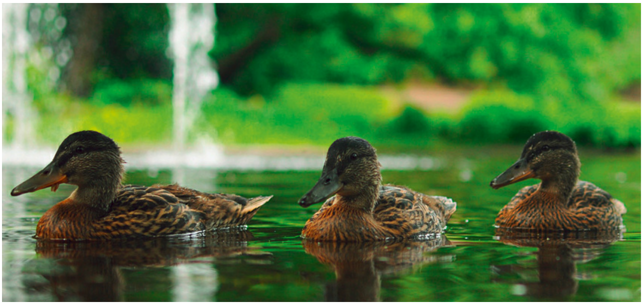
Universität Potsdam



Oda-Christina Brandt-Kobele

Comprehension of verb inflection in German-speaking children

Spektrum Patholinguistik – Schriften | 6

Spektrum Patholinguistik – Schriften | 6

Spektrum Patholinguistik – Schriften | 6

Oda-Christina Brandt-Kobele

Comprehension of verb inflection in German-speaking children

Universitätsverlag Potsdam

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.dnb.de/ abrufbar.

Universitätsverlag Potsdam 2014
http://verlag.ub.uni-potsdam.de/

Am Neuen Palais 10, 14469 Potsdam
Tel.: +49 (0)331 977 2533 / Fax: 2292
E-Mail: verlag@uni-potsdam.de

Die Schriftenreihe **Spektrum Patholinguistik – Schriften** wird herausgegeben vom Verband für Patholinguistik e. V.

ISSN (print) 1869-3822
ISSN (online) 1869-3830

Zugl.: Potsdam, Univ., Diss., 2012
Gutachter:
Prof. Dr. Barbara Höhle
Prof. Dr. Isabell Wartenburger
Datum der mündlichen Prüfung: 20.07.2012

Umschlagfotos:
Paul Sapiano, http://www.flickr.com/photos/peasap/518956588/
Mikko Luntiala, http://www.flickr.com/photos/mikko_luntiala/4720244099/

Online veröffentlicht auf dem Publikationsserver der Universität Potsdam
URL http://pub.ub.uni-potsdam.de/volltexte/2014/6204/
URN urn:nbn:de:kobv:517-opus-62046
http://nbn-resolving.org/urn:nbn:de:kobv:517-opus-62046

Zugleich gedruckt erschienen im Universitätsverlag Potsdam
ISBN 978-3-86956-216-2

for Greg

Acknowledgements

This dissertation would not have been possible without the support of many people to whom I would like to express my sincere gratitude.

First and foremost, I would like to thank Prof. Barbara Höhle, my supervisor. Her advice and guidance as well as her teaching and her own work have helped, influenced and inspired me. I thank for her accepting me as her doctoral student and giving me the chance to work and learn in the stimulating and fruitful environment of the Language Acquisition Group at the University of Potsdam. I am very grateful for her continuous support throughout the years.

I am very grateful to Prof. Isabell Wartenburger for accepting to serve on my dissertation commitree.

The empirical research in this thesis would have been impossible to conduct without the help of numerous people. I thank the student assistants who helped me conduct and run the experiments: Katja Hummel, Caroline Magister and Maja Stegenwallner, and I thank all children who participated in the experiments and to their parents for making it possible to conduct the research we are interested in.

A special thanks goes out to my friends and colleagues Silvana Poltrock and Frauke Berger for their support in preparing and running the experiments, as well for inspiring and motivating discussions on the findings. I cannot thank Tom Fritzsche enough for his help and support on all aspects of experimental work throughout the years. He gracefully shared his comprehensive knowledge about eye tracking, 'R', data analysis, statistical analysis and so much more. Working and traveling with you all has been a great experience that I am very thankful for.

Many thanks go also to my friends and colleagues Hans Trukenbrod, Susan Ott, Heike Herrmann and Ulrike Frank for support and friendship throughout the years.

Over the course of the past few years, many people have generously taken their time to discuss various aspects related to this dissertation. I want to thank the organizers and participants of the IASCL 2008 in Edinburgh, the RASCAL 2009 conference in Groningen, the DSGF 2009 conference in Osnabrück, and the IASCL 2011 in Montreal, for the chance to present and discuss my work presented in this thesis.

A special thanks goes to Ana-Teresa Perez-Leroux, Petra Hendriks, Bart Hollebrandse, Irina Sekerina, Anja Müller and Angela Grimm.

I gratefully acknowledge the financial support I have received from the collaborative research project CHLASC (*Characterizing Human Language by Structural Complexity*) which provided a part of the funding for this research.

It has been an enormous privilege to the able to spend the past few years doing what I like most. I want to thank my family, my parents and my sisters, for their unfailing love and support throughout the years.

Finally, words are not enough to express my love and gratitude to Greg. I thank you so much for your continuous support regarding all aspects of my work and my life. I thank you for helpful and stimulating discussions, for proof-reading earlier versions of this thesis, for help on LaTeX, but most importantly for keeping up my spirits and encouraging me to go on. This is for you and Carl, who always remind me of what matters most.

Abstract

Previous studies on the acquisition of verb inflection in normally developing children have revealed an astonishing pattern: children use correctly inflected verbs in their own speech but fail to make use of verb inflections when comprehending sentences uttered by others. Thus, a three-year old might well be able to say something like *The cat sleeps on the bed*, but fails to understand that the same sentence, when uttered by another person, refers to only one sleeping cat but not more than one.

The previous studies that have examined children's comprehension of verb inflections have employed a variant of a picture selection task in which the child was asked to explicitly indicate (via pointing) what semantic meaning she had inferred from the test sentence. Recent research on other linguistic structures, such as pronouns or focus particles, has indicated that earlier comprehension abilities can be found when methods are used that do not require an explicit reaction, like preferential looking tasks. This dissertation aimed to examine whether children are truly not able to understand the connection the the verb form and the meaning of the sentence subject until the age of five years or whether earlier comprehension can be found when a different measure, preferential looking, is used. Additionally, children's processing of subject-ver agreement violations was examined.

The three experiments of this thesis that examined children's comprehension of verb inflections revealed the following: German-speaking three- to four-year old children looked more to a picture showing one actor when hearing a sentence with a singular inflected verb but only when their eye gaze was tracked and they did not have to perform a picture selection task. When they were asked to point to the matching picture, they performed at chance-level. This pattern indicates asymmetries in children's language performance even within the receptive modality.

The fourth experiment examined sensitivity to subject-verb agreement violations and did not reveal evidence for sensitivity toward agreement violations in three- and four-year old children, but only found that children's looking patterns were influenced by the grammatical violations at the age of five.

The results from these experiments are discussed in relation to the existence of a production-comprehension asymmetry in the use of verb inflections and children's underlying grammatical knowledge.

Zusammenfassung

Experimentelle Studien zum Erwerb der Verbflexion bei sprachunauffälligen Kindern haben ein überraschendes Muster aufgezeigt. Kinder im Alter von drei und vier Jahren verwenden Verbflexionsendungen anscheinend korrekt in ihrer eigenen Sprachproduktion, aber sie scheinen unfähig zu sein, Verbflexionen in den Äußerungen anderer zu verstehen. Ein Kind ist also problemlos in der Lage ‚Sie schläft auf dem Bett' zu sagen, wenn es die Position von z.B. einer Katze beschreiben möchte. Gleichzeitig scheint es nicht zu verstehen, dass sich ein Satz wie ‚Sie schläft auf dem Bett' auf nur *eine* schlafende Katze und nicht mehrere bezieht.

Das Verständnis von Sätzen, in denen der einzige Hinweis auf die Anzahl der Handelnden (den Numerus des Subjekts) die Verbflexion ist, wurde bislang nur mit ‚Zeige-Experimenten' untersucht. In solchen Sprachtests soll das Kind durch eine Zeigegeste auf eines von zwei vorgegebenen Bildern explizit anzeigen wie es den vorgegebenen Satz verstanden hat. Aktuelle Studien, die das Verständnis von sprachlichen Elementen wie Pronomen und Fokuspartikeln bei Kindern untersucht haben, lassen erkennen, dass die Testmethodik einen erheblichen Einfluss auf die kindlichen Sprachverständnisfähigkeiten zu haben scheint. Wenn man Methoden verwendet, die keine explizite Reaktion von Seiten der Kinder verlangen, findet man korrektes Verständnis schon bei jüngeren Kindern. Das Ziel dieser Dissertation war es zu untersuchen, ob drei- und vierjährige Kinder tatsächlich nicht in der Lage sind die Beziehung zwischen Verbform (Art der Verbflexion) und Subjektbedeutung (Numerus des Subjekts) zu verstehen oder ob man korrektes Sprachverständnis in jüngeren Populationen finden kann, wenn eine alternative Testmethode, die Messung der Augenbewegungen, verwendet wird. Zusätzlich wurde untersucht ob Kinder im gleichen Alter Verletzungen der Subjekt-Verb-Kongruenz in auditiv präsentierten Sätzen entdecken.

Drei Experimente dieser Dissertation, die das kindliche Sprachverständnis in Bezug auf Verbflexion untersucht haben bringen folgendes Muster zum Vorschein:

Deutsch-sprachige Kinder im Alter von drei bis vier Jahren schauten mehr zu einem Bild, auf dem nur ein Akteur zu sehen war, wenn sie einen Satz mit einem singular flektierten Verb hörten (*Sie streichelt eine Katze*). Andererseits schauten sie mehr zu einem Bild, auf dem zwei Akteure zu sehen waren, wenn sie einen Satz mit einem plural flektierten Verb hörten (*Sie streicheln eine Katze*). Wenn sie hingegen gebeten wurden, auf das korrekte Bild zu zeigen, reagierten sie nicht besser als es der Zufall erwartet hätte, d.h. sie waren nicht in der Lage einen Satz einem entsprechenden Bild zuzuordnen. Dieses Ergebnismuster deutet auf die Existenz von (methodenabhängigen) Asymmetrien innerhalb einer sprachlichen Modalität, dem Sprachverständnis, hin. Das vierte Experiment untersuchte die kindliche Sensitivität gegenüber der Verletzung von Subjekt-Verb-Kongruenz. Hier zeigte sich, dass das Blickverhalten von fünfjährigen Kindern von der Grammatikalität der Testsätze beeinflusst war, während keine Evidenz für das Erkennen von Grammatikalitätsverletzungen bei jüngeren Kindern gefunden werden konnte. Das asymmetrische Performanzmuster innerhalb der rezeptiven Modalität, das in dieser Arbeit gefunden wurde, erlaubt Rückschlüsse auf die Annahme einer Produktions-Verständnis-Aymmetrie und somit auch auf Theorien zur grammatischen Entwicklung bei Kindern.

Contents

List of Figures

List of Tables

1 Introduction

One of the main aims in language acquisition research has been to describe children's linguistic knowledge at different stages in development. One aspect of this has been to describe the developmental time course of the emergence of verb inflection morphology. The acquisition of the grammatical morphemes that mark subject-verb agreement has been of interest because it is assumed to reflect general properties of the morphosyntactic development in children.

Many languages exhibit agreement between subject and verb. This means that the forms of the two words match in some way. In English and German, for example, subject and verb agree in (person and) number, so that in a sentence like *the book is great* both subject and verb are singular, while in a sentence like *the books are great* both are plural. Agreement is viewed as a fundamental aspect of language processing, so studies of grammatical agreement have played a central role in language production (and to a lesser extend in language comprehension) studies in adults.

When children first produce sentences, they very often omit the verb inflectional affixes that mark subject-verb agreement, just as they don't realize most function morphemes in their early speech. Interestingly, when children begin to use inflected verbs in their production (around their second birthday), they do not do so consistently. Thus, children produce finite (adult-like) and non-finite (non adult-like) verbs in sentences at the same time during development. But if they use a finite verb, it is in most cases inflected correctly. This pattern has generated a great deal of empirical and theoretical interest and a wealth of explanations have been put forward. These can be roughly divided into (1) positions that view the instances of inflected verbs found in spontaneous speech as evidence that children have acquired abstract adult-like knowledge of verb inflections and subject verb agreement and (2) positions that view these instances as reflections of the input which children have memorized in the form of unanalyzed (or partly analyzed) chunks of speech. According to the former position, the strong nativist approach, children possess adult-like grammatical knowledge, but either performance factors (e.g. Phillips, 1995) or additional constraints in the child's grammar (e.g. Wexler, 1994) hinder them from applying the adult-like grammatical knowledge in every obligatory context. According to the latter position, the constructivist approach, children lack adult-like knowledge of subject verb agreement and/or verb inflections, but either recall unanalyzed chunks of speech when producing sentences and/or (a little later in development) rely on rules that they have generated from the input. These rules can well differ from the adult grammar, for example in the sense that they rather apply to 'superficial' distributional properties of the speech stream and not to grammatical features (such as '+/-finit' or '+/-singular').

