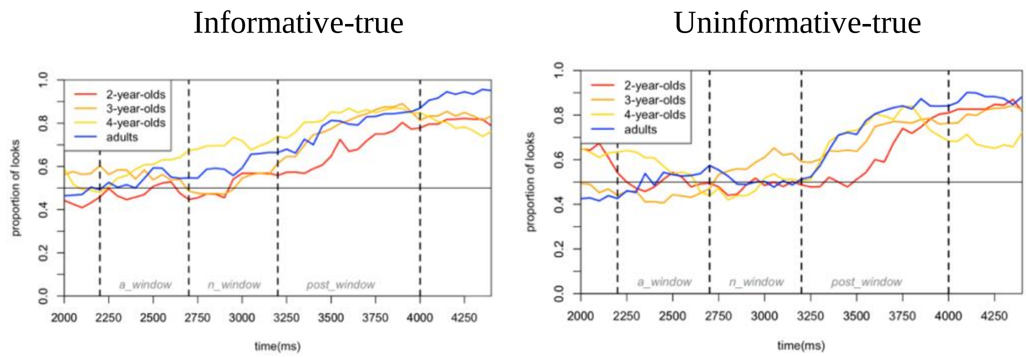


Figure 46. Proportion of fixations to the target over time in the informative-true condition (on the left) and in the uninformative-true condition (on the right) for 2-year-olds (in red), 3-year-olds (in orange), 4-year-olds (in yellow) and adults (in blue). The first vertical line indicates the average onset of the adjective, the second vertical line indicates the average onset of the noun, the third vertical line indicates the average offset of the sentence. The horizontal line indicates chance level.

The observation of the graph from the uninformative-true condition, however, confirms that participants are indeed engaged in the task. As soon as the information from the noun becomes available, all groups direct to their gaze to the target picture. Although 2-year-olds need more time to process the stimulus question, they almost reach the accuracy of older children and adults, corroborating their high level of proficiency when noun-computation only is required, in line with the findings from Experiment 2.

In conclusion, we believe a task bias indeed affected our adult data. Nevertheless, the observation of children’s looking behavior in Figure 46 suggests a plausible developmental pattern similar to what we found for Experiments 1 and 2, i.e., that 4-year-olds are faster and more accurate than 2-year-olds, and that 3-year-old children are in between the other two groups, showing emerging processing abilities.

6.5 General discussion

Experiment 3 set out to explore children’s processing of adjective-noun phrases and their ability to predict the noun based on the adjective meaning. With this aim, we designed a Visual World eye-tracking study, presenting participants with two pictures on screen, and recording eye-movements while hearing Italian predicative yes/no questions with adjective-noun combinations. We aimed to answer 3 research questions.

6.5.1 Predicting nouns based on adjective meaning

The first research question (RQ1) asks whether children between two and five years of age interpret adjective-noun phrases incrementally and predict the noun by interpreting the prenominal informative adjective. Our results demonstrate that this is indeed the case. In conditions in which the adjective was informative about the following noun (e.g., ‘Is soft the pillow?’), the proportion of looks to the target (e.g., the pillow) increased upon hearing the adjective and before the onset of the noun. In contrast, when the adjective was uninformative (e.g., ‘Is white the pillow?’ for the pillow–bone trial in which both the pillow and the bone were white), the proportion of looks to the target between adjective onset and noun onset did not differ from chance. Hence, we did not find evidence of a *Noun-Anchoring* mechanism in the interpretation of adjective-noun combinations in predicative questions. Moreover, contrary to our expectations, children did not need more time than the adults to shift their gaze towards the target picture. This finding contrasts with our results from Experiments 1 and 2, in which a significant group difference was found in a 4- as well as a 2-referent display, and independently of whether the interpretation of the post-nominal adjective was necessary or not. Hence, prenominal adjectives in the present study seem to facilitate sentence processing, allowing for a rapid and accurate incremental interpretation and semantic prediction based on adjective meaning alone, strengthening Tribushinina and Mak (2016)’s findings with silence in-between adjectives and nouns, but contrary to Davies et al. (2021)’s results from natural speed adjective-noun combination processing.

6.5.2 Interpreting prototypical and non-prototypical properties of objects

The second research question (RQ2) explored whether there is a difference in the time-course and in the accuracy of interpretation between prototypical and non-prototypical properties of objects. We manipulated the visual representation of target properties by making use of adjectives that were not visible on the screen (i.e., prototypical adjectives like *soft*) or visible on the screen (i.e., non-prototypical adjectives like *closed*). In Tribushinina and Mak's study only non-prototypical adjectives were tested, inhibiting the possibility to investigate what the source of children's successful prediction was. This experimental design aimed to explore whether the integration of adjective semantics with world knowledge in one case, and the visual context in the other, would result in a delayed target identification in one or the other situation. No effect of prototypicality emerged for children, nor adults, showing that once the adjective meaning is acquired children are not facilitated nor disadvantaged by the representation of the property within the visual context.

6.5.3 Children's and adults' processing of adjective-noun phrases

The third research question (RQ3) targeted the nature of a possible group difference between children and adults with respect to RQ1 and RQ2, by exploring 2-, 3- and 4-year-olds' looking behaviors in comparison. Surprisingly, no significant difference between children and the adult control group was found, neither concerning the interpretation of informative vs. non-informative prenominal adjectives (RQ1), nor in relation to the prototypicality of the properties expressed by the adjectives (RQ2). As already mentioned, this finding was unexpected if compared to what we found in the two previous experiments and raises a question about the nature of the contrasting results we got in Experiment 1 and 2 on the one hand, and Experiment 3 on the other.

One possibility is that our findings may provide experimental evidence of a facilitation in the processing of adjective-noun combinations (in Experiment 3) in comparison to noun-adjective phrases (in Experiments 1 and 2). However, since the experimental tasks in the three experiments were not identical, concluding that the adjective-noun order in Experiment 3 was the only source of facilitation in processing noun-adjective combinations would be misleading. Indeed, we believe there might be a pragmatic motivation for the contrasting outcomes from noun-adjective and adjective-noun processing, which has to do with the differences between the tasks. The design of Experiment 1 cannot be directly compared to that of Experiment 3, whereas the experimental design of Experiment 2 was very similar to the current study. While presenting two pictures on the screen, in Experiment 2 the task explicitly asked participants to find an object which was modified by a postnominal color-adjective. In 2 out of 3 experimental conditions the noun was informative about the target object (e.g., when asked 'Where is the dinosaur green?', in a dinosaur-hat visual context) and, even so, children were overall slower and less accurate than the adults. Thus, from the 2-year-old children's viewpoint - to whom the difference with the adult group is mostly attributable - the (redundant) information provided by the adjective might have been significant and, hence, prompt a set-computation that is costlier for them than for adults. In Experiment 3, by contrast, there was no specific task. Children were asked about the property of an object, and the interpretation of the adjective was necessary to answer the yes-no question. In contrast to Experiment 2, if the prenominal adjective was informative, the noun was nonetheless expected in the sentence for discourse felicity (e.g., asking "Is soft?" would be infelicitous in the context of the experiment). Hence, in Experiment 3, the task makes adjective interpretation both useful and necessary in every condition, favoring incremental processing and anticipatory fixations. This result does not completely undermine the claim that toddlers younger than 36 months process adjectives only after the referent noun, as hypothesized by the *Noun-Anchor Hypothesis* (Ninio, 2004). Rather, we believe the contrasting outcomes from our experiments result from different processes of interpretation, i.e., from developmentally limited set-operation abilities in Experiment 2, and from efficient predicting skills in Experiment 3.

Alternatively, Experiment 3 may have a limitation due to a task effect affecting the adults' data. By the end of the testing session, most adult participants, unaware of the purposes of the experiment, recognized to have manipulated their eye-movement behavior once they realized that in one experimental condition (i.e., the informative-false condition) the adjective was informative (e.g., *soft*), but referred to the distractor object (e.g., the bone). Thus, fearing to be deceived by the stimulus question, they tended to ignore the adjective and wait for the noun, which was uniquely referred to the "correct" object, negatively influencing the effect of the informativeness of the adjective. To explore this possibility, we evaluated the graphs in Figures 45 and 46 and we observed that 4- and often 3-year-olds outperformed adults in the informative conditions, as they rapidly focused their fixations towards the object possessing the labeled property. 2-year-olds, by contrast, were not able to make use of the information provided by the adjective and waited for the noun to shift to the labeled object.

We believe these two explanations do not entirely exclude each other. The analysis of children' data showed that children between two and five years of age do make use of the informative adjective to predict the upcoming noun. This signals that children understand adjectival meaning and can draw inferences, surfacing in an efficient incremental processing of adjective-noun combinations. In other words, children might be better at predicting than at computing set-operations, and if the adjective alone is sufficient for task resolution, they base on that, associating e.g., *soft* to a pillow, and they stick to it, probably avoiding the unnecessary set-operation. A difference between the interpretation of informative vs. non-informative adjectives was also found in the adult group indicating that, despite the conscious manipulation of the looking behavior adults acknowledged, they unconsciously reacted to the informative adjective. Nonetheless, it is plausible that the adults' cautious looking behavior indeed influenced the outcome of the experiment, resulting in the absence of a group effect, contrary to our predictions and to our results in Experiments 1 and 2.

6.5.4 Conclusion

Although some limitations have been detected, this study has shown that children between two and five years of age process yes/no questions with prenominal adjectives incrementally. They do so by interpreting adjectives as they hear them and by predicting the following noun based on the adjective lexical meaning. Furthermore, our findings revealed that once the adjective is acquired children successfully process it, independently of whether, for its interpretation, they need to rely on world knowledge or explore the visual scene. By means of eye-tracking, we were able to reveal that children as young as 2;4 years are successful parsers and that once a novel adjective is acquired they improve their processing skills becoming adult-like within a few months and very early in development.

7 Conclusions

This thesis investigated the acquisition and processing of adjectives in Italian-speaking children between the age of two and five years. Although language processing is a widely investigated phenomenon in the acquisition literature, psycholinguistic research has rarely focused on the adjectival class. Furthermore, the scarce literature on the topic has provided contrasting evidence on the timing and mode of adjectival processing in children younger than the age of four.

The core aspects under discussion in the present work included the emergence of the ability to integrate noun and adjective meanings, the real-time interpretation of different semantic classes of adjectives and the role of predictive processing in the computation of adjective-noun occurrences. The present thesis was built upon previous studies in the literature, trying to identify and overcome their limitations. Specifically, we aimed to address adjective acquisition and processing from different points of view, bringing improvements in the methodology, answering new research questions, and testing a wider age-range of participants.