To distinguish between those two views on children's early grammatical knowledge, it can be helpful to turn to the receptive modality. In general, research suggests that children understand more about their language than they are themselves producing (e.g. Goldin-Meadow, Seligman, and Gelman, 1976; Naigles, 2002). A handful of studies has investigated children's very early sensitivity to the presence of verb inflections in sentences. These studies have found that children detect verbal inflectional morphemes in speech passages and that they are even aware of (some of) the distributional dependencies that underlie the presence of these grammatical morphemes (e.g. Soderstrom, 2002; Soderstrom, White, Conwell, and Morgan, 2007; Nazzi, Barriere, Goyet, Kresh, and Legendre, 2011).

Even though children's ability to detect grammatical dependencies between subject noun phrase and verb form before their second birthday is impressive, this is only half the story of language comprehension. Other studies that have investigated if and when children are able to understand the connection between the *form* of the verb and the *meaning* of the subject noun have found a surprisingly late ability to do so. Children below the age of five or six years were not able to select a picture showing one referent when presented with a sentence containing a singular inflected verb (*The fish swims*) and to select a picture showing multiple referents when presented with a sentence that contained a plural inflected verb (*The fish swim*).[1] This inability has not only been found in children speaking English (Johnson et al., 2005), a language that is known for its impoverished morphological system, but also in children speaking Spanish (Pérez-Leroux, 2005), a language with a richer inflectional system that allows for subject-less sentences. The problem of finding a singular or plural referent cannot be explained by claiming that children below the age of five would not know about linguistic manifestations of the difference between 'one' and 'more than one'. Recent research has shown that children can correctly interpret sentences that contain a number marked noun, a deteminer/quantifier and a number marked auxiliary (*There is a car* vs. *There are some cars*) at the age of two years (e.g. Kouider, Halberda, Wood, and Carey, 2006; Wood, Kouider, and Carey, 2009)

The findings that children are able to use verb inflections in production starting at age two but that they are unable to use the same grammatical morphemes in comprehension until three or four years later indicate the existence of a rather large production-comprehension asymmetry. Such an asymmetry, if truly present, has different logical consequences for the two types of theories regarding children's early grammatical knowledge. If one assumes that children possess adult-like rules to produce verb inflections rather early, comprehension deficits have to be explained by performance factors. Since comprehension is usually viewed as being 'easier' than production (a claim that will be discussed in this thesis), a production-preceding-comprehension pattern is challenging to explain. If one on the other hand assumes that children's first verb inflection productions are rather based on memorized constructions

[1] Critically, in those test sentences, the number information on the subject noun was not available, either due to the use of irregular plural (homophone to the singular form), like *the deer is running / the deer are running* or due to verbs that disguised the number information on the subject due to coarticulation, like *the cat sleeps / the cats sleep*.

than on adult-like rules, a production-comprehension asymmetry does not come too surprising. It would be explained by the assumption that children possess a set of rules that allows them to produce many seemingly correct utterances but that does not relate to meaning (yet).

But before further discussing possible theoretical implications of a production-comprehension asymmetry, it is important to experimentally verify whether the claim of such an asymmetry is actually warranted. Recent research on the comprehension of pronouns and focus particles as well as word order regularities indicates that children (sometimes) display linguistic knowledge earlier in development when they are tested with a methodological paradigm that poses as little task demands as possible on them (e.g. Bergmann, Paulus, and Fikkert, 2011; Höhle, Berger, Müller, Schmitz, and Weissenborn, 2009; Sekerina, Stromswold, and Hestvik, 2004; Chan, Meints, Lieven, and Tomasello, 2010).

Therefore, the aim of the present thesis is to investigate whether German-speaking children can make use of verb inflections in sentence comprehension to determine the semantic number of the sentence subject when tested using the preferential looking paradigm. If children show earlier comprehension than found with a picture selection task, the claim of a production-comprehension asymmetry needs to be reevaluated. The research questions of the present thesis are as follows:

1. Are German-speaking children aged three to four years able to understand the connection between the *form* of a verb inflection and the *meaning* of a sentence subject in a preferential looking task?

2. Is the ability to make use of verb inflections during sentence comprehension (to infer the number of a sentence subject) influenced by methodological factors, ,i.e. by task demands?

3. Are German-speaking children aged three to six years able to detect the dependency between the *form* of the subject and the *form* of an inflected verb, i.e. are they sensitive to subject-verb agreement violations?

Four experiments will be presented that aimed to find answers to the questions. Experiment 1 was pursued determine whether German-speaking children were able to detect subject-verb agreement violations in simple SVO-sentences in a preferential looking task. In accordance with other preferential looking studies that investigated sensitivity to well-formedness of sentences (e.g. Kedar, Casasola, and Lust, 2006; Zangl and Fernald, 2007), it was measured whether children showed 'disruption' in finding a visual referent when the target noun was presented in an ungrammatical (with subject-verb agreement violation) as opposed to a grammatical sentence (with correct subject verb agreement). Experiment 2, 3 and 4 investigated whether German-speaking children were able to use verb inflection information in sentence processing as a cue to subject number. Experiment 2 was a pure eye tracking task, while children in Experiment 3 and 4 had to perform a picture selection task. The aim was to compare comprehension abilities in the two types of task. The comprehension experiments were administered to shed further light on the proposed production-comprehension asymmetry in the acquisition of verb inflection and the rules that might underlie children's processing of verb inflections.

The present thesis is organized as follows. In the first part, 'Theoretical Background', I will lay out two broad theoretical accounts that were proposed to account for children's acquisition of verb inflection (Chapter 2) and present the production data that has given rise to those accounts. In Chapter 3 will present production data from English-, German-, Spanish-, Italian- and Dutch-speaking children, because those are the languages that have been examined with regard to the receptive modality as well. Then I will focus on receptive grammar. First, receptive studies that tap children's sensitivity to the *form* of subject-verb agreement morphology will be presented in Chapter 4. Such sensitivity to form has been examined in very young children, using preferential listening procedures, and in school-aged children, using grammaticality judgement tasks that require metalinguistic processing. While presenting data on children's receptive abilities, theoretical frameworks that have been claimed to account for the data (like the Competition Model), will be described. In Chapter 5, I will present studies that examined whether children are able to use verb inflection to infer the semantic meaning of the utterance. As mentioned, only very late comprehension abilities have been found in more than one language, giving rise to a production-preceding comprehension pattern. This pattern will be discussed more thoroughly in Chapter 6. I will present other areas of language acquisition in which such a counterintuitive pattern has been observed. Additionally, I will discuss the logical implications for the theories of verb inflection acquisition if a true production-comprehension asymmetry is in place. In the second part, 'Empirical Investigations', the four experiments mentioned above will be presented and the results of each will be discussed. In the third part, 'General Discussion', I will try to reconcile the data presented in this thesis and gathered in the experiments, and try to formulate one account that can explain children's use of verb inflections in the productive and receptive modality.

Part I

Theoretical Background

2 Acquisition of verb agreement morphology

In this chapter I lay out why the emergence of verb inflections that mark subject-verb agreement has been a topic of language acquisition research for nearly half a decade. One reason, among others, is the early absence of these grammatical morphemes in child speech, combined with fact that these verb inflections are only produced variably once they start to be used by children. This co-occurrence of adult-like and non adult-like verb forms has generated considerable theoretical and empirical interest. I first specify what aspects of subject-verb agreement are under investigation in this thesis, and then present two theoretical approaches to the acquisition of verb inflection. Under one approach, the child is assumed to have adult-like knowledge from early on, so that variable realization of subject-verb agreement morphemes have to be explained via performance factors. Under the alternative approach, the child is assumed to have linguistic knowledge that differs from the adult-like one, giving rise to variable production patterns.

2.1 Subject-verb agreement in child language

It is commonly recognized that children's very early language production lacks functional elements, with verb inflection being no exception. Language production in the second and third year of life is then marked by co-occurrence of finite (adult-like) and non-finite (non adult-like) verb forms. This suggests that children acquire some knowledge about subject-verb agreement rules and verb inflection morphology early on, since they (at least sometimes) produce finite forms. But they do not use inflected finite verbs in every obligatory grammatical context, which calls for explanations. The phenomenon of non-finite verbs in children's early sentences is often referred to as 'root infinitives' (e. g. Radford, 1990; Hoekstra and Hyams, 1998) or as 'optional infinitives' (e. g. Wexler, 1994). Theories on the acquisition of verb inflections are mainly put forward to accommodate the co-occurrence pattern of finite and non-finite verb forms in early production.

Such theories mainly differ in whether they attribute early abstract morphosyntactic knowledge to children, or whether they assume that children initially rely on alternative rules, which are different from the target-like grammar (or even simply recall memorized chunks of speech). The former account is mainly supported by proponents of the nativist framework who assume a strong continuity between child and adult grammars. According to this view, children have adult-like linguistic competence and differences in the use of morphosyntactic rules have to be ascribed to performance factors. The latter account still attributes some kind of linguistic knowledge to children that allows them to produce target-like utterances, but this knowledge might be quite different from the rules employed by adults. This position is rather supported by proponents of a weak continuity framework or even constructivist accounts to language acquisition.

Critically, most theoretical work regarding the acquisition of subject-verb agreement and verb inflections relies on production data. The matter of subject-verb agreement is not of primary concern in most studies investigating root infinitives (but see Clahsen and Penke (1992)). The common consensus within the nativist framework is that subject-verb agreement does not seem to be problematic for children. This is based on the observations that children produce either a non-finite verb, which of course lacks agreement, or a finite one, which is in most cases inflected to correctly agree with the sentence subject (e. g. Phillips, 1995). Thus, spontaneous speech data often creates the impression that the acquisition of subject-verb agreement is easy and effortless for normally developing children.

Interestingly, adult-like comprehension of verb inflection morphology has been found to appear much later than adult-like production.[2] This pattern of production preceding comprehension is counter-intuitive and poses difficulties on theories of language acquisition. On the surface, the acquisition path for verb inflection morphology, based on previous studies, seems to be as follows. At first, verb inflections are neither produced nor understood, later they are produced correctly, but not understood, and only later have children mastered verb inflection morphology to the point where they can make use of them in sentence comprehension and production. To explain such a pattern, theories that credit children with early adult-like knowledge of morphosyntactic rules have to explain how performance factors can have a stronger detrimental effect in children's comprehension than in their production. Theories that assume non adult-like knowledge need to posit rules that can give rise to seemingly correct production while not allowing the child to make use of verb inflections in sentence comprehension.

The main topic of the current work is *number agreement* between subject and verb. This language phenomenon deals with the fact that in English and German (and many other languages), the subject and the verb have to agree in number (and person), at least in indicative present tense. The morpheme that marks the agreement relation is attached to the verb, but the number value associated with it is determined by the number value of the subject. Put differently, the subject is the agreement controller and the verb is the agreement target or dependent element (e. g. Nicol, Forster, and Veres, 1997; Eberhard, Cutting, and Bock, 2005). The number of the sentence subject can be marked morphologically (e. g. *dog*, *dogs*), lexically in the

[2]Throughout this work, the term *comprehension* is used to refer to children's ability to make use of verb inflection during sentence comprehension. It is often thought that verb agreement marking does not in itself carry meaning, and therefore might not be comprehended or understood in isolation, as for example a word form like *dog* can. But for reasons of simplicity, the process of using verb inflection information during sentence comprehension is called comprehension throughout this thesis.

case of irregular noun plurals (e. g. *child, children*), by a pronoun (e. g. *we, they*) or via quantifiers (e. g. *some, all, many*) and numerals (e. g. *one, two*). Semantically, the number marking on the noun canonically indicates the numerosity of the referent, as either one or more than one. The morphological information attached to the verb, however, is neutral with respect to, for example, whether a plurality of events is denoted. This morphological information rather refers to the number value of the subject, and that is why it is called agreement marking (e. g. Bock, Nicol, and Cutting, 1999; Pawlowska, Leonard, Camarata, Brown, and Camarata, 2008).[3]

The agreement relation between subject and verb is one of the syntactic dependencies a child has to learn in the process of language acquisition, sooner or later. To establish agreement, a property inherent to the sentence subject is matched against a formal property of the verb, the inflectional category (e. g. Nazzi et al., 2011; Foote and Bock, 2011). The agreeing elements can occur in either adjacent (e. g. *the girl sings*) or non-adjacent (e. g. *the girl that is playing in the garden every day likes to sing*).[4] The detection of syntactic dependencies involves the recognition of distributional dependencies (e. g. Nazzi et al., 2011) and/or the matching of underlying grammatical features (adult-like) (e. g. Eberhard et al., 2005; Nevins, Dillon, Malhotra, and Phillips, 2007). It is hypothesized that the former is a developmental prerequisite for the latter (e. g. Soderstrom, 2002), but it is an open question when a transition from 'surface-based' distributional processing to abstract linguistic (feature-based) processing occurs. The answer to this question critically depends on the theoretical stand one takes.

[3] In subjectless sentences as in a pro-drop language like Spanish, the number information attached to the verb can be the only cue as to subject number. The same is true for English subjects that are not overtly marked for plural, e. g. *deer*. Such cases are exploited in experimental settings, as will be seen in Chapter 5.

[4] The dependency can only be called adjacent if one refers to the words. If one assumes the single morphemes to be the critical entities of language processing, even the dependency between a singular subject and the verb inflectional morpheme would be considered to be non-adjacent. (Nazzi et al., 2011).

The current work focuses on lexical main verbs. Further, it takes a non-lexicalist view of morphology, according to which at least main verbs and inflectional morphemes enter syntactic operations in production as independent elements *in the adult language processing system*. How children produce and comprehend verb inflection is unclear and a matter of debate, as the different claims from the nativist and constructivist approaches show. It is unclear whether auxiliaries and modals are stored as whole forms in the lexicon or whether morphological building rules apply here in sentence production as well. It is for example claimed that auxiliaries are generated in inflectional heads and do not have to be raised to INFL (e. g. Wexler, 1994). [5] Thus, the present work deals with regular present tense inflectional morphology on main verbs that encodes number contrast. Other verbs or inflectional phenomena are only mentioned for reasons of comparison.

In the remainder of this chapter, I will present the pertinent theoretical accounts on the acquisition of subject-verb agreement and verb inflection that are relevant for this thesis. This includes a sketch of the different kinds of rules young children could employ when processing and producing verb inflections. Then, I will present previous research on children's use of verb inflections which is considered relevant for this thesis, in particular on English, Spanish and German. First, verb inflection in language production will be presented, with the aim of determining when normally developing children start to produce inflected verbs and when they do so at an adult level. Additionally, factors which critically influence the production of verb inflection will be discussed.

Following this, the receptive side of processing verb inflections will be considered. First, I will present studies from English, Dutch and French that demonstrate children's early sensitivity to the presence of verb inflection morphology. Whether such sensitivity points to abstract grammatical knowledge of subject-verb agreement or can be explained with statistical processing

[5]But is has repeatedly been noted that auxiliaries are always finite in young children's production (e. g. Phillips, 1995). Thus, utterances like **be tired* or *müde sein* are virtually unattested in child speech.

mechanisms will be discussed. Second, I will present research on the later developing, metalinguistic detection of verb inflection that is necessary in grammaticality judgement tasks. This is combined with a rough sketch of the Competition Model. This framework provides interesting explanations for children's use of verb agreement in grammaticality judgement tasks. Finally, a more elaborate view of the receptive modality will be presented, including not only sensitivity to verb morphology but actual understanding of what semantic aspects inflectional morphemes can infer.

It will become evident that there seems to be great variation in children's ability to produce, recognize and comprehend verb inflections. More precisely, an asymmetry between production and comprehension abilities in children seems to arise. The experiments presented in this work try to answer some of the questions related to the apparent production-comprehension asymmetry in the processing of verb inflections.

2.2 Theoretical accounts of the acquisition of verb inflection

Children's early utterances are characterized by the co-occurence of finite and non-finite verb forms over a prolonged period of time. One usually finds utterances like those in 2.1 and 2.2 in the speech of a normally developing two year old learning German (examples taken from Clahsen, Penke, and Parodi (1993)) and utterances like those in 2.3 and 2.4 in a two year old learning English (examples taken from Harris and Wexler (1996)). Interestingly, when children use inflectional morphology, they typically do it right; either they produce a correctly inflected verb or they omit agreement morphology at all. This acquisition pattern has repeatedly been noticed since the early days of child language research (Brown, 1973; Cazden, 1968; Phillips, 1995; Wexler, 1994). Research in languages other than English has revealed that children are indeed able to distinguish between the linguistic properties that are associated with non-finite and finite verb forms (e. g. Verrips and Weissenborn, 1992). This evidence is mostly based on the relationship between the finiteness status of the verb and the placement of the verb in a sentence. It has repeatedly been found that children position verbs according to their finiteness status. Thus, it cannot be argued that children persist in using non-finite forms concurrently with finite verbs because they would not know about the difference.

(2.1) *die Ente kommt nicht hin.*
the duck come-3SG not here
the duck does not come here

(2.2) *ein Schal haben.*
a shawl have-INF
(I) have a shawl

(2.3) This goes in there.

(2.4) *Patsy need a train.

Cross-linguistic research has further revealed that ungrammatical non-finite forms are only very rarely found in the early sentences of children acquiring Romance languages like Spanish (Grinstead, 1994) or Italian (Guasti, 1993).[6] An overview is provided in Chapter 3.4. There seems to be a connection between the morphological richness of a language and the proportion of incorrect use of non-finite verbs in children's language. The degree of morphological richness of a language (and its use in child directed speech) is positively related to the rate at which children produce verb inflections in their own speech (e. g. Phillips, 1995; Xanthos, Laaha, Gillis, Stephany, Aksu-Koc, Christofidou, Gagarina, Hrzica, Ketrez, Kilani-Schoch, Korecky-Kröll, Kovacevic, Laalo, Palmovic, Pfeifer, Voeikova, and Dressler, 2011). Such a positive correlation is not necessarily intuitive, as one could rather assume that languages with only sparse morphological paradigms would be easier to acquire because there are simply fewer forms to learn. But rich morphological paradigms imply variation in the input, and

[6] But see Pratt and Grinstead (2007) for a proposal that there are indeed uninflected verbs in child Spanish, which take the form of bare stems.

variation has been found to aid in the language learning process (see e. g. Naigles and Hoff-Ginsberg (1998) on the role of variable sentence structures in verb learning and Gómez (2002) for the role of variability in artificial language learning). Therefore, early exposure to a variety of inflectional forms may help children to focus on these forms and the differences in meaning that they carry, or the different morphosyntactic dependencies they enter into (Xanthos et al., 2011).

Several theories have been put forward to explain the lack of inflectional categories in early child grammar and the following co-occurrence of finite and non-finite verbs in child production. Some of these theories follow the assumption that the child's grammar is fundamentally similar to that of an adult, reflecting the Strong Continuity approach to language acquisition. That is, the linguistic competence is assumed to be in place early on, and the presence of non-finite (non adult-like) verbs in children's root clauses has to be explained by performance factors. These appear in the form of extra-syntactic difficulties, e. g. processing demands (Phillips, 1995) or pragmatic principles (Hyams, 1999). Hyams (1999) for example claims that "given that children use agreeing forms of the verbs with a high degree of accuracy, it cannot the the case that root infinitives arise from a lack of knowledge of the specifier-head agreement requirement or of the specific forms themselves." (pg. 395). Wexler (1994) argues along the same lines.

Other theories follow the assumption that children's linguistic knowledge is different from the adult's knowledge, such that children employ rules that differ from the adult ones, or that their language processing is based on other (less abstract) aspects of language. For example, children might base their preference for a grammatical structure in an experiment tapping early sensitivity to grammatical well-formedness (e. g. Santelmann and Jusczyk, 1998), on phonological properties and statistical regularities of the input instead of abstract grammatical rules. Thus, children would still base their preference on a linguistic rule, but this rule would be fundamentally different from the one employed by the adult. Interestingly, this latter way of viewing children's early linguistic knowledge subsumes proponents of the weak continuity assumption and of the no continuity assumption. Strict no-continuity constructivist accounts go a step further and claim that children's early linguistic representations are simply made up of memorized chunks (*constructions*) that reflect the input but lack any underlying structure (e. g. Tomasello, 2000; Abbot-Smith and Tomasello, 2006; Theakston, Lieven, and Tomasello, 2003). Weak continuity theories acknowledge differences between children's and adult's grammatical knowledge, thus their linguistic competence, but still assume that children's early representations can be described in grammatical terms like adult's knowledge. This allows that certain aspects of the child's grammar can be different from the target grammar. Both theoretical stands put a greater burden on the question of how children finally end up with an adult grammar. [7]

With regard to verb inflection, the 'rules' a child could have acquired when she begins to use verb inflections could have one of the following forms (see 2.5 and 2.6).[8]

(2.5) $V_{INF} \rightarrow V_{3SG}$ iff SubjNP_{3SG}

(2.6) V+/Ø/ $\rightarrow$ V+/s/ following he, she, it, NP+/Ø/

The rule stated in 2.5 reflects a rule employed by adults under the assumptions of non-lexical morphology and an encapsulated model on language production (e. g. Thornton and MacDonald, 2003; Bock and Miller, 1991).[9] The model is clearly simplified, but sufficient to incorporate the number distinction. In

[7] The acquisition of verb morphology is certainly not the main battleground between proponents of the nativist approach and the constructivist approach to language acquisition. At the heart of the debate is the level of abstractness that should be ascribed to children's early linguistic representations. Thorough discussions are found for example in Tomasello (2000); Fisher (2002); Lidz, Gleitman, and Gleitman (2003); Goldberg (2004); Naigles (2004).

[8] The rule presented in 2.6 is originally suggested by Johnson et al. (2005). See additionally the discussion of their experiment in Chapter 5.2.1.

[9] Under the 'encapsulated' model, production processes are assumed to be separate into distinct processing stages that begin with the formulation of a conceptual representation. Then grammatical encoding takes place in which each lexical concept is mapped onto abstract syntactic representations. This is where agreement processes are thought to take place, thus agreement is

this case, the child would know that the morphological form of the verb depends on the morphosyntactic features of the sentence subject, thus the child would have adult-like knowledge of the subject-verb agreement relation. If a child produced correctly inflected verbs, this assumption would further entail that the child knows the phonological realization of the morphological form (i. e. an English-speaking child would know that 3^{rd} person singular is phonologically marked with an *-s* affixed to the lexical verb stem). Finally, the child would have to be granted with morphological knowledge of the subjects number features. Theoretically, a rule like 2.5 should be applicable in production and comprehension processes alike.

If one on the other hand assumes that a young child has internalized a rule that rather relies on surface properties of the input instead of abstract grammatical features, a rule as stated in 2.6 seems more likely. This rule simply tells the child to add an *-s* when she produces a verb following either the pronouns *he, she* or *it* or when she produces a (subject) noun that does not have a (plural) marking.[10] Such a rule certainly differs from the adult one and leaves the child with many unclear cases (e. g. what do to with irregular noun plurals? What to do with a word like *horse* or *bus*, which does end with an *-s*?). But such a rule would still enable a child to produce many correctly inflected verbs in her own speech. Most important for the present work is that the latter rule does not allow a child to infer the number value of the subject based on the verb form, i. e. this rule does not help the child in semantic comprehension. The rule stated in 2.6 is devoid of meaning in the sense that is does not relate to the number value of the subject at all. To connect this kind of rule with subject number, an additional rule is required which draws the link between the phonological form of

viewed as a primarily syntactic phenomenon (morphosyntactic features are copied or co-indexed (e. g. Bock, Eberhard, Cutting, and Meyer, 2001; Thornton and MacDonald, 2003).