To conclude this thesis, chapter 7 recapitulates the theoretical background that motivated Experiment 1 (§7.1), Experiment 2 (§7.2), and Experiment 3 (§7.3), summarizing the main findings and reporting the answer to each research question. §7.4 discusses the theoretical implications of the main findings. §7.5 examines the limitations and future directions of the present work. To ease reading, in this chapter bibliographical references have been avoided and replaced by the related chapters' sections.

7.1 Experiment 1

Experiment 1 (Ch. 4) was conceived in response to the observation that the interpretation of nouns and adjectives in combination has rarely been investigated using refined online technologies and never in languages with noun-adjective order, such as Italian. Furthermore, the interpretation of subsective adjectives has only been tested offline and never using eye-tracking with children, much less in comparison to adults.

In this dissertation we have seen that most of the previous studies on adjective acquisition have found evidence of a *Noun-Anchoring* mechanism in the interpretation of noun-adjective combinations by children younger than 36 months of age. Specifically, young children have been found to show difficulties or failure in integrating the adjective with the referent noun, both offline and online. This was taken as evidence that nouns are easier to process, and that their interpretation is always privileged, while adjectives are either ignored or processed more slowly after the noun. This has been observed in languages with postnominal adjectives (e.g., Hebrew) as well as with prenominal adjectives (e.g., English). Fewer studies found that nouns and adjectives are processed equally quicky and that, when adjectives are prenominal, children successfully interpret the combination incrementally. However, we have seen that these studies were limited in various ways and, above all, did not test children younger than the age of three.

As we have widely discussed, the few studies investigating the acquisition and interpretation of adjective semantics reported that relative and absolute adjectives are interpreted in different ways, in that the cut-off point is identified in the middle of the scale for relative adjectives and on one boundary for absolute adjectives. These findings led to the observation that absolute adjectives share properties with both relative and intersective adjectives. Similar to relative adjectives, absolute adjectives are context sensitive (though to different degrees); like for intersective adjectives, objects described by absolute adjectives are always judged as either possessing the property or not, and are never vague. Although interesting, these results came from offline Scalar Judgment Tasks administered to children older than three years, or online eye-tracking studies testing adults.

Based on these considerations, we designed a Visual World eye-tracking study and we tested Italian-learning children between 2;4 and 5;3 years in comparison to adult controls. We recorded fixations to four pictures on the screen of two objects crossed by two properties (e.g., a black shoe, a white shoe, a black sock, and a white sock), while each stimulus sentence unfolded (i.e., “Where is the Noun-Adj?”, e.g., *Dov’è la scarpa nera?* “Where is the black shoe?”). Adjectives were either intersective (e.g., *black*), relative (e.g., *big*) or absolute (e.g., *closed*). We addressed the following research questions:

RQ1. Is there evidence of a *Noun-Anchoring* mechanism in children's online processing of Italian noun-adjective combinations? Do children focus on the interpretation of the noun-word while omitting or delaying the integration of the adjective meaning? Do children and adults differ in processing noun-adjective combinations?

AW1. Our findings indicate that children between two and five years of age are able to successfully integrate noun and adjective meanings to identify the referent object. In our results we do not find evidence of a *Noun-Anchoring* mechanism, as children do not omit adjective interpretation, nor they delay it in comparison to the noun. However, when confronting their processing pattern with the adults', children's processing limitations emerge, as they are found to be overall slower and less accurate. While children focus on the labeled objects at the noun-offset and shift to the target at the end of the referring expression, adults interpret each word during the time-window in which it is uttered, resulting in a faster and more efficient reference resolution. The nature of the different outcome from the two age populations is addressed by the next two research questions (§4.5.1).

RQ2. Are there differences in the way children interpret different semantic classes of adjectives in real-time? Are children sensitive to contextual cues affecting the interpretation of each adjective class? Does adults' processing vary across the different classes of adjectives?

AW2. Results revealed that children successfully resolved reference in all adjective-conditions, but that differences between adjective classes emerged. In line with previous offline studies and with the semantic theories on subsective adjectives, we found that intersective adjectives (e.g., *black*) are integrated faster than relative adjectives (e.g., *big*). This finding suggests that, once the noun is processed and the labeled objects are identified, in the interpretation of relative adjectives children look back and forth the target and the noun-competitor, sensibly checking their choice against the contrast object. Upon hearing e.g., *the big teddy*, children first identify and focus on the two teddies on the screen, and shift between the two in interpreting the adjective *big*. This second step is not required for the interpretation of intersective adjectives, as the identification of the shoe counting as *black* is not result of a comparison with the show that is white. Consequently, the interpretation of relative adjectives requires extra processing time and results in a delayed reference resolution. By contrast, the process of interpretation of absolute adjectives (e.g., *closed*) did not differ from neither of the two other adjective-conditions, as the time required to identify the target was in-between intersective and relative adjectives. This result matches the theoretical observation that absolute adjectives share properties with both intersective and relative adjectives and is in line with previous results from offline studies, indicating that children as young as 28 months are sensitive to the semantic properties of adjectives.

The results from the adults' group revealed that they were fast and accurate in all adjective-conditions, but that the interpretation of absolute adjectives was slower in comparison to intersective and relative adjectives. We proposed two possible explanations for this finding. First, this pattern could result from differences within the group of absolute adjectives, that either had two-closed-boundaries (e.g., *open/closed*) or just one-closed-boundary (e.g., *clean/dirty*). Second, the delayed interpretation could be attributable to the visual stimuli in this adjective-conditions, that required a multi-dimensional manipulation of the pictures on the screen (§4.5.2).

RQ3. With respect to the first two research questions, do children show developmental differences in processing noun-adjective combinations? Do 2-, 3- and 4-year-old children differ in processing different classes of adjectives?

AW3. Our findings constitute evidence that children younger than age 3 have the most difficulties in interpreting noun-adjective combinations in comparison to 3- and 4-year-olds, attributable to their less-developed cognitive and linguistic abilities. Interestingly, however, 2-year-olds' looking pattern in the interpretation process was essentially the same as older children and adults. This finding contrasts with previous studies testing children younger than 3, who delayed or failed adjective processing but successfully interpreted the noun regardless of the relative order of the words in the combination. A closer look at the interpretation of the three adjective-conditions confirmed the pattern we found for the whole group of children. The difficulties we observed with relative adjectives are shared among the three age-groups, while the intersective adjectives were faster and earliest integrated. This finding corroborates the claim that children's knowledge of semantic differences between subsective adjectives emerges very early in development (§4.5.3).

In conclusion, our results from Experiment 1 provided new compelling evidence covering the integration of nouns and adjectives and the awareness of the contextual cues affecting the interpretation of intersective, relative and absolute adjectives. Unlike previous research in the field, through the eye-tracking technology we found evidence of developing adult-like processing skills. Indeed, for children as young as 28 months no *Noun-Anchoring* mechanism was observed in the interpretation and integration of noun and adjective meanings. This result contributes to the literature by showing that, once adjective meaning is acquired, children are able to integrate it with the noun, though at a very slow speed which gets faster with age.

7.2 Experiment 2

Experiment 2 (Ch. 5) was designed building upon the results from Experiment 1, where we ascertained that children can successfully integrate noun and adjective information and that children's performance gets faster and more efficient with age. Nevertheless, no evidence of the *Noun-Anchor Hypothesis* was found. One possible limitation of Experiment 1 in this respect is that both noun and adjective always needed to be interpreted and integrated to resolve reference. Hence, the experimental design of Experiment 1 did not allow for a comparison between conditions in which adjective interpretation is either required or not. This was central in Experiment 2. We reduced the experimental complexity of Experiment 1 by presenting children a two-picture visual scenario and we asked questions containing nouns combined with color (intersective) adjectives only, which in the previous experiment resulted to be easier and faster integrated with nouns.

Although most previous studies found evidence of a *Noun-Anchoring* mechanism in interpreting noun-adjective and adjective-noun combinations, an alternative hypothesis was proposed, i.e., the *Sequential Integration Hypothesis*. According to this view, children have difficulty with the integration of any two content words, regardless of their relative order. Consequently, the integration is easier when the informative word comes last, as this arrangement is least demanding on working memory. What this proposal suggests is that a possible asymmetry between the interpretation of nouns and adjectives does not come from a priority of nouns over adjectives in acquisition and processing patterns. Rather, when an informative word comes first, the parser's attention might be disrupted by an unnecessary additional information, which negatively affects the computation of the combination. No evidence in support of this hypothesis was found in Looking-While-Listening tasks with prenominal nor postnominal adjectives. However, this issue has never been investigated with the eye-tracking, nor with children as young as 28 months in a language with postnominal adjectives like Italian.

Based on our findings in Experiment 1, and on the experimental literature discussed, we developed a Visual World eye-tracking study. We tested Italian-learning children (2;4 – 5;2) and adults and we recorded fixations to two objects on the screen, while listening to questions about one of them (e.g., *Dov'è il dinosauro verde?*, "Where is the green dinosaur?"). The two pictures on the screen were manipulated so that the noun and the adjective varied in their informativeness with respect to the visual scene. Our research questions were the following:

RQ1. Do children show a *sequential-integration* pattern in interpreting color-adjectives combined with informative and non-informative nouns? Are they disrupted by uninformative postnominal adjectives?

AW1. Our results contrast with the *Sequential Integration Hypothesis*. In the discussion we have seen that, in conditions in which the noun was informative, children identified the referent early and did not appear to suffer any disruption from the adjective, regardless of whether it was informative or uninformative. However, in the visual condition in which the adjective was informative – and its interpretation was required to resolve reference – a computational asymmetry was found at the expense of the adjective. In other words, when adjective interpretation is essential to complete the task, children are less accurate than in the conditions in which it is not, showing processing difficulties specifically linked to adjective interpretation. This finding contrasts with previous results with English and Spanish children between 3;0 and 3;6 years but matches with results from younger children (§5.5.1).

RQ2. Do children and adults differ in the pattern of interpretation with respect to speed and accuracy?

AW2. We found that children and adults differ both in speed and accuracy of interpretation. While adults rapidly interpret each informative word as soon as it is heard, children's performance is slower and less accurate, especially when the post-nominal adjective is the only informative word of the combination (§5.5.2).

RQ3. Are there differences between 2-, 3- and 4-year-old children in processing noun-adjective combinations across different informativeness conditions?

AW3. From our results age-differences emerged in all conditions. In the two conditions in which the noun was informative, we found that 4-year-olds outperformed 2-year-olds, while 3-year-olds' performance was in between, signaling a developmental pattern of interpretation. Nevertheless, all groups successfully resolved reference. By contrast, when the integration with the adjective was fundamental for target identification, while 4-year-olds were as accurate as in the other two conditions, 2-year-olds failed the task and kept looking at both pictures around chance, indicating that they could not integrate the adjective meaning. 3-year-olds' performance was in between the other two groups, corroborating the observation from Experiment 1, i.e., that they are improving their processing skills but are still not as efficient as older children (§5.5.3).