[10] It might be the case that the child just attends to the noun that is closest to the verb, assuming that this is the NP which the rule refers to. Haskell and MacDonald (2005) found in corpus analyses of adult language that the closest noun to the verb is typically the subject.

the nominal with semantic number (e. g. N+/Ø/ $\rightarrow$ N+/s/ iff N denotes more than one.[11]). A rule as stated in 2.6 would suffice to enable a child to distinguish grammatical from ungrammatical sentences. As a side note, this kind of rule credits the child with more sophisticated linguistic knowledge than some proponents of the constructivist approach assume for the early stages of language development, because they assume that the child does not even have the notion of V in those early stages (e. g. Tomasello, 2000). Early language production is rather thought to be entirely based on the recall of memorized chunks that have been encountered input. But such a rule could well be what a child extracted from the first memorized constructions, in the form of "child's grammatical knowledge comprise[d] of general but pared down constructions" (Wilson, 2003, pg. 85). Such a rule precisely reflects a step on the child's way from item-based constructions to abstract grammatical knowledge according to constructivist accounts.

Importantly, the two rules stated above make different predictions with regard to the production and comprehension of verb inflections.[12] If a child has internalized a rule like 2.5 from early on, she should in theory do the following: (1) regarding sentence production, all verbs (lexical verbs, auxiliaries, modals, copula) should be correctly inflected (if the child 'knows' the number value of the sentence subject), (2) regarding sensitivity to agreement violations, all sentences containing 3^{rd} person singular subjects should be 'judged' correctly (if only 3^{rd} person singular and 3^{rd} person plural subjects and verbs are presented, but children might be correct on the plural sentences as well, simply because this would be the complementary case) and (3) regarding comprehension of verb inflections, children should be able to infer that a 3^{rd} person singular inflected verb

[11]Such a rule would again lead to predictions that children overgeneralize the regular noun plural *s* to irregular nouns and produce nouns such as 'the deers' or 'mans'. Such cases are certainly documented in child speech (e. g. Marchman, Plunkett, and Goodman, 1997; Marcus, 1995; Zapf and Smith, 2007).

[12]The following predictions apply for English. For another language, like German, other predictions, especially with regard to noun plural would have to the posited.

refers to a singular subject, and therefore to a singular referent (again similar rates of comprehension should be found for lexical verbs, auxiliaries, modals and copula). Any deviance from these predictions found in children's language use needs to be explained by performance factors. Children can show frequency and familiarity effects (which are easily explained by performance factors), but their ability to produce or comprehend verb inflection should not be influenced by the phonological structure of the subject noun, such that e. g. nouns that end of an *-s* are harder to process.

If a child has build up a rule like 2.6 at some point in development, she should rather show the following pattern within and across modalities: (1) regarding sentence production, main verbs should carry the *-s* affix if they are preceded by a noun that does end in *-s* or by a 3^{rd} person singular pronoun. For auxiliaries, copula etc., additional rules would need to be posited (such as '*be* $\rightarrow$ *is* following he, she, it, NP+/Ø/'), (2) regarding sensitivity to agreement violations, sentences should be 'judged' adequately if they contain a 3^{rd} person singular pronoun or an NP that does not end in *-s* and if verbs are marked by an *-s* (as correct) or lack an *-s* inflection (as incorrect). Irregular nouns (with regard to plural) or other pronouns should be difficult for children, unless they have build up additional rules tacking these nominals. (3) regarding comprehension of verb inflections, children should not be able to use verb inflections as a cue to subject number because this rule does not make any predictions about meaning. To infer meaning, children need additional rules which state that NPs with *-s* affix refer to plural entities. Still, even with a combination of such rules, children should not be able to infer the number of irregular plural or nouns that end with *-s* (like 'horse' or 'bus').

In the following, theories that have either been influential in the discussion of non-finite verbs in children's root sentences or that are considered relevant for the current thesis will be presented.[13] Wexler, as a proponent of the nativist view, pro-

[13] There are certainly more accounts that have been proposed to explain the phenomenon of co-occurence of finite and non-fininte verbs in children's early

posed the existence of an 'Optional Infinitive-Stage' (OI-stage), in which non-finite forms are allowed in young children's sentences because Tense (and/or Agreement) features can optionally be left underspecified, due to the Unique Checking Constraint (UCC), in child grammar. This is not possible in adult grammar, as all sentences have to be finite (Wexler, 1998). According to this grammatical model, an unchecked feature has to be eliminated, thereby producing the co-occurence pattern of finite and non-finite verb forms found in young children's language production. This theory takes a maturational view on language development, as it assumes that the UCC and thus OIs disappear because the child's grammatical systems itself matures (Wexler, 1998; Ionin and Wexler, 2002). Still, the theory attributes 'Full Competence' (Poeppel and Wexler, 1993) to the child. Wexler's theory has had a strong influence on the study of non-finite verbs in children's sentences. Ingham (1998) suggested an account in which root infinitives are explained by the absence of a subject agreement projection (AgrSP).[14] The theory of a missing AgrSP follows work by Clahsen and colleagues (e. g. Clahsen and Penke, 1992; Clahsen et al., 1993), who first argued that children go through a period in which a finiteness projection higher than VP exists, which only holds finite forms that do not require an analysis in terms of agreement.

Phillips (1995) proposed that the use of non-finite verb forms is not due to missing grammatical competence, but due to processing difficulties. The assumption of processing difficul-

sentences, see for example Ferdinand (e. g. 1996). A very early theory (Gleitman and Wanner, 1982) suspected that the phonological properties of functional elements (being short and less heavily stressed) prevented children from recognizing those elements in the input. This again was regarded as underlying reason for children's omission of functional elements in production. But further research has shown that young children are indeed sensitive to the presence of function words, even though they do not always produce them themselves (e. g. Shady, 1996; Gerken, 2002; Gerken and McIntosh, 1993; Gerken, Landau, and Remez, 1990; Höhle, Weissenborn, Kiefer, Schulz, and Schmitz, 2004; Höhle and Weissenborn, 2003).

[14] Ingham (1998) proposed that all children go through a stage in which they mark their sentences for Tense but fail to show agreement between the verb/modal and the sentence subject, although at some point he argues that this phase might follow the OI-stage proposed by Wexler.

ties is incorporated in a Principles-and-Parameter framework. Phillips assumed that overt V-I movement, moving the verb from the VP to the functional head INFL, is required to realize verb inflection.[15] Accessing the inflectional paradigm in sentence production (under the assumption of non-lexicalist morphology) is thought to have a cost to it, as does the V-I movement itself, but V-I movement is additionally thought to facilitate spell out and has thus a compensatory advantage. For adults, it is assumed that accessing the inflectional paradigm is a highly automatic process with only very little cost. Therefore the advantage of spelling-out the inflectional form outweighs the costs of accessing it. For children, accessing the morphological form is, at least in the beginning, considered to be a not-yet-automized process. Therefore, the cost for accessing the inflectional form may outweigh the advantage of spelling it out in children. This lack of enough processing capacities to access the form of the verb inflection is though to give rise to sentences containing root infinitives. It is important to note that Phillips takes a very dynamic view on the 'cost' of a syntactic operation (contrary to e. g. Theakston et al. (2003) who attribute processing costs mostly to sentence length).[16] Nevertheless, this theoretical account assumes that children possess adult-like knowledge of agreement morphology, but might not be able to access it in every obligatory context.

Thus, all the above mentioned nativist theories explaining root infinitives still assume that children have adult-like rules to construct subject-verb agreement and use verb inflections, similar to the claim made by Hyams (1999). A different approach is taken by proponents of the constructivist view. Here, the basic assumption is that children's early linguistic representations have the form of memorized chunks (constructions), and that the early use of 'rules' is tied to specific lexical item (item-based). Only very limited productive use of linguistic

[15] This applies at least to languages in which inflectional features have to be attached to the verb, like French and German.

[16] Phillips (1995) theory is based on a non-lexicalist approach to morphosyntax, where verbs and inflectional features enter the syntactic operations in language production as independent syntactic elements.

rules is supposed to be found in child language. Further, all patterns in child speech are supposed to mirror the input children receive, especially with regard to input frequency (e. g. Tomasello, 2000; Childres and Tomasello, 2001; Abbot-Smith and Tomasello, 2006; Wilson, 2003).

The question of 'productivity' in early language acquisition has been applied to children's early use of verb inflections (e. g. Theakston et al., 2003; Wilson, 2003). Most experimental studies define productivity as the ability to add an inflection to a novel word stem, or to drop an inflectional affix from a newly learned word (Hohenstein and Akhtar, 2007).[17] Other experimental studies have examined the co-occurence patterns of verbs, auxiliaries or copulas and specific subjects, e. g. pronouns, in spontaneous speech (e. g. Wilson, 2003) or calculated the number of morphological forms for each verb in the child's productive use (e. g. Pizutto and Caselli, 1994). The constructivist approach explains co-occurence pattern of finite and non-finite verbs in children's sentence production in the second and third year of life with the input children receive. Since they hear declarative sentences like *He jumps.* and questions like *Can he jump?*, they might be confused about whether or not a 3^{rd} person singular marker is required in English or not (Theakston et al., 2003). This contrasts with the nativist view that even the earliest use of 3^{rd} person singular forms reflect abstract knowledge of tense and agreement. According to the constructivist view, children first learn various instances of inflected verbs, in the form of lexically based constructions, and only later in their linguistic development generalize from this set of data they have accumulated (Theakston et al., 2003).

Interestingly, most studies investigating verb inflection rely on production data. Comprehension data, as presented in this thesis, might shed a different and hopefully clarifying light on the question of children's early knowledge of verb inflection and regarding the development of subject-verb agreement knowledge. As Fisher (2002) states, "[a]nother way to assess

[17]Most studies have focused on the present progressive affix *-ing* or on past tense *-ed* (Hohenstein and Akhtar, 2007).

abstract knowledge is through comprehension tasks." (pg. 271). But to set the stage, I will first turn to production studies, focussing on the developmental time course of the emergence of verb inflection in child language

3 Production of verb inflection

In this chapter, I will review various studies that investigated children's use of verb inflection in sentence production. This is done separately for studies examining English-, German- and Spanish/Dutch-speaking children. The overview is restricted to these languages because these are the ones for which comprehension data exist. The main focus of this production overview lies on the time course of children's use of verb inflections. One major concern regards the methodologies used and how they give rise to varying rates of development.

3.1 Methods for testing production

The majority of studies on the acquisition of verb inflection have examined children's use of these functional morphemes in the productive modality.[18] Researchers who aimed to examine children's productions of 3^{rd} person singular and plural inflections either investigated children's spontaneous speech at vari-

[18]Many studies have tested children's production and comprehension of the English past tense affix *-ed* or the English present progressive *-ing* (e. g. Akhtar and Tomasello, 1997; Hohenstein and Akhtar, 2007), mainly to gain insight about children's syntactic and morphological productivity, i. e. to find out whether and when children have a general understanding of the role verb inflectional affixes play in language. Other studies have included the production of copula and auxiliary forms of *BE* and the auxiliary *DO*, since these verbs mark finiteness and especially *BE* displays a rather elaborate morphological paradigm, in contrast to the other verbs in the English language (Rice et al., 1995). Since number contrasts on lexical verbs (and therefore the 3^{rd} person singular and 3^{rd} person plural inflectional affixes) are in the focus of this thesis, research on other verb inflectional affixes (e. g. past tense, present progressive) and other verbs (e. g. *BE* and *DO*) will only be discussed marginally.

ous ages (Brown, 1973; Clahsen, 1986) or they employed elicited production tasks (Berko, 1958; Rice and Wexler, 2001).[19]

When spontaneous speech data from children is analyzed for the presence or absence of grammatical morphemes, it is critical to verify whether the syntactic context actually calls for the use of a certain grammatical morpheme (e. g. Cazden, 1968; Stromswold, 1996). If the syntactic context is unambiguous and it is clear that adults would always use a certain grammatical morpheme in the position in question, this is called an *obligatory context*. For example, a child would only need to add the 3^{rd} person singular inflection to a main verb (adding an *-s* to the stem), if she was producing a sentence with a 3^{rd} person singular subject. Thus, if one would find the utterance 'the dog bite' in a spontaneous speech sample, one could clearly state that the verb was not correctly inflected in this case. If one on the other hand would not find sentences containing such 3^{rd} person singular subjects (but for example only 1^{st} and 2^{nd} person subjects, which can happen in conversational speech) in a spontaneous speech sample, one could not argue that the child in question does not produce the critical grammatical morpheme. Importantly, there should always be a critical mass of obligatory contexts in children's spontaneous speech samples to draw justified conclusions (Balason and Dollaghan, 2002). One important advantage of elicited production studies is that the experimenter has better control over the critical aspects just mentioned (e. g. Thornton, 1996).