To conclude, results in Experiment 2 are consistent with our previous findings in Experiment 1 and show a processing pattern rapidly developing from the second to the fifth year of age. A direct confrontation between conditions in which adjective interpretation was either required or not allowed us to find evidence of the *Noun-Anchor Hypothesis*, as children younger than the age of 3 failed to integrate adjective meaning even with familiar intersective color adjective.

7.3 Experiment 3

Experiment 3 (Ch. 6) expanded the investigation on children's processing of nouns and adjectives by examining eye-movements during the online interpretation of prenominal adjectives. We targeted children's ability to process adjectives as they are being heard and to predict the following noun based on the lexical meaning of the adjective. While Experiment 1 and 2 provided contrasting results with respect to the integration of nouns and adjectives by the youngest children, in Experiment 3 we took on a new approach by reversing the order of nouns and adjectives. What the *Noun-Anchor Hypothesis* predicts in this respect is that children younger than the age of 3 would initially listen through the adjective and start the interpretation of the combination only at the noun. Previous studies addressing this issue with the Visual World paradigm found that 3-year-olds are able to make use of the prenominal adjective to predict the following noun, but only when language speed is manipulated to allow more time for processing and only with adjectives expressing prototypical properties of the noun (e.g., *heavy* for a stone).

To test the computation of adjective-noun occurrences in Italian, we showed participants two pictures on the screen (e.g., a white pillow and a white bone) and we asked predicative yes-no questions containing adjectives that were either informative (e.g., *soft*) or uninfomative (e.g., *white*) about the noun (e.g., *È morbido il cuscino?* lit., 'Is soft the pillow?', "Is the pillow soft?"). Moreover, adjectives were either prototypical (e.g., *soft*) or non-prototypical (e.g., *open*) properties of the objects. 39 Italian-speaking children (2;4 – 5;3) participated in the experiment. The following research questions have been addressed:

RQ1. Are children able to predict the noun based on the pre-nominal adjective in Italian predicative constructions? Do they show a *Noun-Anchoring* mechanism in the interpretation of prenominal adjectives? Does their looking pattern differ from the adults'?

AW1. Our results showed that children between two and five years of age successfully predict the noun based on the lexical meaning of the prenominal adjective. In conditions in which the adjective is informative about the noun (e.g., *soft* for a pillow), children shift their gaze towards the pillow before hearing the noun. Interestingly, no difference between children and adults is found in the interpretation of adjective-noun combinations. This contrasts with our findings in Experiments 1 and 2, where children were always overall slower and less accurate than adults, even in conditions in which the noun was already informative about the target object and the interpretation of the adjective was not necessary (§6.5.1).

RQ2. Are there differences between the interpretation of adjectives expressing an inherent property of the object and do not have a visual representation (e.g., *soft*, for a pillow), and adjectives that are visible on the screen (e.g., *open*, for a window)? Do adults show processing differences in this respect?

AW2. Our findings showed that no difference emerges in the interpretation of prototypical and non-prototypical adjectives and, thus, that children integrate adjective semantics with world knowledge and the visual context equally quickly and accurately. The same was found for adults. This result indicates that, once the meaning of an adjective is acquired, the visual representation of a property does not enhance nor inhibit children's successful processing (§6.5.2).

RQ3. If children and adults differ with respect to the first two research questions, do 2-, 3- and 4-year-old children show developmental differences in processing informative vs. uninformative adjectives? And in the interpretation of inherent vs. non-inherent properties of objects?

AW3. Both in relation to the interpretation of informative vs. non-informative prenominal adjectives (RQ1), and to the prototypicality of the properties expressed by the adjectives (RQ2), no difference between children and adults has emerged. This raised a question about the source of this unexpected result in comparison to our findings in Experiment 1 and 2. Two possible explanations have been proposed. First, we suggested that the adult-like performance children show in Experiment 3 cannot be directly attributable to the adjective-noun order, as the task was not comparable to those administered in Experiments 1 and 2, mainly for pragmatic reasons. Take Experiment 2, where two pictures were showed while a noun-adjective combination described the object to be found. Here, the task explicitly required participants to look for an object, and, in 2 out 3 conditions, the noun was informative and sufficient for task resolution (e.g., while hearing 'Where is the dinosaur green?', in a dinosaur-hat visual context). Even in this simple context, children were always found to be slower and less accurate than adults, mostly due to the poor performance of 2-year-olds. This finding might be the result of the expectation children have about the post-nominal adjective which, if added to the already-informative noun, triggers an unnecessary set-computation which is difficult and takes time. Experiment 3, by contrast, did not have an explicit task, but required participants to answer a yes/no question about one of the two objects on the screen. For the question to be answered, in all conditions the adjective needed to be processed. Moreover, for pragmatic felicity, the noun was always expected, regardless of the informativity of the prenominal adjective. Hence, children might be better and more efficient at predicting the upcoming noun – as required in Experiment 3 – than in set-operations – triggered by Experiment 2 (and Experiment 1). A second explanation was formulated based on the looking behavior of the adult participants, who we believe might have been biased by the task. We observed that adults were cautious in shifting their gaze towards the target object after hearing an informative noun, fearing to be deceived by the conditions in which the adjective was informative (e.g., *soft*) but the noun labeled the distractor (e.g., the bone). We concluded that the two possibilities do not exclude each other and that our findings might be a combination of (i) efficient anticipatory abilities that children develop very early in development and (ii) some adults' prudent looking behavior (§6.5.3).

In conclusion, our results in Experiment 3 do not contrast with previous findings in Experiment 1 and 2, but rather complete the picture and provide new insights on children's powerful ability to exploit all possible cues (visual and conceptual) and to process language incrementally even before the age of 3.

7.4 Theoretical implications

This thesis has largely focused on the broad theoretical debate on the acquisition and processing of noun-adjective combinations. By adopting a comprehensive approach, we answered new research questions and we tested a wide age-range of participants, investigating adjective processing from multiple perspectives. Throughout the discussion, we have seen that the findings reported by offline experimental paradigms investigating the acquisition of noun-adjective combinations (e.g., picture-pointing tasks) and unsophisticated online techniques (e.g., the looking-while-listening paradigm) have provided unconclusive evidence in this respect. Moreover, previous studies exploiting our same technology (i.e., eye-tracking) were limited in their research questions and in the age-range of participants.

Our three experiments aimed to test the two hypotheses that have been proposed to account for the difficulties in young children's interpretation of nouns and adjectives in combination, i.e., the *Noun-Anchor Hypothesis* (Ninio, 2004) and the *Sequential Integration Hypothesis* (Weisleder & Fernald, 2009). According to the first, children under the age of 3 interpret noun-adjective combinations in two-steps, by first processing the noun, regardless of whether the language under discussion has prenominal or postnominal adjectives. Adjectives are always integrated later or even ignored. The second hypothesis claims that word combinations are always hard to interpret, and that, to avoid a processing overload and working memory costs, the interpretation is easier when the informative word comes last, as no information needs to be kept in memory and retrieved at a later point. To verify these hypotheses, we constructed three eye-tracking experiments in which we manipulated the types of adjectives used, the visual representation of the properties, and the relative order of nouns and adjectives.

In Experiment 1, for the task to be resolved, participants were required to always integrate the meaning of nouns and adjectives, as neither of the two was sufficient to identify the described object. Our findings provide evidence of 2-year-old's efficient and adult-like pattern of interpretation of such combinations. The differences that emerged between children and adults – and between 2-year-olds and older children – concern the speed and the accuracy of interpretation, but never the process of interpretation. We believe this might be attributable to different factors. First, children might have a slower and time-consuming access to the lexicon in comparison to experienced adults, which could be further slowed down by a slower phonological activation. This is supported by the fact that adults were able to interpret nouns and adjectives upon hearing the first syllable, while children needed to listen to the entire word before shifting their gaze accordingly. Second, the difficulties children show might have to do with the complexity of the experimental design. The time-window during which a preview of the images was displayed on the screen might have been too short for children who needed more time to process the visual stimuli, delaying the beginning of the computation of the linguistic information. We believe that these two explanations do not exclude each other. Interestingly, however, children's still-developing cognitive abilities did not result in an asymmetry between the processing of nouns and adjectives.

Moreover, Experiment 1 was built to investigate the online processing of subsective adjectives by comparing intersective, relative and absolute adjectives. This allowed us to demonstrate that the different set-operations required for each of these classes played a role in the process of interpretation, with intersective adjectives being the easiest and faster and relative adjectives being the most demanding and time-consuming. At the same time, we provided evidence that children as young as two years of age are aware of the semantic differences among subsective adjectives.

In Experiment 2, we only tested intersective color adjectives and we manipulated the informativeness of nouns and adjectives with respect to the visual scenario. Participants were asked to look for an object which, in 2 out of 3 conditions, was identifiable at the noun, while the interpretation of the adjective was required in only one condition. Our results showed that, contrary to the predictions made by the *Sequential Integration Hypothesis*, children did not suffer from any disruption due to the additional uninformative post-nominal adjective. Rather, in the two conditions in which noun interpretation was enough, the oldest children's performance was almost target-like. By contrast, when the post-nominal adjective was informative, children's process of interpretation was overall negatively affected, and 2-year-olds failed reference resolution. Contrary to Experiment 1, in Experiment 2 we found evidence of a *Noun-Anchoring* mechanism in the interpretation of noun-adjective combinations.

The differences between the tasks in Experiments 1 and 2 clearly play a role in the contrasting outcomes from the two experiments. All adjective-conditions in Experiment 1 required participants to process both nouns and adjectives to identify the target object and, even so, children did not show an asymmetry between noun and adjective processing. By contrast, in Experiment 2, although the adjectives used were found to be the easiest to interpret in Experiment 1 (i.e., intersective), the youngest children failed the identification of the target when the integration with the color-adjective was necessary. Since Experiment 1 evidenced that children as young as 28 months do not lack the linguistic and cognitive resources to compute set operations, we cannot assume that this is the case in Experiment 2. Rather, we argue that the computation of the whole combination in Experiment 2 was statistically less likely to be required and hence, the youngest children tended to avoid engaging in costly set operations, even when necessary to resolve reference. A further possible explanation for the poor performance of the 2-year-olds in this experiment could be due to the fact that, during the testing phase, Experiment 2 was always administered last. Hence, the total duration of the testing session (around 40 minutes) and, specifically, the duration of Experiment 2 (i.e., 24 trials) with respect to the other two (i.e., 12 trials in Experiment 1, 16 in Experiment 3), might have been too long and tiring for the youngest children, whose engagement in the task was negatively affected. In conclusion, we believe that integrating nouns and adjectives is not beyond the reach of this age group and their looking pattern does not mirror a *Noun-Anchoring* mechanism per se, but rather a shortcut children take to avoid cognitively costly operations, within an experimental design which arguably prompted this behavior in the youngest children.

Experiment 3 provided further evidence of efficient processing resources children develop very early in development, as well as their ability to predict the noun based on the lexical meaning of the prenominal adjective. Here, the task required participants to interpret both words of the combination to answer the stimulus question. We found that children were successful and, surprisingly, adult-like. We have discussed that a possible task bias might have negatively influenced adults' data but that nonetheless children's ability to predict a noun is well developed. Contrary to the predictions from the *Noun-Anchor Hypothesis*, children as young as 28 months do not process such combinations in a noun-first/adjective-later fashion. Rather, they successfully interpret the combination incrementally. This finding demonstrates that when the linguistic context makes adjective interpretation useful, children as young as 28 months interpret it promptly and correctly.

Once again, Experiments 2 and 3 are not directly comparable as the tasks they exploited are different and required different processes of interpretation. Nonetheless, Experiment 3 evidenced that the participants' interpretation of pre-nominal adjectives can benefit from multiple cues, i.e., the visual representation of the property and the conceptual knowledge of familiar objects, which help inferring the referred object. This effect was likely further intensified by the linguistic stimulus, which raised a pragmatic expectation about the following noun (e.g., asking 'Is *soft*?' would have been pragmatically infelicitous). In contrast, the task in Experiment 2 required a set computation that was not necessary most of the times, as the adjective was over-informative (e.g., when asking for the dinosaur *green* in a dinosaur-hat context). This has probably had the opposite effect of Experiment 3. Specifically, the pragmatically infelicitous presence of the postnominal adjective might have refrained the youngest children from performing cognitive operations that recruit sophisticated abilities. In conclusion, our findings suggest that children as young as 28 months possess efficient prediction and set-operation skills. Before the age of three, however, while predicting abilities are less demanding, the computation of set operations is cognitively costlier and, when possible, avoided.

7.5 Limitations and future research directions

The main limitation of the present work is represented by the language under investigation. The initial goal of the present thesis – and of the cotutelle agreement with the University of Konstanz – was to test our research questions in both Italian and German. The two languages were chosen for the specular position of adjectives in attributive and predicative use, as well as their morphological similarities in such constructions. Specifically, Experiment 2 was constructed to test Italian and German monolingual children, as well as German-Italian bilinguals. This would have allowed us to observe possible differences in strategies and patterns of interpretation by the application of experimental paradigms that were identical except for the relative order of nouns and adjectives. This was hampered by the Covid-19 outbreak, which made the recruitment of Italian participants particularly challenging, and the data collection with German participants technically impracticable.

A further limitation of this work concerns the possible task bias we identified and discussed in Experiment 3. We have seen that our experimental conditions brought adult participants to the conscious manipulation of their looking behavior, affecting the outcome of the experiment. It would be interesting to run this experiment by removing the tricky conditions. A second possibility would be to modify the linguistic stimulus with an exclamative intonation (e.g., *È morbido il cuscino!*, "The pillow is soft!"), to avoid asking questions which can bias participants towards a forced shift to the referent object.

In general, more comprehensive insights could come from adapting our experimental design to German and to other languages, allowing for a cross-linguistic comparison with respect to the role of adjective semantics, incremental processing and predicting abilities. Relatedly, a further future research development might be to extend the investigation to early bilingual and/or language impaired populations, allowing for the validation of our results to adjective acquisition and processing across languages and populations.

References

Abney, S. P. (1987). The English noun phrase in its sentential aspect. (Doctoral dissertation, Massachusetts Institute of Technology).

Aikhenvald, A. Y. (2004). *Explorations in Linguistic Typology*. R. M. W. Dixon (Ed.). Oxford University Press.

Alexiadou, A. (2001). Adjective syntax and noun raising: Word order asymmetries in the DP as the result of adjective distribution. *Studia linguistica*, *55*(3), 217-248.

Alexiadou, A. (2004). On the development of possessive determiners. *Diachronic clues to synchronic grammar, 72*, 31.

Alexiadou, A., Haegeman L. & Stavrou M. (2007). *Noun Phrase in the Generative Perspective*. Berlin/New York, Mouton de Gruyter.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of memory and language*, *38*(4), 419-439.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition, 73*, 247-64.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*(3), 191-238.

Altmann, G., Garnham, A., Henstra, J. (1994). Effects of syntax in human sentence parsing: evidence against a structure-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory and Cognition 20*, 209–216.

Ambridge, B., & Rowland, C. F. (2013). Experimental methods in studying child language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(2), 149-168.

Aparicio, H., Xiang, M., & Kennedy, C. (2016). Processing gradable adjectives in context: A visual world study. In *Semantics and Linguistic Theory* (Vol. 25, pp. 413-432).

Aparicio, H., Kennedy, C., & Xiang, M. (2018). Perceived Informativity and Referential Effects of Contrast in Adjectivally Modified NPs. In E. Castroviejo, L. McNally, & G. Sassoon (Eds.), *The Semantics of Gradability, Vagueness, and Scale Structure. Experimental perspectives* (p. 199-220). Berlin: Springer.

Arias-Trejo, N., & Plunkett, K. (2010). The effects of perceptual similarity and category membership on early word-referent identification. *Journal of Experimental Child Psychology*, *105*(1-2), 63-80.

Baayen R.H., van Rij J., de Cat C., Wood S. (2018) Autocorrelated Errors in Experimental Data in the Language Sciences: Some Solutions Offered by Generalized Additive Mixed Models. In: Speelman D., Heylen K., Geeraerts D. (eds) Mixed-Effects Regression Models in Linguistics. *Quantitative Methods in the Humanities and Social Sciences*. Springer, Cham. https://doi.org/10.1007/978-3-319-69830-4_4

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Baayen, H., Vasishth, S., Kliegl, R., & Bates, D. (2017). The cave of shadows: Addressing the human factor with generalized additive mixed models. *Journal of Memory and Language, 94,* 206-234.

Bach, E.W. (1974). *Syntactic Theory*. New York, Holt, Rinehart and Winston.

Backscheider, A. G., Shatz, M., & Gelman, S. A. (1993). Preschoolers' ability to distinguish living kinds as a function of regrowth. *Child Development*, *64*(4), 1242-1257.

Baker, M. (2003). *Lexical categories: verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.

Bally, C. (1944). *Linguistique générale et linguistique française*. Berne, Francke.

Barr, D. J. (2008). Analyzing 'visual world'eyetracking data using multilevel logistic regression. *Journal of memory and language*, *59*(4), 457-474.

Bartlett, E. (1976). Sizing things up: the acquisition of the meaning of dimensional adjectives. *Journal of Child Language*, *3*(2), 205-219.

Bartning, I. (1976). Remarques sur la syntaxe et la sémantique des pseudo-adjectifs dénominaux en français. (Doctoral dissertation, Stockholms Universitet).

Bartsch, R., & Vennemann, T. (1972). Relative adjectives and comparison. *UCLA papers in syntax*, *2*, 80-107.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv. org:1506.04967.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* *67*(1). 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Bates, E. (1993). Comprehension and production in early language development: Comments on Savage-Rumbaugh et al. *Monographs of the Society for Research in Child Development*, *58*(3-4), 222-242.

Beck, D. (2002). *The typology of parts of speech systems: The markedness of adjectives*. Psychology Press.

Becker, M. (2000). The Development of the Copula in Child English: The Lightness of 'Be'. (Doctoral dissertation, University of Pennsylvania).

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language*, *6*, 183–200.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253-3258.

Berman, A. (1974). Adjectives and Adjective Complement Constructions in English. (Doctoral dissertation, Harvard University).

Bernstein, J.B. (1992). On the Syntactic Status of Adjectives in Romance. *CUNYForum*, *17*, 105-122.

Bever, T. G. (1970). Cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley.

Bhat, D.N.S. (1994). The adjectival category, criteria for differentiation and identification. *Studies in Language companion, 24*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bierwisch, M. (1987). Dimensionsadjektive als strukturierender Ausschnitt des Sprachverhaltens. In M. Bierwisch & E. Lang (Eds.), *Grammatische und konzeptuelle Aspekte von Dimensionsadjektiven* (p. 91-286). Berlin: Akademie-Verlag.

Bierwisch, M. (1989). The semantics of gradation. In M. Bierwisch & E. Lang (Eds.), *Dimensional adjectives: Grammatical structure and conceptual interpretation* (p. 71- 261). Berlin: Springer.

Bittner, D. (1998). Entfaltung grammatischer Relationen im NP-Erwerb: Referenz. *Folia Linguistica*, XXXI, 255-283.

Blackwell, A. (2001). On the Acquisition of the Syntax of English Adjectives. In A. Okrent & J. Boyle (Eds.), CLS 36, Volume 2: *The Panels*. Chicago: University of Chicago Press.

Blackwell, A. (2005). Acquiring the English adjective lexicon: relationships with input properties and adjectival semantic typology. *Journal of Child Language*, *32*(3), 535-562.

Bloom, L. (1973). *One word at a time: The use of single-word utterances before syntax*. The Hague: Mouton.

Boersma, Paul & David Weenink (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.19, retrieved 13 June 2016 from http://www.praat.org/

Bolinger, D. (1967). Adjectives in English: attribution and predication. *Lingua*, *18*, 1-34.

Bornstein, M. H. (1985). On the development of color naming in young children: Data and theory. *Brain and language*, *26*(1), 72-93.

Bornstein, M. H., Kessen, W., & Weiskopf, S. (1976). Color vision and hue categorization in young human
infants. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(1), 115.

Bosque, I. (1993). Sobre las diferencias entre los adjetivos relacionales y los calificativos. In *Revista Argentina de Lingüística, 9* (1/2), 9-48.

Bosque, I., & Picallo, C. (1996). Postnominal adjectives in Spanish DPs. *Journal of Linguistics*, 349-385.

Bouchard, D. (1998). The distribution and interpretation of adjectives in French: A consequence of Bare Phrase Structure. *Probu,s 10* (2), 139-183.

Bouchard, D. (2002). *Adjectives, Number and Interfaces: Why Languages Vary*. Amsterdam, Elsevier.

Bowers, J. (1993). The syntax of predication. *Linguistic Inquiry*, *24*(4), 591-656.

Braine, M. D. S. (1963). The ontogeny of English phrase structure: The first phase. *Language*, *39*, 1–14.

Britt, M. A. (1994). The interaction of referential ambiguity and argument structure in the parsing of prepositional phrases. *Journal of Memory and Language*, *33*(2), 251-283.