Especially those studies that used the presence and position of RIs in children's early sentences to argue for the presence or the absence of a certain structure of functional categories in child grammar relied almost solely on spontaneous speech production(e. g. Poeppel and Wexler, 1993; Ingham, 1998; Clahsen et al., 1993). The data from such studies will be examined to provide information about the rate and correctness of verb inflec-

[19] English studies have usually only focused on the production of 3^{rd} person singular inflection *-s*, because this is the only overt affix in the present tense inflection paradigm, and because in some studies the acquisition path of the *-s* verb affix has been contrasted with the acquisition path of the homophonous *-s* noun plural affix.

tion in children's language production. The reasoning works as follows: all RI-studies cited report the percentage of finite and non-finite verbs children produced at a certain point in development. Finite verbs are reported to be almost always used with correct inflection. Further, researchers usually only included those instances in which they were sure about the sentence subject, which was mostly singular, or in which they could clearly distinguish whether a bare stem or an inflected verb was used (which results in the predominant research on 3^{rd} person singular *-s* in English). Therefore, we take the reported percentages of (correctly inflected) finite verbs as an indicator for the rate of children's use of verb inflection morphology at various ages. This is done to determine a point of 'productive mastery'. Additionally, data from elicited production studies will be presented to broaden and complete the picture. Critically, spontaneous speech and elicited production differ in their linguistic and non-linguistic demands. First, spontaneous speech allows the child to pick and choose the lexical items he or she is familiar with, while elicited production tasks force the child to work with whatever she gets presented. Thus, frequency and familiarity of the verbs (not to mention the use of pseudo verbs in elicited production tasks) can differ significantly across these two tasks. Second, spontaneous production transcripts do not allow to control for the semantic aspect of language production. With regard to the number contrast between singular and plural, one could imagine that a child had more than one entity of whatever she is talking about in mind, but because she was not able to access the plural noun and/or the plural verb, she simply produced a singular sentence. This sentence might well be correct with regard to subject-verb agreement and the use of inflection, but it might not convey what the child actually intended. In an elicited production task that examined the number contrast, such a singular inflected sentence in the context of e. g. a two-actor picture would be regarded as error. These differences should be kept in mind when data from spontaneous speech and elicited production are compared.

Since no separate production experiment was conducted as part of the present study, the already published data on children's verb inflection production needs to serve as a basis for comparison when the later comprehension experiments are discussed. Additionally, we can call on the the data gathered in elicited production experiments. The main findings for English and German verb inflection will be presented in the following.

3.2 Early verb inflection in English

In his longitudinal study of the spontaneous speech of three English-speaking children, Brown (1973) investigated the acquisition of 14 grammatical morphemes relative to the mean length of utterance (MLU) and the developmental stage of the children. He found correct production of the 3^{rd} person singular *-s* in 90 % of obligatory contexts at the age of 2;3 (years; months) for Eve, at 3;6 for Adam and at 3;8 for Sarah. This age range already indicates substantial variance across children even within one language. In this study, Brown additionally defined a widely accepted, although of course arbitrary, criterion for the 'acquisition' of a morpheme in the sense of productive mastery. According to this criterion, a morpheme has to be present in 90 % of all obligatory contexts in three successive speech samples from a particular child.

De Villiers and de Villiers (1973) pursued a cross-sectional approach to investigate the acquisition of English grammatical morphemes. They analyzed spontaneous speech samples of 21 children, aged 1;4 to 3;4 years of age, and searched for the productive use of the same 14 grammatical morphemes that Brown analyzed.[20] For the production of the 3^{rd} person singular *-s*, 14 children provided enough obligatory contexts to be included in the analysis. Results showed that the use of the 3^{rd} person singular *-s* ranged from 0 % to 100 % with a mean of 44.7 % at the time point of testing. Because the age and the MLUs of the children in this study varied considerably, but the amount

[20] Only transcripts that provided five or more obligatory contexts for a given morpheme were used for that morpheme.

of correct use of this grammatical morpheme is critical for the present study, we list the percentages found by de Villiers and de Villiers (1973a) in Table 3.1.

Initial	MLU	Age (months)	3^{rd} person singular *-s* (reg.)
J.	1.44	21	12.5
J.	1.58	19.5	0
R.	2.04	26	33.3
H.	2.08	26	0
C.	2.24	30	20
C.	2.31	28	0
J.	2.45	31	75
K.	2.79	30	55.5
H.	2.99	21	14.3
E.	3.03.	33	42.8
A.	3.16	28.5	88.5
G.	4.23	29.5	94.3
M.	4.29	36	90
M.	4.67	40	100

Table 3.1: Percentage of obligatory contexts in which 3^{rd} person singular *-s* morpheme was used correctly, taken from de Villiers and de Villiers (1973a).

It can be seen that the percentage of correct use of the English 3^{rd} person singular verb inflection is rather variable across children. While one child aged 28 months produced no correctly inflected verb at all, there is another one at almost the same age (28.5 months), who inflected the verbs correctly to almost 90 % (critically, the MLUs of these children differ as well). It is further interesting to note that there are only four children who add the inflectional morpheme *-s* more than two thirds of the time. According to the data from de Villiers and de Villiers (1973a), English learning children at the age of three years have still not truly mastered the regular 3^{rd} person singular inflection.

Berko (1958) was the first one to elicited grammatical morphemes from English learning children in an extensive elicited production task. Critically, in order to not only tap grammatical forms that might have been learned via rote-memory and that might be stored as holistic forms, Berko used neologistic word forms (and came to fame as the mother of the 'Wug test'). Among other morphemes, she was interested in the production of the 3^{rd} person singular *-s*. To elicit agreement morphemes, she introduced two new verbs (*loodge* and *naz*), once as an infinitive ('This man knows how to naz') and once as a present progressive form ('Look, he is nazzing'). Then children were presented with a gap sentence that was supposed to elicit the correctly inflected present tense form ('Every day the man...').[21] Preschoolers aged four to five years (N=33) correctly inflected the item *loodges* 57 % of the time and the item *nazzes* 47 %. First graders aged five and a half to seven years (N=61) inflected *loodges* to 56 % correctly and *nazzes* to 49 % correctly.

These rates of correct inflection are much lower than the ones found in the spontaneous speech studies reported earlier. The differences might be due to general cognitive demands, verb selection (familiarity and frequency) and/or 'semantic demands' that differ in both tasks. This will be discussed at the end of this chapter. The lower rates of correctly inflected verbs cannot be explained with appeal to possible phonological problems the children tested in this sample might have had, because the children had no problems adding *-s* to nouns to inflect those for plural. Further, there was no overall problem with verb inflection, because children added the present progressive morpheme *-ing* in 90 % of the contexts in which they were asked to do so.

Rice et al. (1995) used both methodological approaches, i. e. analysis of spontaneous speech and elicited production samples, to investigate the OI-stage in English-speaking SLI-children, children matched for chronological age (CA), and children matched for mean length of utterance (MLU). Of main interest for the present study are the results of the control children, since they provide a rough estimate of when the OI-stage in English-learning children ends. Thus it tells us when those children can be credited with target-like use of 3^{rd} person singular verb inflection in the productive modality.[22]

[21] Note that this elicited production task does not contrast the number feature of the verb in form in model and test, but rather focuses on aspect.

[22] The authors investigated the use of the 3^{rd} person singular *-s* morpheme as a

Method	SLI	three year olds	five year olds
spontaneous speech*	34 % (*SD*=26 %)	51 % (*SD*=40 %)	85 % (*SD*=20 %)
elicited production	26 % (*SD*=23 %)	38 % (*SD*=41 %)	95 % (*SD*=13 %)

Table 3.2: Grammatical morpheme 3^{rd} person singular *-s*: percentage correct (*in obligatory contexts), taken from Rice et al. 1995.

Table 3.2 shows the mean percentage of 3^{rd} person singular *-s* use in the spontaneous speech and in elicited production probes.[23] Present tense singular *-s* was elicited showing pictures to the children depicting a person in a particular occupation. This was accompanied by the experimenter's comments, e. g. 'I am a teacher and I teach. This here is a fire fighter and he...' Interestingly, the three year olds (who still outperformed the five year old SLI-children, at least considering the numerical means) produced the 3^{rd} person singular *-s* inflections in only half of the obligatory contexts in spontaneous speech and even less in the elicited sentences. The five year olds on the other hand produced the agreement morpheme to at least 85 %. Thus, the 20 three year old children produced significantly less verb inflections in the spontaneous speech samples and in the elicited production probes than the 22 five-year olds did. One can conclude that there has to be a transition or a shift from the age of three to the age of five concerning the use of verb inflection morphology in English-speaking children, at least in those children without SLI. An important additional finding was that all of the children's errors were omissions, leaving the bare stem (**he walk*). Children never added an incorrect inflectional affix.

marker for tense, since in the OI-theory, tense is thought to be the optional feature in child grammar. Additionally, they investigated the use of the past tense affix *-ed* and the use of *BE* and *DO*, arguing that all these four English morphemes mark tense and therefore finiteness.

[23]Children needed to produce at least three obligatory contexts to be included in the spontaneous speech analysis.

Rice, Wexler, and Hershberger (1998) performed a longitudinal study with normally developing and SLI-children and thereby replicated their findings. The normally developing children mastered 3^{rd} person singular *-s* around the age of four, irrespective of the task used to measure the production of verb inflection (spontaneous speech and elicited production). Similarly, but using only elicited production data, Rice and Wexler (2002) found that normally developing English learners on a group level did not reach the 90 % criterion of correct use of the 3^{rd} person singular *-s* before the age of four.

Theakston et al. (2003) finally elicited 3^{rd} person singular *-s* in English-speaking children as well. Critically, the children (aged 2;6 to 3;0) first heard the known and novel verbs either in declarative (finite) or question (non-finite) form and were then asked to produce a declarative sentence. This was done to investigate the influence of the input on children's productive use of verb inflections. For the known verbs, no influence of input was found, as all of these were correctly inflected about 70 % of the time. For the novel verbs, an influence of input on the production of correctly inflected finite verbs and non-finite verbs was found. Children produced more non-finite forms when they had heard the novel verbs in questions, and thus in the non-finite form, than when they had heard them in declaratives. The authors take this pattern as evidence for a constructivist, item-specific view on morphological development in children. Others argue that a demonstration of input effects on the rate of non-fininte verbs is no compelling evidence that input is actually the cause of the OI-phenomenon (Soderstrom, 2002).

The conclusions made by Theakston et al. (2003) are corroborated by a spontaneous speech examination carried out by Wilson (2003). He investigated longitudinal transcripts of five children acquiring English (aged 1;6 to 3;5) to determine lexical variety of inflectional use and whether all elements marking inflection (i. e. copula *be*, auxiliary *be* and 3^{rd} person singular) would develop in parallel as proposed by Rice et al. (1998). He finds the latter mentioned claim not to be true, since all children included in the study showed considerable variation in their use of the three different elements marking inflection. Further he found very little lexical variation and claimed that "the evidence that many if not all early uses of inflection are tied to lexically specific frames shows that the production of correctly agreeing forms cannot be taken as evidence for the child knowing the relevant morphemes and principles of agreement" (Wilson, 2003, pg. 111). Pine, Lieven, and Rowland (1998) as well found that children showed a lot of 'lexical specificity'. The number of verb types that were inflected for 3^{rd} person singular was typically very small within each of the 12 children (aged 1;4 to 2;7) under investigation, suggesting that the use of the grammatical morpheme was limited to a handful (maximally 8) of unanalyzed forms.

Summarizing, English-speaking children start to produce correctly inflected verbs as early as two years of age, but it takes children a few years to show productive mastery in all kinds of tasks. In tasks that put more demands on children's language processing and that require them to inflect verbs they have not themselves chosen, adult-like performance appears not before the age of four or five. At the age of three, the rate of verb inflection production seems to be very variable. These facts can so far be explained by both kinds of rules posited to underly children's processing of verb inflections (see 2.5 and 2.6). If one assumes an adult-like rule to be in place, performance factors that affect elicited production more than spontaneous speech need to be considered. A non adult-like rule could explain task-related differences in a straight-forward way (i. e. without appeal to particularities of performance mechanisms): if children produced verbs along with pronouns or familiar nouns, they can easily add the *-s* to the verb if necessary. If they are asked to produce a specific (inflected) verb form in relation to semantic information (number of the sentence subject or aspect or the action), they fail or they have to guess whether the verb inflection *-s* should be added to the verb stem.

3.3 Early verb inflection in German

The German verb inflection paradigm is richer than the English one. See Table 3.3 for an example of the regular present tense indicative inflectional affixes that mark person and number in German lexical verbs, taken from Bittner (2003c).[24] Studies on the acquisition of German verb morphology have often compared the time course of acquisition of the single inflectional forms (Bittner, 2003c; Clahsen, 1986). Most studies report the following time course: *-en* appears earlier than *-t* and *-Ø* (which represents the bare stem), *-st* usually appears last (e. g. Bittner, 2003c).

This relative time course of the appearance of verb inflectional forms is not very informative regarding the time point of acquisition of the morphological paradigm in general, whether children are using verb inflectional affixes in a productive way or not and what kind of rule might be underlying their productions.[25] So far, almost all studies on the acquisition of verb inflections by German-speaking children are based on spontaneous speech data, except the data gathered by Ott (2011). However, as we only tested comprehension in the present study, these data have to serve as a general benchmark for the production of inflections by German children.

Clahsen (1986) investigated the productive use of verb inflections in two German-learning children aged 1;6 to 3;6 by conducting frequency analyses of all present tense inflectional mor-

[24] German verbs are further characterized as belonging to the strong or weak class. Strong verbs have a stem-vowel change when inflected (e. g. *seh-en*, 'to see' vs. *er sieh-t*, 'he sees') and weak verbs are those without stem vowel change (e. g. *sag-en*, 'to say' vs. *er sag-t*, 'he says'). In all experiments presented in this thesis, only weak verbs are used. The difference between strong and weak verbs is not considered relevant in this work, because the inflectional affixes are identical in both cases, at least in present tense.

[25] Some analyses even collapse all uses of *-en*, even though this verbal affix marks 1st and 3rd person plural *and* the infinitival form (Bittner, 2003b). This is a good example for challenges that arise when analysing children's spontaneous speech and why it is important to establish 'obligatory contexts of use'. Analyzing verbs solely according to the phonological form of the inflectional affix reveals no conclusive information about the acquisition of verb inflection.

	Singular	Plural
1^{st} person	mach-**e**/-**Ø**	mach-**en**
2^{nd} person	mach-**st**	mach-**t**
3^{rd} person	mach-**t**	mach-**en**

Table 3.3: Regular German present tense person-number inflection, taken from Bittner 2003.

phemes at various ages. Correct use, in Brown's sense of more than 90 % correct, for the 3^{rd} person singular inflectional morpheme *-t* did not occur before the age of 2;11. For the 3^{rd} person plural inflection *-n*, correct use emerged even slightly later, at around 3;1.

Clahsen and Penke (1992) analyzed the spontaneous speech of one German-learning child, Simone, from the age of 1;7 to 2;8. For the 3^{rd} person singular inflection *-t*, they found it used in 83 % of the obligatory contexts, thus this child only produced non-finite verbs in main clauses about 17 % of the time. As usual, agreement errors were almost absent in the corpus, since appropriate use was found in 98 % of the cases. Only very rarely did the child use a plural subject with a singular inflected verb. The usual errors were omissions of verb inflection.

Further, a causal relationship between the morphological acquisition of the verb inflectional paradigm and the syntactic development of verb movement (which is obligatory in German main clauses) has been claimed by Clahsen and colleagues. According to the *lexical learning hypothesis*, German-learning children first learn the complete verb inflection paradigm (including the last acquired form *-st*, which marks 2^{nd} person singular) and this triggers the development of the full syntactic clause including all functional categories (IP and CP). This process is described as 'morphological bootstrapping' (Clahsen, Eisenbeiss, and Penke, 1996). Verrips and Weissenborn (1992) on the other hand proposed that acquisition of finiteness and verb placement does not depend on the acquisition of subject-verb agreement morphology. This view is stated in their *independence hypothesis*, which claims that "finiteness and verb placement in German develop independently of subject verb agreement morphology, more specifically, that V-to-I-to-C movement develops before the paradigm of subject-verb agreement has been completely mastered" (Verrips and Weissenborn, 1992, pg. 286). Interestingly, 'mastery of subject-verb agreement' is usually only defined in terms of children's use of verb inflections in spontaneous speech. In the nativist tradition, an inflectional form is regarded as 'acquired' as soon as it is apparent in a child's spontaneous speech. No further claims regarding productivity or 'form of the rule' are made, it is rather implicitly assumed that the child acquires the adult-like rule from the start.

In another study, Poeppel and Wexler (1993) analyzed the spontaneous speech of Andreas, a German-learning child aged 2;1. They found that in all utterances that contained a 3^{rd} person singular subject and a finite verb, the verb was correctly inflected with the *-t* marker. This pattern again reflects the over and over repeated finding that when children inflect verbs, they do it correctly. When the data are checked for the amount of inflected, thus finite verbs, opposed to non-finite verbs, we find 82 % of the verbs correctly inflected and 18 % of the verbs in their non-finite form.[26] Interestingly, all agreement errors consisted of a plural subject occurring with a 3^{rd} person singular inflected verb (e. g. 'Alle Tiere liegt da', *All animals lies there*), thus the number agreement between subject and verb was violated.[27] These patterns lead the authors to conclude that the

[26]The authors note that the child predominantly used singular subjects with correct agreement morphology, and that 2^{nd} person singular subjects are very rare, but they do not provide precise numbers on how many of the child's sentences contained a 1^{st} person singular subject versus a 3^{rd} person singular subject.

[27]The example provided by Poeppel and Wexler (1993) is an interesting one, because it shows how difficult it can be to infer the child's intention when analyzing spontaneous speech. One can hypothesize that the child actually intended to to say something like *Jedes Tier liegt da* 'Every animal lies there'. This noun phrase with a different quantifier ('every' instead of 'all') is nearly synonymous, but is grammatically singular. If Andreas actually intended to say 'Jedes Tier...', he might have selected the correct verb inflection. Of course, subject and verb did not agree in the sentence he uttered, but this might be due to 'wrong' noun phrase selection or Andreas' assumption that 'alle X' refers to a singular entity. Then, the agreement error not due to incorrect verb inflection. Recent research has shown that the acquisition of quantifiers like 'some' is challenging for three year olds (e. g. Hurewitz,

agreement system is basically available at age of two years (although they acknowledge that the agreement paradigm is not fully available, because there are some number mismatches found.)

The discrepancies between Poeppel and Wexler's findings (showing that Andreas was produced most of the verb inflections correctly) and those of Clahsen (1986) can be accounted for by differences in the way the data were analyzed: while Clahsen considered utterances containing a non-finite verb and an overt subject as agreement violations, Poeppel and Wexler restricted their analysis to utterances containing finite and thus overtly inflected verb forms. It should nevertheless be kept in mind that spontaneous speech data might overestimate children's production abilities to inflect verbs correctly, due to high familiarity and frequency of the verbs that children usually produce. Further, it is not possible to estimate how many of the inflected verb forms the child might have memorized as full constructions. The authors of both studies emphasize that frozen expressions like *Was macht der?* 'What is he doing?' were removed from analysis, but it is in my view very difficult to decide which verb forms should be counted as frozen expressions and which should be regarded as 'created in a productive way'.

Finally, Bittner (2003a) provided analyses of the spontaneous speech of one child learning German, Anna, recorded between 1;8 and 2;1. All finite verbs the child was using were 3^{rd} person singular present tense lexical verbs. The proportion of self-produced finite forms out of all self-produced forms varies from 25 % at the beginning of the recordings to around 70 % at the end of recordings. The author noted that the repertoire of lexical verbs was rather limited at the beginning, but expanded considerably until 2;1.

Clearly, problems can arise when findings from case studies, which are based on the data of one or two children, are used to serve as a basis for generalization (e. g. Blom, 2003). Clahsen et al. (1993) analyzed that data of 7 German-learning chil-

Papafragou, Gleitman, and Gelman, 2006). Poeppel and Wexler (1993) raise this possible interpretation as well (footnote 15).

dren, aged 1;8 to 2;9, to investigate functional categories in child grammar. Since the focus of the research paper was on the presence and number of functional categories in German children's initial grammar, verb inflection was always set in relation to verb placement. Nevertheless, the authors reported a certain amount of sentences with incorrect subject-agreement marking, in which the verb occurred at the first or second position of the sentence V1/V2 (e. g. 'du schreibt das', *you*-2SG *write*-3SG *this*). The proportion of incorrect agreement in V1/V2-position (of all verb patterns) ranges from 0 to 64 %. Clahsen et al. (1993) took their data as providing evidence that only one functional projection exists in early child grammar, which only marks finiteness, but not agreement. They further claimed that children at this stage in development have not yet acquired the complete subject-verb paradigm.

Ott (2011) was to the best of my knowledge the first to publish elicited production data on verb inflections from German-learning children. The main aim of this experiment was to investigate whether frequency of sub-syllables has an influence on the production of *-t* verb inflections to mark 3^{rd} person singular in German. Thus, she compared the rates of correct verb inflection in high frequent sub-syllables containing a short vowel $VCt]_\sigma$ (e. g. *fäll-t*, 'fall-3SG') to those in low frequent sub-syllabes containing a long vowel $VVCt]_\sigma$ (e. g. *fehl-t*, 'miss-3SG') in German-speaking SLI-children (N=16, mean age = 4;8) as well as CA-matched (N=16, mean age = 4;8) and language-matched (LM) controls (N=14, mean age = 3;5). Critically, pseudo verbs were used to (1) prevent children from producing forms that might be stored holistically in the lexicon and (2) to have better control of the phonological properties of the verbal material. Children were presented with pictures of multiple persons performing the same unknown action. This was accompanied by the introduction of the pseudo verb, always presented in the 3^{rd} person plural form (*Schau mal, sie telen.*, 'Look, they teel'). Then children were shown the picture of one person performing the just introduced action and they were asked to complete a sentence like 'There is another girl. What is she doing? She...'. Of main interest for the present study are the rates of correct inflections (i. e. adding the verb inflection *-t* to the pseudo verb) that children reached in this task. Overall, CA-matched children aged 4;8 produced 76 % correctly inflected verbs, without showing an effect of frequency of the sub-syllable. LM-matched children did not show an effect of the frequency of the sub-syllable either, but overall only inflected 26 % of the pseudo verbs correctly. SLI-children finally inflected 41 % of the verbs correctly, and additionally showed a significant effect of sub-syllabic frequency. They were much better at inflecting verbs with a high frequent short vowel sub-syllable as opposed to verbs with a low frequent long vowel sub-syllable, as predicted by Ott (2011), following Marshall and van der Lely (2006).

Since Ott's study focused on the interface of phonology and morpho-syntax, she classified instances in which children added the inflectional allomorph *-et* to the verb stem as mistakes.[28] If we pay less attention to the phonological specifications, but count the instances of *-et*-affixation as correct inflections according to the morphological paradigm of German, the rates of correct inflection rise for all groups. Especially the youngest children substituted the inflectional allomorph *-et* for the expected inflectional affix *-t* very often, namely in 30 % of the cases. Simply adding these allomorph instances to the correctly inflected counts raises the percentage of correctly inflected verbs up to 56 % for the LM-matched children (mean age 3;5). For the older children (mean age 4;8), the inclusion of the allomorph inflected verbs (14.5 %) raises the percentage of the correctly inflected verbs up to almost 90 %.

Most relevant for the current thesis are the findings that three and a half year old children only inflect about half of the pseudo verbs (and only about one quarter when strict phonological criteria are applied) and four and a half year old normally developing children inflect almost 90 % of pseudo verbs correctly (but only about 75 % when strict phonological criteria are ap-

[28] *-et* is an inflectional allomorph of the 3^{rd} person singular inflection *-t*, which occurs on verb stems that end on the alveolar plosives /t/ or /d/. According to Wiese (2000), the epenthesis of schwa is strictly related to the application of a morphological rule.

plied). The German data is very similar to the English data gathered by Rice et al. (1995), showing a transition in normally developing children from about three years of age to about four and a half to five years of age. Productive mastery of 3^{rd} person singular verb inflection does not seem to be in place before the age of four and a half, at least when the verbs are elicited.