Brown-Schmidt, S., Campana, E., & Tanenhaus, M. K. (2002). Reference resolution in the wild: On-line circumscription of referential domains in a natural, interactive problem-solving task. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 24, No. 24).

Brown, R. (1958). How shall a thing be called?. *Psychological review*, *65*(1), 14.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Bruner, J. S. (1978). Berlyne Memorial Lecture: Acquiring the uses of language. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *32*(4), 204.

Bruner, J. S., Olver, R. R., & Greenfield, P. M. (1966). Studies in cognitive growth.

Buse, J. E. (1965). MAK Halliday and others: The linguistic sciences and language teaching. Mak Halliday, Angus Mclntosh, Peter Strevens.(Longmans' Linguistics Library.) XIX, 322 pp. London: Longmans, Green and Co. Ltd., 1964. 35s. *Bulletin of the School of Oriental and African Studies*, *28*(3), 675-676.

Cabredo Hofherr, P. & Mathushansky, O. (éds.). (2010). *Adjectives. Formal Analyses in Syntax and Semantics*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Cardinaletti, A., & Giusti, G. (2011). The acquisition of adjectival ordering in Italian. In *Variation in the input* (pp. 65-93). Springer, Dordrecht.

Cardinaletti, A., & Guasti, M.T. (1995). Small clauses: some controversies and issues of acquisition. *Small Clauses* (pp. 1-23). Brill.

Carey, S. (1982). Semantic development: The state of art. *Language acquisition: The state of art*.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word.

Carlson, G. N. (1977). *Reference to kinds in English*. University of Massachusetts Amherst.

Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, *10*(2), 159-199.

Chomsky, N. (1957). *Syntactic Structures*. The Hague/Paris, Mouton.

Chomsky, N. (1965). Persistent topics in linguistic theory. *Diogenes, 13* (51), 13-20.

Christiansen, M. H., Allen, J., & Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes*, *13*(2-3), 221-268.

Cinque, G. (1994). On the evidence for partial N-movement in the Romance DP. In *Paths towards universal grammar. Studies in honor of Richard S. Kayne*. Georgetown University Press.

Cinque, G. (2004b). Issues in adverbial syntax. *Lingua 114* (6), 683-710.

Cinque, G. (2005b). *The dual source of adjectives and phrasal movement in the Romance DP*. (Ms., Università degli Studi Ca' Foscari di Venezia.)

Cinque, G. (2007b), *The Syntax of Adjectives. A comparative study*. (Ms., Università degli Studi Ca' Foscari di Venezia).

Cinque, G. (2010). *The syntax of adjectives: A comparative study* (Vol. 57). MIT press.

Clahsen, H., Eisenbeiss, S., & Vainikka, A. (1994). The seeds of structure. A syntactic analysis of the acquisition of case marking. In T. Hoekstra & B. Schwartz (Eds.), *Language Acquisition Studies in Generative Grammar* (p. 85-118). Amsterdam: John Benjamins.

Clark, H. H. (1970). Word associations and linguistic theory. *New horizons in linguistics*, *1*, 271-286.

Clark, R., & Roberts, I. (1993). A computational model of language learnability and language change. *Linguistic Inquiry*, *24*(2), 299-345.

Clark, T., & Warmer, J. (Eds.). (2003). *Object modeling with the OCL: the rationale behind the Object Constraint Language* (Vol. 2263). Springer.

Colombo, J., & Bundy, R. S. (1981). A method for the measurement of infant auditory selectivity. *Infant Behavior & Development*, *4*, 229–231.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*, 84−107.

Crain, S., & Steedman, M. J. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing.* Cambridge: Cambridge University Press.

Crisma, P. (1990). *Functional categories inside the noun phrase: A study on the distribution of nominal modifiers* (Tesi di Laurea, Università degli Studi Ca' Foscari di Venezia).

Crisma, P. (1993). On adjective placement in Romance and Germanic event nominals. *Rivista di Grammatica Generativa, 18,* 61-100.

Croft, W. (1991). The evolution of negation. *Journal of Linguistics*, *27*(1), 1-27.

Crystal, T. H., & House, A. S. (1988). Segmental durations in connected-speech signals: Current results. *The journal of the acoustical society of America*, *83*(4), 1553-1573.

D'Addio, W. (1974). *La posizione dell'aggettivo italiano nel gruppo nominale*.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: evidence from eye movements. *Cognitive Psychology*, *42*, 317–367.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 125-127.

Darwin, C.H. (1877). A biographical sketch of a young infant. *Kosmos, 1*, 367-376.

Davies, C., Lingwood, J., Ivanova, B., & Arunachalam, S. (2021). Three-year-olds' comprehension of contrastive and descriptive adjectives: Evidence for contrastive inference. *Cognition*, *212*, 104707.

Déchaine, R. (1993). Predicates across categories. (Doctoral dissertation, University of Massachussetts).

Delsing, L. O. (1993). On attributive adjectives in Scandinavian and other languages. *Studia linguistica*, *47*(2), 105-125.

Demonte, V. (1999a). A Minimal Account of Spanish Adjective Position and Interpretation. In Jon Franco, Alazne Landa & Juan Martin (eds.), *Grammatical Analyses in Basque and Romance Linguistics. Papers in honor of Mario Saltarelli*, Amsterdam/Philadelphia, John Benjamins, 45-75.

Demonte, V. (1999b). El adjetivo: clases y usos. La posición del adjetivo en el sintagma nominal. In Ignacio Bosque & Violeta Demonte (dir.), *Gramática Descriptiva de la Lengua Española* (vol. 1: *Sintaxis básica de las clases de palabras*), Madrid, Espasa, 129-215.

Demonte, V. (2011). Adjectives. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning* (Vol. 2, 1314-1340). Berlin: De Gruyter.

den Dikken, M. (2006). *Relators and Linkers. The syntax of predication, predicate inversion, and copulas*. Cambridge, MA: MIT Press.

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition*, *11*(2), 159-184.
Di Sciullo, A. M., & Williams, E. (1987). *On the definition of word* (Vol. 14). Cambridge, MA: MIT press.

Dimitrova-Vulchanova, M. (2003). *Modification in the Balkan nominal expression. From NP to DP*. John Benjamins Publishing Company. Amsterdam/Philadelphia.

Dixon, R.M.W. (1982). Where have all the adjectives gone?: and other essays in Semantics and Syntax. *Janua Linguarum* (*107*). Series Maior. Walter de Gruyter, 1982.

Dixon, R. M. (2004). Adjective classes in typological perspective. In Dixon, R. M., Aikhenvald, A. I., Aĭkhenvalʹd, A. I., & Aikhenvald, A. Y. (Eds.). (2004). *Adjective classes: A cross-linguistic typology 1*, 1-49. Oxford University Press.

Dobrovie-Sorin, C. (1987). A propos de la structure du groupe nominal en Roumain. *Rivista di Grammatica Generativa 12*, 123-152.

Dowty, D. (1979). *Word meaning and Montague grammar*. Dordrecht: Reidel.

Dromi, E. (1987). *Early lexical development*. University Press.

Durrant, S., Delle Luche, C., Cattani, A., & Floccia, C. (2014). Monodialectal and multidialectal infants' representation of familiar words. *Journal of Child Language, 21*, 1–19.

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of psycholinguistic research*, *24*(6), 409-436.

Eilers, R. E., Oller, D. K., & Ellington, J. (1974). The acquisition of word-meaning for dimensional adjectives: The long and short of it. *Journal of Child Language*, *1*(2), 195-204.

Eisenbeiss, S. (2000). The acquisition of the Determiner Phrase in German child language. In M.-A. Friedemann & L. Rizzi (Eds.), *The Acquisition of Syntax: Studies in Comparative Developmental Linguistics* (p. 26-62). Harlow: Longman.

Emonds, J.E. (1976). *A Transformational Approach to English Syntax*. New York, Academic Press.

Fantz, R. (1958). Pattern vision in young infants. *The Psychological Record*, *8*, 43–47.

Fantz, R. (1964). Visual experience in infants: Decreased attention familiar patterns relative to novel ones. *Science*, *146*, 668–670.

Fantz, R. L. (1963). Pattern vision in newborn infants. *Science,* 19 April 140 (no. 3564), 296 – 297.

Fehri, A.F. (1997). Arabic antisymmetrical adjectives and possessive structure. *Linguistic Research 4*, 1-32.
Fenson, L., Marchman, V. A., Thal, D., Dale, P. S., & Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual. Second Edition.* Baltimore, MD: Brookes.

Fehri, A. F. (1998). Layers in the Distribution of Arabic Adverbs and Adjectives and their Licensing. In *Perspectives on Arabic Linguistics* (p. 9). John Benjamins.

Fehri, A. F. (1999). Arabic modifying adjectives and DP structures. *Studia linguistica*, *53* (2), 105-154.

Fernald, A., McRoberts, G., & Herrera, C. (1992). The role of prosodic features in early word recognition. *8th International Conference on Infant Studies*. Miami, May, 1992.

Fernald, A., McRoberts, G. W., & Herrera, C. (1994). Effects of prosody and word position on infants' ability to recognize words in fluent speech. *Unpublished manuscript*.

Fernald, A. (2002). Hearing, listening, and understanding: Auditory development in infancy. In: G. Bremner & A. Fogel (Eds.), *Handbook of infant development* (pp.35-70). Blackwell Publishers: Oxford, UK.

Fernald, A., Perfors, A. F., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the second year. *Developmental Psychology*, *42*, 98–116.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language. *Developmental psycholinguistics: On-line methods in children's language processing*, *44*, 97.

Fernald, A., Thorpe, K., & Marchman, V. A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective–noun phrases. *Cognitive psychology*, *60*(3), 190-217.

Floccia, C., Delle Luche, C., Durrant, S., Butler, J., & Goslin, J. (2012). Parent or community: Where do 20-month-olds exposed to two accents acquire their representation of words?. *Cognition*, *124*(1), 95-100.

Foppolo, F., & Panzeri, F. (2011). Do children know when their room counts as clean. *Proceedings of NELS42. Amherst: GLSA Publications.*

Foppolo, F., & Panzeri, F. (2013). Do children know when their room counts as clean? In S. Kan, C. M. Cantwell, & R. Staubs (Eds.), NELS 40: *Proceedings of the Fortieth Annual Meeting of the North East Linguistic Society*: Volume 1 (p. 205-218). Amherst, MA: GLSA.

Fox, J., & Weisberg, S. (2011). An R companion to applied regression. Thousand Oaks, California. Online: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178–210.

Frazier, L., Clifton, C., & Stolterfoth, B. (2008). Scale structure: Processing minimum standard and maximum standard scalar adjectives. *Cognition, 106*, 299-324.

Gasser, M., & Smith, L. B. (1998). Learning nouns and adjectives: A connectionist account. *Language and cognitive processes*, *13*(2-3), 269-306.

Gelman, S. A., & Markman, E. M. (1984). Implicit contrast in adjectives and nouns: Implications for word learning in preschoolers. *Journal of Child Language*, 12, 125–143.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *Center for the Study of Reading Technical Report;* no. 257.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj, *Language development: Language, thought, and culture* (Vol. 2). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. *Language acquisition and conceptual development*, *3*, 215-256.

Giegerich, H. J. (2005). Associative adjectives in English and the lexicon-syntax interface. *Journal of Linguistics*, 571-591.

Giorgi, A. & Longobardi G. (1991). *The syntax of Noun Phrases: Configuration, parameters and empty categories*. Cambridge, Cambridge University Press.

Girard, G. (1747). *Les vrais principes de la langue françoise, ou la parole reduite en méthode, conformément aux lois de l'usage en 16 discours*. chez Le Breton petit-fils d'Houry, imprimeur ordinaire du roi, rue de la Harpe.

Girard, G. (1982). 1747. *Les Vrais Principes de la Langue Française.* Lebreton, Paris.

Givón, T. (1970). Notes on the semantic structure of English adjectives. *Language*, 816-837.

Givón, T. (1979). Language typology in Africa: a critical review. *Journal of African Languages and Linguistics*, *1*(2), 199-224.

Gleitman, L. R., & Wanner, E. (1982). Language acquisition: The state of the art.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language learning and development, 1*(1), 23-64.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehen- sion in a new paradigm. *Journal of Child Language, 14,* 23–45.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental psychology*, *28*(1), 99.

Golinkoff, R. M., Hirsh-Pasek, K., Mervis, C. B., Frawley, W. B., & Parillo, M. (1995). Lexical principles can be extended to the acquisition of verbs. *Beyond names for things: Young children's acquisition of verbs*, 185-222.

Golinkoff, R. M., Ma, W., Song, L., & Hirsh-Pasek, K. (2013). Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned?. *Perspectives on Psychological Science*, *8*(3), 316-339.

Graff, D. (2000). Shifting sands: An interest-relative theory of vagueness. *Philosophical topics*, *28*(1), 45-81.

Grimm, J. (1870). *Deutsche Grammatik*. Vol 1. (2 ed). Berlin.

Grosjean, F. (1983). How long is the sentence? Prediction and prosody in the on-line processing of language. 501-530.

Gross, D., Fischer, U., & Miller, G. A. (1989). The organization of adjectival meanings. *Journal of memory and language*, *28*(1), 92-106.

Guasti, M. T. (2017). *Language acquisition: The growth of grammar*. MIT press.

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In *The cognitive neurosciences, 4th ed.* (pp. 819-836). MIT press.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304,* 438-41.

Hajek, J. (2004). Adjective classes: What can we conclude. *Adjective classes: A cross-linguistic typology*, 348-361.

Halberda, J., Sires, S. F., & Feigenson, L. (2006). Multiple spatially overlapping sets can be enumerated in parallel. *Psychological science*, *17*(7), 572-576.

Hall, D. G., & Graham, S. A. (1999). Lexical form class information guides word-to-object mapping in preschoolers. *Child Development*, *70*, 78–91.

Hall, D. G., Waxman, S. R., & Hurwitz, W. M. (1993). How two- and four-year-old children interpret adjectives and count nouns. *Child Development, 64,* 1651–1661.

Hallé, P. A., & de Boysson-Bardies, B. (1994). Emergence of an early receptive lexicon: Infants' recognition of words. *Infant behavior and development*, *17*(2), 119-129.

Hammer, A. E. (2002). *Hammer's German Grammar and Usage* (ed. M. Durrell). London.

Hay, J., Kennedy, C., & Levin, B. (1999). Scalar structure underlies telicity in "degree achievements". In *Semantics and linguistic theory* (Vol. 9, pp. 127-144).

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in cognitive sciences*, *9*(4), 188-194.

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child development*, 1021-1034.

Heller, D., & Chambers, C. (2014). Would a blue kite by any other name be just as blue? Effects of descriptive choices on subsequent referential behavior. *Journal of Memory and Language, 70*, 53-67.

Henderson, J. M. (2006). Eye movements. In C. Senior, T. Russell & M. S. Gazzaniga (Eds), *Methods in Mind* (pp. 171–191). Cambridge, MA: MIT Press.

Henderson, J. M., & Ferreira, F. (Eds.). (2004). The interface of language, vision, and action: Eye movements in the visual world. New York: Psychology Press.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The intermodal preferential looking paradigm: A window onto emerging language comprehension.* In: D. McDaniel, C. McKee, & H. S. Cairns (Eds.), *Language, speech, and communication. Methods for assessing children's syntax* (p. 105–124). The MIT Press.

Hoff, E. (Ed.). (2011). *Research methods in child language: A practical guide*. New York, NY: Wiley-Blackwell.

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). I. What does it take to learn a word?. *Monographs of the society for research in child development*, *65*(3), 1-16.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Hopper, P. J., & Thompson, S. A. (1984). The discourse basis for lexical categories in universal grammar. *Language, 60* (4), 703-752.

Huang, Y. T., & Snedeker, J. (2013). The use of lexical and referential cues in children's online interpretation of adjectives. *Developmental psychology*, *49*(6), 1090.

Huddleston, R., & Pullum J.K. (2002). The Cambridge Grammar of the English Language. Cambridge: Cambridge University Press.

Huddleston, R.D. (1976). *An Introduction to English Transformational Syntax*. London, Longman.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, *57*(4), 460-482.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta psychologica*, *137*(2), 151-171.

Inhelder, B., & Piaget, J. (1964). The early growth of logic in the child. New York: Harper & Row.

Jackendoff, R. (1977). X syntax: A study of phrase structure. *Linguistic Inquiry Monographs Cambridge, Mass, (2),* 1-249.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child development, 64*(3), 675-687.

Jusczyk, P. W., & Hohne, E. A. (1997). Infants' memory for spoken words. *Science, 277*(5334), 1984-1986.

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133-156.

Kamp, H. (1975). Two theories about adjectives. In *Formal Semantics of Natural Language*, ed. Edward L. Keenan, 123-155: Cambridge University Press.

Kamp, H. & Partee, B.H (1995). Prototype theory and compositionality. *Cognition, 57*(2), 129–191.

Katz, J. J. (1964). Semantic theory and the meaning of good. *Journal of Philosophy*, *61*(23), 739-766.

Katz, J.J. & Postal, P.M. (1964). *An integrated theory of linguistic descriptions*. Cambridge, MA, MIT Press.

Katz, N., Baker, E., & McNamara, J. (1974). What's in a name? A study of how children learn common and proper names. *Child Development, 45*, 469-473.

Kauschke, C. (2000). Der Erwerb des frühkindlichen Lexikons: Eine empirische Studie zur Entwicklung des Wortschatzes im Deutschen. Tübingen: Narr.

Kayne, R.S. (1994). *The Antisymmetry of Syntax*, Cambridge, MA, MIT Press.

Kennedy, C. (1999). *Projecting the adjective: The syntax and semantics of gradability and comparison*. Routledge.

Kennedy, C., & McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language, 81*(2), 345-381.

Kennedy, C. (2006). Comparatives, Semantics of. In K. Brown (Ed.), *Encyclopedia of language & linguistics* (p. 690-694). Amsterdam: Elsevier.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and philosophy, 30*(1), 1-45.

Kennedy, C., & Levin, B. (2008). Measure of change: The adjectival core of degree achievements. *Adjectives and adverbs: Syntax, semantics and discourse*, 156-182.

Kennedy, C., & McNally, L. (2010). Color, context, and compositionality. *Synthese*, *174*(1), 79-98.

Klibanoff, R. S., & Waxman, S. R. (1999). Syntactic cues to word meaning: Initial expectations and the development of flexibility. In *Proceedings of the 23nd Annual Boston University Conference on Language Development* (pp. 361-372).

Klibanoff, R. S., & Waxman, S. R. (2000). Basic level object categories support the acquisition of novel adjectives: Evidence from preschool-aged children. *Child Development*, *71*, 649–659.

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word n + 2. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1250–1255.

Knittel, M.L. (2005). Some remarks on adjective placement in the French NP. *Probus 17* (2), 185-226.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition, 95*(1), 95-127.

Kowalski, K., & Zimiles, H. (2006). The relation between children's conceptual functioning with color and color term acquisition. *Journal of Experimental Child Psychology*, *94*(4), 301-321.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature reviews neuroscience*, *5*(11), 831-843.

Kukona, A., Fang, S. Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. *Cognition, 119*(1), 23-42.

Kupisch, T. (2014). Adjective placement in simultaneous bilinguals (German–Italian) and the concept of cross-linguistic overcorrection. *Bilingualism: Language and Cognition*, *17*(1), 222-233.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of statistical software*, 82(13), 1-26.

Laenzlinger, C. (2000). French adjective ordering: perspectives on DP-internal movement types. *Generative Grammar in Geneva 1*, 55-104.

Laenzlinger, C. (2005a). French adjective ordering: perspectives on DP-internal movement types. *Lingua 115* (5), 645-689.

Laenzlinger, C. (2005b). Some Notes on DP-internal Movements. *Generative Grammar in Geneva 4*, 227-260.

Lamarche, J. (1991). *Problems for N-movement to Num-P*.

Landau, B., & Gleitman, L. R. (1985). Language and experience: Evidence from the blind child. Cambridge, Mass.: Harvard University Press.

Larson, R. K. (1995). *Olga is a Beautiful Dancer*. (Ms. In *SUNY Stony Brook, text of a paper delivered at the 1995 Winter Meeting of the LSA, New Orleans*).

Larson, R.K. (1998). Events and Modification in Nominals. In Devon Strolovitch & Aaron Lawson (eds.), *Proceedings from Semantics and Linguistic Theory VIII*, Ithaca, NY, CLC Publications, 145-168.

Larson, R.K. (1999). *Semantics of adjectival modification*. Talk at the *LOT Winter School* (18-22 January 1999, Dutch National Graduate School, Amsterdam).

Lassiter, D., & Goodman, N. D. (2013). Context, scale structure, and statistics in the interpretation of positive-form adjectives. In *Semantics and linguistic theory* (Vol. 23, pp. 587-610).

Lees, R.B. (1968). *The Grammar of English Nominalizations*. The Hague/Paris, Mouton.

Leffel, T., Xiang, M., & Kennedy, C. (2016, November). Imprecision is pragmatic: Evidence from referential processing. In *Semantics and linguistic theory* (Vol. 26, pp. 836-854).