Summing up, spontaneous speech passages gathered from German-speaking children seem to indicate rather early knowledge of subject-verb agreement and verb inflection morphology. Poeppel and Wexler (1993) conclude that "he [Andreas 2;1] basically knows the agreement system" (pg. 5). Clahsen, who claims that the agreement system evolves slightly later, also assumes that the knowledge is in place when instances of agreement production can be found in children's spontaneous speech (Clahsen, 1986). This seems to be evidence for the claim that children have acquired an adult-like rule early in development. No my best knowledge, no study on German inflection has yet analyzed lexical variation with regard to early verb inflections. Recent findings from an elicited production task broaden the picture of German-speaking children's productive use of verb inflections. Ott (2011) found that three year olds have considerable difficulties when asked to inflect pseudo verbs for 3^{rd} person singular (the model verb was presented in 3^{rd} person plural). These findings indicate that children either do not yet have an adult-like rule established that allows them to add the 3^{rd} person singular affix *-t* to every verb when the subject is singular or that serious performance limitations hinder the child to display her grammatical knowledge in elicited production tasks. Alternatively, the spontaneous speech data can be interpreted such that young children mostly produce inflected verbs as they have memorized them from the input. The present German data does not allow to distinguish between the two possibilities. But critically, it should be noted that German-speaking children would have to build more non adult-like rules relating noun form and verb form than their English-speaking peers, due to the greater variability of the German noun plural paradigm (e. g. Behrens and Tomasello, 1999). While English-speaking children could in principle operate on a rule like

'V+/Ø/ → V+/Ø/ following N+/s/', German-speaking children would have to phrase a similar rule for the *-s*-plural (*Auto, Autos*, the *-(e)n*-plural (*Hase, Hasen*, *-e*-plural (*Tier, Tiere*) etc.

3.4 Early verb inflection in other languages

Children's receptive processing of verb inflection has been examined not only in English but also in Spanish, Dutch and very recently in Italian. For comparison, the developmental patterns of verb inflection production for Dutch, Spanish and Italian will be roughly described in the following. The necessity of cross-linguistic research in the field of language acquisition to take into account typological differences between languages has been stressed by various researchers (e. g. Bates, MacWhinney, Caselli, Devescovi, Natale, and Venza, 1984; Xanthos et al., 2011; Phillips, 1995; Höhle, Schmitz, Santelmann, and Weissenborn, 2006).

Dutch children start producing inflected verbs around two and a half years of age and to develop beyond three years (e. g. Blom, 2003). In a longitudinal study examining the emergence of finite sentences in Dutch-speaking children's spontaneous speech, Blom (2003) found that inflection (on main verbs) appeared fairly late, since the first finite sentences were all created using modal verbs and the auxiliary *be*. As was the case for English and German, Dutch learning children first produced uninflected verbs, then inflected and uninflected verbs at the same time (were thus said to be in an OI-stage), and finite sentences became more and more frequent over time. Interestingly, Blom (2003) noted that she found only little lexical variation within the inflected forms and and little lexical overlap between inflected forms and non-inflected infinitives.[29]

To examine Dutch-speaking children's knowledge of verb inflection experimentally, Polisenska (2010) performed an elicited

[29] Since Blom's work focuses on the semantic and syntactic restrictions on children's RI-use (in terms of stages) and how these can be explained using various models and hypotheses from the nativist framework, no numbers on children's use of correctly inflected verbs in relation to age ranges are provided.

production task with children aged three to six years. An additional goal was to examine the extend to which a rule for finite inflection were productive, therefore familiar and novel verbs (introduced in the past participle form) were used. Overall, the children tested were found to be very good at producting correctly inflected verbs. The rate of correct inflection ranged from 65 % correct in the three year olds to about 95 % correct in the six year olds, with the four and five year olds in between. That is, the proportion of incorrect inflection decreased with age. Interestingly, no effect of verb familiarity was found, since children performed equally well on familiar and pseudo verbs. Especially high levels of correct inflection were found for the 3^{rd} person singular and plural forms. The finding that three year olds inflected nonce verbs for 3^{rd} person singular 100 % correctly contradicts earlier reported findings from Theakston et al. (2003). The English-speaking children added the *-s* in only about 50 % of the cases. It was concluded that Dutch-speaking children had acquired the finite verb inflection by the age of three. No specific claims regarding the type of rule are made in the study. Because children behave on an adult-level early on, the application of adult-like rules is taken for granted. The finding that verb familiarity did not affect the children's rate of correct inflection can be viewed as a hint that children applied an adult-like rule.

Spanish is a Romance language and it is, like Italian, known for the richness of the inflectional system, at least when compared to English (e. g. Aguado-Orea, 2004). It has repeatedly been reported that Spanish- and Italian-learning children produce verb inflections earlier in the course of development and that root infinitives, as found in German and English, are almost completely absent from Spanish and Italian children's speech (e. g. Guasti, 1993, 2002; Grinstead, 1994). It has been proposed that the richness of the morphological system is the driving force behind early acquisition and use of verb inflections (e. g. Phillips, 1995). Grinstead (1994) for example found only 5 and 10 % of RIs in two children acquiring Spanish aged 2;6

to 2;8 and 2;1 to 2;4, respectively.[30] Berger-Morales et al. (2005) found the cross-linguistic pattern to hold true even within children that acquired two native languages at once. The spontaneous speech of one German-Italian and one Spanish-English bilingual child, both aged 2;0 to 2;6 (longitudinal), reflected the pattern found cross-linguistically. The German-Italian bilingual child used root infintives, thus failed to inflect main verbs, at about 70 % in German, but only at about 5 % in Italian, and the Spanish-English biligual child produced RIs in 27 % of the English utterances, but only in 3 % of the Spanish ones. These patterns seem to confirm the assumption that children acquire verb inflections earlier and possibly easier when learning a language that provides them with a rich inflectional database (Xanthos et al., 2011).

Rubino and Pine (1998) investigated the production of subject-verb agreement in one child acquiring Brazilian-Portugese, which is similar to Spanish and Italian with regard to morphological richness and the possibility of producing grammatical subjectless sentences. Employing a thorough analysis of the child's spontaneous speech from 3;2 to 3;4, the authors provided data not only on the error rates (which were indeed very low) but also on co-occurrence rates of verb types and specific inflectional affixes. They found significantly lower agreement error rates in sentences with singular subjects than with plural subjects and indications that input frequeny contributed mostly to children's correct use of subject-verb agreement. The authors interpret their data as evidence that children acquire verb inflections marking agreement in a piecemeal fashion and that an abstract knowledge of agreement rules is doubtful.

Similar evidence is provided by Pizutto and Caselli (1994), who investigated the grammatical morphology in three Italian-speaking children. They found that 47 % of the verbs used by the children appeared only in one form (of 6 possible forms), and another 40 % appeared only in two or three forms. The remaining 13 % of verbs that appeared in four or more forms

[30] As reported in Berger-Morales, Salustri, and Gilkerson (2005).

were found to be highly frequent and highly irregular. This was assumed to reflect that children had merely memorized the irregular, high frequent verbs via rote learning. The authors concluded that Italian children do not master the whole verb paradigm for all known verbs at once but that they produced certain endings on certain verbs.

The findings by Rubino and Pine (1998) and Pizutto and Caselli (1992) are viewed as evidence that children's early constructions including verb inflections revolve around particular lexical items (c.f. Tomasello, 2000) and only become increasingly abstract and adult-like as the child's grammar develops. The children whose spontaneous speech was examined might either already have build up a non-adult like rule as presented in 2.6 or they might simply produce memorized chunks of speech (verbs plus inflectional affixes) they had earlier encountered in the input.

3.5 Factors influencing production of verb inflection

One factor that seems to have an influence on the rate and correctness of verb inflection production is the task in which the children's language output is collected, i. e. whether one examines spontaneous speech or elicited production data. But it is important to note that not only the demands of the task per se can account for differences found between the data from these two methods, but that verb selection itself might play a significant role. Assuming that verb frequency and familiarity have an impact on children's ability to use an inflected form correctly (e. g. Aguado-Orea, 2004; Theakston et al., 2003; Rubino and Pine, 1998)[31], verb selection itself might be a factor explaining different ages of mastery found with different methods: if children spontaneously produce familiar and highly frequent verbs, the chance of errors in spontaneous speech is lower

[31] Noun familiarity also has an influence on the production of determiners, as found by Boyle and Gerken (1997).

compared to elicited production.[32] Additionally, semantic demands are different in elicited production tasks, since the child is encouraged to produce a linguistic form that matches a meaning provided by the experimenter. When transcripts of spontaneous speech are analyzed, intentions and semantic influences cannot be controlled for.[33]

Other factors that were found to affect the production of grammatical morphemes were the frequency of the grammatical morpheme in the input (Lukács, Leonard, Kas, and Plé, 2009; Warlaumont and Jarmulowicz, 2011; Hsieh, Leonard, and Swanson, 1999; Theakston et al., 2003), the position of the verb in the utterance, and the length of the utterance (Song, Sundara, and Demuth, 2009). Other researchers found semantics to explain variability in grammatical morpheme production, since they noted that children produced past tense morphemes (*-ed*) earlier with verbs that denoted accomplishments (e. g. *dropped*) than with verbs that denoted activities (e. g. *played*) (Bloom, Lifter, and Hafitz, 1980). New research has identified 'neighborhood density' as an additional factor (Hoover, Storkel, and Rice, 2011).

Variability of the input seems to play a critical role in verb inflection production as well. Miller and colleagues (Miller, 2012) have found that input variability strongly affects age of productive mastery of regular noun plural inflection. Children who were presented with consistent input acquired morphological knowledge earlier than children who were faced with variable

[32]Norbury, Bishop, and Briscoe (2001) used a task that tries to combine the advantages of spontaneous and elicited production (following van der Lely (1998)). The child is asked to tell the experimenter things that his family and friends do every day. Importantly, it has to be something different each time the child tells something. The experimenter gives examples like 'Every day my mom cooks dinner' or 'Every day Mark watches TV'. If the child has difficulty thinking of verbs, he is encouraged to talk about typical routine in his house, e. g. what everybody does in the morning etc. Such a task has the advantage that the child produces only known and rather frequent verbs, like in spontaneous speech, but additionally he is encouraged to produce 3^{rd} person singular inflected verbs, the grammatical morpheme under investigation.

[33]See Soderstrom (2008) and Naigles (2002) for discussions on how semantic demands make linguistic tasks much more difficult for children.

input.[34] Very similar differences related to variability of the input were found for English-speaking children regarding agreement morphemes attached to the auxiliary *do* (Miller, 2012).

Still other research has found that children's production of verb inflection is constrained by phonological and prosodic factors. Already Gerken (1996) has noticed that omissions of articles in the speech production of English-speaking two year olds was influenced by the prosodic structure of the sentence. Relatedly, Song and colleagues have examined the influence of sentence position (medial vs. final) and phonological complexity of the verb stem (simple coda vs. complex coda) on the rate of accuracy in children's production of 3^{rd} person singular *-s* (Song et al., 2009). In longitudinal spontaneous speech data from one to three year old children, they found that children produced verb inflection morphemes less accurately when the verb had a complex coda (e. g. *needs*) compared to verbs that had a less complex coda (e. g. *sees*).[35] Using a logistic regression model, Song et al. (2009) found that coda complexity was the most effective factor in predicting the contexts where 3^{rd} person singular *-s* was produced correctly. But position of the verb in the utterance was found be to an important factor as well, with children producing more correct verb inflections in final compared to medial position. Interestingly, this even runs counter to frequency-based expectations (Hsieh et al., 1999), because English-learning children hear inflected verbs most frequently in a sentence-medial position.[36] Another factor that proved to

[34] In a series of studies, Miller and colleagues found that Spanish-speaking children from Mexico City produced significantly more plural *-s* on regular nouns that their Spanish-speaking peers from Chile. The critical difference between those two dialects lies in the variability of the overt plural marking with *-s*. In the mexican version, nouns, determiners and adjectives are consistently marked with an *-s* for plural, while the plural morpheme can be realized as *-s*, *h* or Ø in Chilean Spanish. Interestingly, Chilean children were also less adult-like in comprehension than their mexican peers. The consistency of the noun inflection in Mexican Spanish seems to help children learn the morphological paradigm of noun plural inflection.