Levi, J.N. (1973). Where Do All Those Other Adjectives Come From?. In Claudia Corum, T. Cedric Smith-Stark & Ann Weiser (eds.), *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*, Chicago, IL, Chicago Linguistic Society, 332-345.

Levi, J.N. (1974). On the Alleged Idiosyncrasy of Non-predicate NP's. In Michael W. La Galy, Robert A. Fox & Anthony Bruck (eds.), *Papers from the Tenth Regional Meeting of the Chicago Linguistic Society*, Chicago, IL, Chicago Linguistic Society, 402-415.

Levi, J.N. (1975). The syntax and semantics of non-predicating adjectives in English. (Doctoral dissertation, University of Chicago).

Levi, J.N. (1978). *The syntax and semantics of complex nominals*. New York, Academic Press.

Lewis, M. M. (1936). *Infant speech: A study of the beginnings of language.* London: Routledge & Kegan Paul.

Liao, S., & Meskin, A. (2015). Aesthetic adjectives: Experimental semantics and context-sensitivity. *Philosophy and Phenomenological Research*, 1-28.

Lidz, J., Waxman, S., & Freedman, J. (2003). What infants know about syntax but couldn't have learned: experimental evidence for syntactic structure at 18 months. *Cognition*, *89*(3), 295-303.

Macnamara, J. (1982). Names for things: a study of human learning. Cambridge, Mass.: MIT Press.

MacWhinney, B., & Snow, C. (1990). The Child Language Data Exchange System: An update. *Journal of Child Language*, *17*, 457–472.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.

Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science*, *31*, 133–156.

Mak, W. M., Tribushinina, E., & Andreiushina, E. (2013). Semantics of connectives guides referential expectations in discourse: an eye-tracking study of Dutch and Russian. *Discourse Processes, 50*(8), 557-76.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252-272.

Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological specificity of vowel contrasts at 18-months. *Language and Speech, 51*, 3–21.

Marchand, H. (1966). On the Attributive and Predicative Derived Adjectives and Some Problems Related to the Distinction. *Anglia LXXXIV* (2), 131-149.

Marchand, H. (1967). Expansion, Transposition, and Derivation. *La Linguistique 1*, 13-26.

Marinis, T. (2013). Online Psycholinguistic Methods in Second Language Acquisition Research.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive psychology*, *20*(2), 121-157.

Markman, E. M. (1989). Categorization and naming in children. Cambridge, MA: The MIT Press.

Markman, E. (1990). Constraints children place on word meanings. *Cognitive science, 14*, 57-77.

McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, *6*(2), 203-229.

McNally, L. (2009). The relative role of property type and scale structure in explaining the behavior of gradable adjectives. In R. Nouwen, R. van Rooij, U. Sauerland, & H.- C. Schmitz (Eds.), *ViC 2009* (Papers from the ESSLLI 2009 Workshop on Vagueness in Communcation) (p. 151-168). Berlin: Springer.

McNally, L. (2011). Color terms: a case study in natural language ontology. In *Workshop on the syntax and semantics of nounhood and adjectivehood.* Barcelona.

McNally, L. (2016). Modification. In M. Aloni & P. Dekker (Eds.), *Cambridge Handbook of Semantics*. Cambridge: Cambridge University Press.

McNally, L., & Boleda, G. (2004). *Relational adjectives as properties of kinds*. Colloque de Syntaxe et Sémantique. Paris.

Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*, *20*(2), 158-190.

Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology,* 20, 158–190.

Menuzzi, S. (1994). Adjectival positions inside DP. In Reineke Bok-Bennema & Crit Cremers (eds.), *Linguistics in the Netherlands 1994,* Amsterdam/Philadelphia, John Benjamins, 127-138.

Menyuk, P., Liebergott, J. W., & Schultz, M. C. (1995). Early language development in full term and premature infants. Hillsdale, NJ: Erlbaum.

Mintz, T., & Gleitman, L. (1998). Incremental language learning: two and three year olds' acquisition of adjectives. In *Proceedings of the 20th Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Erlbaum*.

Mintz, T. H. (2005). Linguistic and conceptual influences on adjective acquisition in 24- and 36-month-olds. *Developmental Psychology, 41,* 17–29.

Mintz, T. H., & Gleitman, L. R. (2002). Adjectives really do modify nouns: The incremental and restricted nature of early adjective acquisition. *Cognition*, *84*(3), 267-293.

Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive science*, *26*(4), 393-424.

Moro, A. (1997). *The raising of predicates*. Cambridge, Cambridge University Press.

Nadig, A.S. & Sedivy, J.C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science, 13*, 329–336.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*, 357–374.

Naigles, L. G., & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax. *Child Development*, *64*(6), 1665-1687.

Naigles, L. G., & Gelman, S. A. (1995). Overextensions in comprehension and production revisited: Preferential-looking in a study of dog, cat, and cow. *Journal of Child Language*, *22*(1), 19-46.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the society for research in child development*, 1-135.

Nelson, K. (1976). Some attributes of adjectives used by young children. *Cognition, 4*, 13-30.

Ninio, A. (2004). Young children's difficulty with adjectives modifying nouns. *Journal of Child Language, 31*, 255-285.

O'Grady, W. (1987). *Principles of grammar & learning*. University of Chicago Press.

Panzeri, F. (2011). L'acquisizione degli aggettivi. Quello che i bambini ci dicono sul significato degli aggettivi relativi. *Acquisizione del linguaggio, 4*, 84-95.

Panzeri, F., & Foppolo, F. (2012). Can children tell us something about the semantics of adjectives? In M. Aloni, V. Kimmelman, F. Roelofsen, G. W. Sassoon, K. Schulz, & M. Westera (Eds.), *Logic, Language and Meaning. Lecture Notes in Computer Science* (p. 170-179). Berlin, Heidelberg: Springer.

Panzeri, F., Foppolo, F., Guasti, M.T. (2013). Acquisition meets comparison: an investigation of gradable adjectives. In I. Caponigro, & C. Cecchetto (a cura di), *From Grammar to Meaning - The Spontaneous Logicality of Language*, 266-293. Cambridge University Press.

Partee, B. (1995). Lexical Semantics and Compositionality. In D. Osherson, L. Gleitman, & M. Liberman (Eds.), *An invitation to Cognitive Science: Part 1 Language* (2nd ed., 311-360). Cambridge, MA: MIT Press.

Porretta, V., Tucker, B. V., & Järvikivi, J. (2016). The influence of gradient foreign accentedness and listener experience on word recognition. *Journal of Phonetics, 58,* 1-21.

Postal, P.M. (1969). Anaphoric Islands. In Robert I. Binnick, Alice Davison, Georgia M. Green & Jerry L. Morgan (eds.), *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, Chicago, IL, Chicago Linguistic Society, 205-239.

Pullum, G.K., & Huddleston, R. (2002). Adjectives and adverbs. *The Cambridge grammar of the English language*, 525-595.

Pustejovsky, J. (1995). The generative lexicon. Cambridge (Mass.). London: the MIT Press.

Pustejovsky, J. & Jezek, E. (2008). Semantic Coercion in Language: Beyond Distributional Analysis. *Rivista di Linguistica 20* (1): 181-214.

Pyykkönen, P., & Järvikivi, J. (2010). Activation and persistence of implicit causality information in spoken language comprehension. *Experimental Psychology, 57,* 5-16.

R Development Core Team. 2005. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Ramaglia, F., (2008). La sintassi degli aggettivi e la proiezione estesa del NP. (Doctoral dissertation, Università degli studi di Roma Tre).

Rayner, K., & Clifton Jr, C. (2009). Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biological psychology*, *80*(1), 4-9.

Renzi, L., Salvi G., Cardinaletti, A. (2001). *La grande grammatica italiana di consultazione*. Bologna: Il Mulino.

Reznick, J. S. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics*, *11*(2), 145-166.

Rice, M. (1980). *Cognition to language: Categories, word meanings, and training*. University Park Press.

Rice, M.L. (1989). Children's language acquisition. *American Psychologist*, *44* (2), 149.

Robins, R.H. (1989). *A short history of linguistics*. London: Longman.

Roehrs, D., & Julien, M. (2014). Adjectives in German and Norwegian. *Adjectives in Germanic and Romance*, *212*, 245-261.

Ronat, M. (1974). Échelles de base et mutations en syntaxe française. (Doctoral dissertation, Université de Paris VIII).

Rosch, E. (1987). Linguistic relativity. *ETC: A Review of General Semantics*, 254-279.

Rotstein, C., & Winter, Y. (2004). Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural language semantics*, *12* (3), 259-288.

Sadler, L., & Arnold, D. J. (1994). Prenominal adjectives and the phrasal/lexical distinction. *Journal of linguistics*, 187-226.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27-52.

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*(1), 172-196.

Salverda, A. P., & Tanenhaus, M. K. (2017). The visual world paradigm. *Research methods in psycholinguistics: A practical guide*, 89-110.

Salverda, A. P., Kleinschmidt, D., & Tanenhaus, M. K. (2014). Immediate effects of anticipatory coarticulation in spoken-word recognition. *Journal of memory and language*, *71*(1), 145-163.

Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology, 35*(3), 668.

Sandhofer, C. M., Smith, L. B., & Luo, J. (2000). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of Child Language*, 27, 561–585.

Sandhofer, C. M., & Smith, L. B. (2001). Why children learn color and size words so differently: Evidence from adults' learning of artificial terms. *Journal of experimental psychology: General*, *130*(4), 600.

Sassoon, G., & Toledo, A. (2011). *Absolute and relative adjectives and their comparison classes*. Unpublished manuscript.

Scarano, A. (1999). Storia grammaticale deli'aggettivo: da sottoclasse di parole a parte del discorso. *Studi di Grammatica Italiana 18,* 57-90.

Schachter, P. (1985). Parts-of-speech systems. *Shopen I*, 3-61.

Scott, G.J (2002). The syntax and semantics of adjectival modification. (Doctoral dissertation, SOAS, University of London).

Scott, G.J. (1998). Stacked Adjectival Modification and the Structure of Nominal Phrases. *SOAS Working Papers in Linguistics and Phonetics 8*, 59-89.

Scott, G.J. (2002a). Stacked Adjectival Modification and the Structure of Nominal Phrases. In Guglielmo Cinque (ed.), *Functional Structure in DP and IP* (*The Cartography of Syntactic Structures*, vol.1), Oxford/New York, Oxford University Press, 91-120.

Sedivy, J. C. (2005). 17 Evaluating Explanations for Referential Context Effects: Evidence for Cricean Mechanisms in Online Language Interpretation. *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*, 345.

Sedivy, J. C. (2010). Using eyetracking in language acquisition research. *Experimental methods in language acquisition research*, 115-138.
Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*(2), 109-147.

Sedivy, J.C (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistic Research* 32(1). 3–23.

Seki, L. (1990). Kamaiurá (Tupí-Guaraní) as an active-stative language. *Amazonian linguistics: Studies in lowland South American languages*, 367-392.

Shatz, M. (1978). Children’s comprehension of their mothers’ question-directives. *Journal of Child Language*, *5*, 39–46.

Shipley, E.F., Smith, C. S., & Gleitman, L. (1969). A study of the acquisition of language: Free responses to commands. *Language, 45,* 322-342.

Siegel, M.E.A. (1976). Capturing the Russian Adjective. In Barbara H. Partee (ed.), *Montague Grammar*, New York, Academic Press, 293-309.

Siegel, M.E.A. (1979). Measure adjectives in Montague grammar. *Linguistics, philosophy, and Montague grammar*, 223-262.

Siegel, M.E.A. (1980). *Capturing the Adjective*. New York/London, Garland Publishing.

Smith, C. S. (1961). A Class of Complex Modifiers in English. *Language 37* (3), 342-365.

Smith, C. S. (1964). Determiners and Relative Clauses in a Generative Grammar of English. *Language 40* (1), 37-52.

Smith, C. S. (1991). *The Parameter of Aspect*. Dordrecht, Kluwer.

Snyder, W., Senghas, A., & Inman, K. (2001). Agreement morphology and the acquisition of noun-drop in Spanish. *Language Acquisition: A Journal of Developmental Linguistics*, *9*(2).

Soja, N. N. (1994). Young children's concept of color and its relation to the acquisition of color words. *Child Development*, *65*(3), 918-937.

Solt, S. (2011). Notes on the comparison class. In R. Nouwen, R. van Rooij, U. Sauerland, Solt, & H.-C. Schmitz (Eds.), *ViC 2009* (Papers from the ESSLLI 2009 Workshop on Vagueness in Communcation) (p. 189-206). Berlin: Springer.

Spelke, E. S. (1976). Infants' intermodal perception of events. *Cognitive Psychology*, *8*, 553-560.

Spelke, E. S. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology*, *15*, 626636.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive psychology*, *45*(4), 447-481.

Sproat, R., & Shih, C. (1988). Prenominal adjectival ordering in English and Mandarin. In *Proceedings of NELS* (Vol. 18, No. 2, pp. 465-489).

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*(6640), 381-382.

Svenonius, P. A. (1994). Dependent nexus: Subordinate predication structures in English and the Scandinavian languages. (Doctoral dissertation, University of California, Santa Cruz).

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*(2), 147-166.

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological science*, *13*(5), 480-484.

Syrett, K., Bradley, E., Kennedy, C., & Lidz, J. (2006). Shifting Standards: Children's Understanding of Gradable Adjectives. In K. U. Deen, J. Nomura, B. Schulz, & B. Schwartz (Eds.), *Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition* - North America (GALANA) (Vol. 4, p. 353-364). Cambridge, MA: UConn Occasional Papers in Linguistic.

Syrett, K. (2007). Learning about the structure of scales: Adverbial modification and the acquisition of the semantics of gradable adjectives. (Doctoral dissertation, Northwestern University).

Syrett, K., Kennedy, C., & Lidz, J. (2010). Meaning and Context in Children's Understanding of Gradable Adjectives. *Journal of Semantics, 27,* 1-35.

Tanenhaus, M. K., Spivey-Knowlton, M. J, Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 1632–1634.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of psycholinguistic research*, *29*(6), 557-580.

Tanenhaus, M. K., & Brown-Schmidt, S. (2008). Language processing in the natural world. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 1105-1122.

Taylor, M., & Gelman, S. (1988). Adjective and Nouns: Children's Strategies for Learning New Words. *Child Development, 59*(2), 411-419.

Thomas, D. C., Campos, J. J., Shucard, W., Ransay, D. S., & Shucard, J. (1981). Semantic comprehension in infancy: A signal detection analysis. *Child Development*, *52,* 798–803.

Thompson, S. (1988). A discourse approach to the category 'adjective'. In John A. Hawkins (ed.), *Explaining language universals*, Oxford, Blackwell Publishers, 167-185.

Thorpe, K., & Fernald, A. (2006). Knowing what a novel word is not: Two-year-olds 'listen through' ambiguous adjectives in fluent speech. *Cognition*, *100*(3), 389-433.

Thorpe, K., Baumgartner, H., & Fernald, A. (2006). Children's developing ability to interpret adjective–noun combinations. In *Proceedings of the 30th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Toledo, A., & Sassoon, G. W. (2011,). Absolute vs. relative adjectives-variance within vs. between individuals. *Semantics and linguistic theory* (Vol. 21, pp. 135-154).

Tremblay, A., & Ransijn, J. (2015). LMERConvenienceFunctions: Model selection and post-hoc analysis for (G) LMER models. *R package version, 2*(10).

Tribushinina, E. (2013). Adjective Semantics, World Knowledge and Visual Context: Comprehension of Size Terms by 2- to 7-Year-Old Dutch-Speaking Children. *Journal of Psycholinguistic Research, 42*, 205-225.

Tribushinina, E., Gillis, S., & De Maeyer, S. (2013). Infrequent word classes in the speech of two- to seven-year-old children with cochlear implants and their normally hearing peers: A longitudinal study of adjective use. *Journal of Pediatric Otorhinolaryngology, 77*, 356-361.

Tribushinina, E., van den Bergh, H., Ravid, D., Aksu-Koç, A., Kilani-Schoch, M., Korecky-Kröll, K., . . . Gillis, S. (2014). Development of adjective frequencies across semantic classes. A growth curve analysis of child speech and child-directed speech. *Language, Interaction and Acquisition, 5,* 185-226.

Tribushinina, E., & Mak, W. M. (2016). Three-year-olds can predict a noun based on an attributive adjective: evidence from eye-tracking. *Journal of child language*, *43*(2), 425.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: separating effects of lexical preference from garden-paths. *Journal of Experimental psychology: Learning, memory, and Cognition*, *19*(3), 528.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of memory and language*, *33*(3), 285-318.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of memory and language*, *35*(4), 566-585.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*(2), 89-134.

Trueswell, J. C., & Tanenhaus, M. K. (Eds.). (2005). Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions. Cambridge, MA: MIT Press.

Trueswell, J.C. (2008). Using eye movements as a developmental measure within psycholinguistics. In *Developmental Psycholinguistics: On-line Methods in Children's Language Processing* [Language Acquisition and Language Disorders 44], I.A. Sekerina, E.M. Fernandez & H. Clahsen (eds), 73–96. Amsterdam: John Benjamins.

Valois, D. (1991a). The internal syntax of DP and adjective placement in French and English. In Tim Sherer (ed.), *Proceedings of NELS 21*, Amherst, MA, GLSA, 367-381.

Van de Velde, D. (2006). Les adjectifs de groupe. *Travaux de Linguistique 53*, 135-154.

Van Rooij, R. (2011). Implicit versus explicit comparatives. In *Vagueness and language use* (pp. 51-72). Palgrave Macmillan, London.

Vanzon, C. A. (1834). *Grammatica ragionata della lingua italiana*. Livorno: Angeloni.

Vendler, Z. (1968). *Adjectives and Nominalization*. The Hague/Paris, Mouton.

Vincent, N. (1998). On the grammar of inflected non-finite forms (with special reference to Old Neapolitan). *Copenhagen studies in language*, *22*, 135-158.

Wagner, K., Storage, D., Dobkins, K., & Barner, D. (2013). Children's acquisition of subsective and intersective adjectives (Poster presented at BUCLD 38, George Sherman Union, Boston University).

Warren, B. (1984). *Classifying Adjectives*. Göteborg, Acta Universitatis Gothoburgensis.

Warren, B. (1988). Ambiguity and Vagueness in Adjectives. In *Studia Linguistica 42* (2), 122-172.

Wass, S. V. (2016). The use of eye-tracking with infants and children. In *Practical research with children* (pp. 50-71). Routledge.

Waxman, S. R., & Kosowski, T. D. (1990). Nouns mark category relations: Toddlers' and preschoolers' word-learning biases. *Child development*, *61*(5), 1461-1473.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. *Cognitive psychology*, *29*(3), 257-302.

Waxman, S. R., Senghas, A., & Benveniste, S. (1997). A cross-linguistic examination of the noun-category bias: Its existence and specificity in French- and Spanish-speaking preschool-aged children. *Cognitive Psychology*, *32*(3), 183–218.

Waxman, S.R. & Markow D.B. (1998). Object properties and object kind: twenty-one-month-old infants' extension of novel adjectives. *Child development, 69,* 1313-1329.

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology*, 36, 571–581.

Waxman, S., & Booth, A. (2001). Seeing Pink Elephants: Fourteen-Month-Olds' Interpretations of Novel Nouns and Adjectives. *Cognitive Psychology, 43*, 217-242.

Waxman, S., & Booth, A. (2003). The origins and evolution of links between word learning and conceptual organization: New evidence from 11-month-olds. *Developmental Science*, *6*(2), 128-135.

Waxman, S. R., & Teresa Guasti, M. (2009). Nouns, adjectives, and the acquisition of meaning: New evidence from Italian-acquiring children. *Language Learning and Development*, *5*(1), 50-68.

Weicker, M. & Schulz, P. (2018). Proceedings of the 42nd annual Boston University Conference on Language Development, ed. Anne B. Bertolini and Maxwell J. Kaplan, 790-802. Somerville, MA: Cascadilla Press.

Weicker, M. (2019). The role of semantic complexity for the acquisition of adjectives. (Doctoral dissertation, Johann Wolfgang Goethe-Universität Frankfurt am Main).

Weisleder, A., & Fernald, A. (2009). Real-time processing of postnominal adjectives by Latino children learning Spanish as a first language. In *Proceedings of the 24th Annual Boston University Conference on Language Development*. 611-621.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development, 7*(1), 49-63.

Winter, W. (1965). Transforms without kernels?. *Language*, *41* (3), 484-489.

Wood, S. N. (2006). On confidence intervals for generalized additive models based on penalized regression splines. *Australian & New Zealand Journal of Statistics, 48*(4), 445-464.

Wood, S. N. (2017). Generalized additive models: an introduction with R. CRC press.

Yoon, Y. (1996). Total and partial predicates and the weak and strong interpretations. *Natural language semantics*, *4*(3), 217-236.

Zahner, K., Kutscheid, S., & Braun, B. (2019). Alignment of f0 peak in different pitch accent types affects perception of metrical stress. *Journal of Phonetics*, *74*, 75-95.

Ziegler, J., & Pylkkänen, L. (2016). Scalar adjectives and the temporal unfolding of semantic composition: An MEG investigation. *Neuropsychologia, 89*, 161-171.