[35] In line with the English data described above, the longitudinal data of four of the six children who provided spontaneous speech samples did not show mastery in Brown's sense of 90 % correct at the oldest age tested (3;6).

[36] But Fernald and McRoberts (1993) found that one year and three month old

have a significant influence on the rate of correct verb inflection production was utterance length. The longer utterance, the less likely it was that children produced the verb inflection correctly. The frequency of the inflected verb in the parental input of the children on the other hand did not seem to play a role in children's verb inflection production. An increasing MLU finally was significantly correlated with children's ability to produce verb inflection morphemes correctly. This indicates that there is indeed a strong correlation between children's general language abilities as measured by MLU and their production of 3^{rd} person singular *-s*. An elicited production experiment with two year olds confirmed the spontaneous speech findings: verb inflection production was more accurate in verbs with simple codas, thus in phonologically simple contexts, and in sentence-final position. These results are consistent with the *prosodic licensing hypothesis*, which states that young children will produce grammatical morphemes first in unmarked, phonologically and prosodically simple contexts (Demuth and McCullough, 2009).[37]

Song's findings were replicated and expanded by Sundara, Demuth, and Kuhl (2011), who tested the influence of sentence position of the verb on children's perception and production of 3^{rd} person singular verb inflection (using an elicited imitation task). They found significantly higher production rates in sentence-final compared to sentence medial position in 22-months old English-speaking children. In the 27-months olds, the rates differed numerically, but failed to reach significance.

children are better at recognizing newly learned words in utterance-final position, so sentence-final processing might generally be easier for children. This does not have to be related to grammatical morphemes in general (Shady and Gerken, 1999).

[37] See additionally Hsieh et al. (1999) for a comparison of 3^{rd} person singular *-s* and plural noun *-s*. The authors claim that plural noun is acquired earlier than its verbal counterpart, because it is more frequent in the input, appears more often in salient sentence-final position and is phonetically longer, thus easier to perceive. It is difficult to generalize these comparisons of plural noun and singular verb marking to other languages, like German, because of the specific case that *-s* marks both forms in English, but not in other languages.

A correlation analysis with production rates and CDI-scores showed that, somewhat unsurprisingly, children with larger vocabularies showed higher verbal morpheme production rates.

For German, only Ott (2011) has examined the influence of phonology on the production of 3^{rd} person singular *-t*. She did not find enhanced performance in 'simpler' phonological contexts in normally developing children, but did in SLI-children. But the children tested by Ott were slightly older, and it might well be the case that effects of phonological complexity were no longer visible. Additionally, the sub-syllables that Ott compared differed in their frequency, but not necessarily in their phonological complexity (depending on the underlying phonological theory, see Ott (2011)).[38]

Therefore, prosodic or phonological influences on verb inflection production cannot be ruled out in younger and linguistically less advanced children or in other phonological contrasts (e. g. simple vs. complex coda in the inflected verb, *rede-t* vs. *bastel-t*) in German. To date there is no study investigating whether the prosodic licensing hypothesis applies to German verb morphology as well.[39] It is subject to further research, whether verb position, utterance length, MLU and verb frequency have an effect on verb inflection production in young German-speaking children.[40] Additionally, it would be of great interest whether one or more of the above-mentioned factors could explain acquisition patterns found in the German spontaneous speech data that were so far only used to discuss early root infinitives or the relationship between finiteness and agree-

[38] Ott (2011) compared long and short vowels in the verb stems and not the complexity of the coda, as Song et al. (2009) did.

[39] Only Frank and Düming (2007), cited in Ott (2011), found a relationship between the phonological ability to produce final plosives and the morphosyntactic ability of marking 3^{rd} person singular verb with the correct *-t* inflection. But this pattern was only found in one single-case study of a boy with a severe language disorder. It is difficult to generalize from these observations to normal developing German-learning children.

[40] Verb position in this context should not be confused with the relationship between verb position and finiteness-status of verbs in German. To test the influence of verb position as a prosodic factor, one would need to compare finite transitive and intransitive verbs, in the first case the verb would be followed by an object noun and in the latter case, not.

ment. Most importantly, all of these factors could play a role in comprehension of verb inflection as well. This has to be kept in mind when the results of the present experiments are discussed.

Summarizing, a variety of factors have been identified that affect children's productive use of verb inflection. These include verb frequency and familiarity, neighborhood density, variability of inflectional marking in the input, sentence length and the position of the verb in the sentence as well as the phonological complexity of the verbs. Still, this does not necessarily entail that constructivist theories of verb inflection acquisition are more appropriate than theories assuming the early presence of linguistic rules and adult-like representations, because the latter can still call on performance factors in production. In this scenario, children would produce less inflected verbs if phonological structure and/or sentence length and/or lexical frequency would put higher demands on their overall processing capacities.

3.6 Summary: Time course of verb inflection production

The various studies on verb inflection production suggest that normally developing English-speaking children start to acquire present tense verb morphology at around two years of age, but that it takes children three additional years to use inflectional morphology at an adult level, i. e. in every obligatory context. Especially the elicitation studies allude that children correctly inflect verbs only at the age of four or five years (Rice et al., 1995, 1998). Familiarity of verbs seems to affect the rate of correct inflection (Berko, 1958; Theakston et al., 2003; Rubino and Pine, 1998), and might be one reason why spontaneous speech samples generally show earlier mastery than found in elicited production studies. For German inflection, the data available so far suggests that the youngest age of productive mastery might be around two and a half years (Poeppel and Wexler, 1993), but this should be taken with caution. First of all, Poeppel and Wexler (1993) presented a single case study, and this child might just be very elaborate and far developed in his morphological knowledge (a similar point is made by Blom (2003)). The findings by Bittner (2003a,c) and Clahsen (1986) already point to a slightly later age of productive mastery, although these studies are based on spontaneous speech data as well. When asked to use novel verbs in elicited production experiment, German children showed productive mastery at a much later age, at around four and a half years (Ott, 2011). The protraced period of acquisition indicates that verb inflection is certainly one of the challenging aspects of language acquisition. This view is supported by numerous studies that show that verb inflection is one of the most error prone and demanding linguistic structures for children with specific language impairment and late L2-learners (e. g. Leonard, Eyer, Bdore, and Grela, 1997; Leonard, Camarata, Pawlowska, Brown, and Camarata, 2007; Oetting and Hadley, 2008; Rice et al., 1998; Ionin and Wexler, 2002; Foote, 2011).

The findings that the factors of input frequency, input consistency, familiarity, sentence position, utterance length, phonological complexity of the verb stem and children's MLU all play a role in verb inflection production might account for the relatively long time it takes children from the first uses of grammatical morphemes to the achievement of adult level competency. Language acquisition theories differ significantly in how they explain the temporal gap between the first productive use and the much later mastery of verb inflection. Nativist theories that assume early abstract morphosyntactic knowledge calls on additional constraints in the child grammar (e. g. Wexler, 1998) or performance factors (e. g. Phillips, 1995) that cause children to continue producing uninflected verbs in finite sentences, while constructivist theories assume early item-based knowledge that only later allows children to generalize all morphological forms to all verbs (e. g. Theakston et al., 2003).

Nevertheless, because incorrect use of verb inflection has only rarely been documented in spontaneous speech, most researchers within the nativist tradition have claimed that children *have the knowledge* of finiteness, verb inflection and subject-verb agreement (Wexler, 1994; Phillips, 1995). This is even taken further in the informal claim that children are "little inflection machines" (Wexler, 1998, pg. 43) who learn the inflectional properties of their native language almost instantaneously. This phenomenon is referred to as the *Very Early Knowledge of Inflection*-account (VEKI). Only some researchers within the nativist framework have claimed that children go through a stage in which they can determine the finiteness-status of a verb but still lack complete knowledge about subject-verb agreement (e. g. Clahsen et al., 1993; Ingham, 1998; Verrips and Weissenborn, 1992). Still, the claim that children *know* about verb inflection very early is found repeatedly, among others stated by Phillips (1995):

> [...] children have a good knowledge of the morphology of their target language at a very early age, as early as anybody has been able to check, in fact. In a large proportion of their utterances, however they fail to realize this morphological knowledge. (Phillips, 1995, pg. 7)

Proponents of the constructivist view claim that low error rates do not inform about children's knowledge of verb inflection since all spontaneous language can simply reflect memorized constructions that are not generated via linguistic rules. The two alternative views on early verb inflection, the influence of task as well as lexical, phonological and syntactic factors on the rate of verb inflection give rise to the question: what does *knowledge of verb morphology* mean? Does it mean that children have to be able to produce correctly inflected verbs in their everyday conversations? Or that children have to be able to extend the productive use to newly learned words, showing morphological productivity? Or does 'knowing verb inflection' mean that children have to be sensitive to the presence of an agreement form in obligatory contexts? Or does it mean that children have to be able to use inflectional morphemes to infer semantics, e. g. the tense of the sentence or the number of a sentence subject? Does knowledge refer to the form or the function of verb inflections or both?

In the following I will present research on the receptive side of verb inflection. This work can roughly be divided into studies that have shown early sensitivity to the presence of verb inflection morphology, studies that have shown that metalinguistic detection of verb inflectional morphology is not easy for children and studies that have demonstrated particular difficulties in using verb inflection morphology in sentence comprehension. As will be seen, these findings call the notion of 'good knowledge of morphology' into question.

4 Detection of verb inflection

In this chapter, children's receptive abilities regarding verb inflections will be presented. Critically, this chapter deals with with children's sensitivity to the well-formedness of sentences in which subject and verb either agree or an agreement violation is present, not with children's use of verb inflection in sentence comprehension. The studies tapping sensitivity to agreement violations can be divided into (1) infant studies, which show that children below age two can detect agreement violations in a preferential listening paradigm, and (2) child studies, which show that school-aged children's abilities to judge the grammaticality of a sentence containing agreement violations is rather poor. Before the relevant studies are presented, I will lay out the importance of including the receptive modality in language acquisition research. Further, I will briefly review the methods used to examine infant's sensitivity to grammatical well-formedness of sentences. With regard to grammaticality judgement tasks, I will additionally sketch out a theoretical framework, the 'Competition Model', which is frequently used to explain children's ability to detect agreement violations in such tasks (or the lack thereof).

4.1 The relevance of receptive grammar

While the earliest studies on language development mostly examined children's language production, the focus has shifted towards the receptive modality, testing children's comprehension abilities. This is not only due to the development of 'new' research methods (like preferential listening, intermodal preferential looking, and event related potentials), but also due to theoretical considerations. First of all, it is irrefutable that comprehension is a very important part of language use, in children and in adults. We do not only talk, but we listen to what other people say, and to have successful conversations, both processes need to work properly.

Concerning child language, receptive studies have shown that young children's knowledge of grammatical form very often exceeds their productive capacities (e. g. Gerken et al., 1990; Höhle et al., 2006). The same pattern was repeatedly found for phonological and lexical knowledge in children (e. g. Höhle and Weissenborn, 2003; Yoshida, Fennell, Swingley, and Werker, 2009; Mani and Plunkett, 2010b). In many cases, earlier receptive knowledge carries great implications for theories of language acquisition. For example, early theories (e. g. Cazden, 1968) which had suggested that children's omissions of functional elements were due to children's inability to recognize these elements in the speech stream, have been proved wrong by infant speech perception studies showing that preverbal children are very adept at recognizing functional elements in the speech stream (e. g. Gerken and Aslin, 2005; Jusczyk, 1997). Moreover, function morphemes even aid children in the process of language acquisition, since they help to segment the speech stream (e. g. Shi and Lepage, 2008) and assign syntactic categories to newly learned words (e. g. Höhle et al., 2004; Mintz, 2006; Bernal, Lidz, Millotte, and Christophe, 2007; Hochmann, Endres, and Mehler, 2010).

Additionally, it has been claimed that receptive performance in children can show children's productivity where productive performance can not. Such productivity, in the sense of productive use of grammatical rules and linguistic knowledge, has often been called on to prove or refute the existence of abstract syntactic categories in early child grammar and the alternative view of exemplar-based learning in language development (Abbot-Smith and Tomasello, 2006; Naigles, 2004; Fisher, 2002).

Further, studies tapping the receptive modality allow researchers to focus on language processing, rather than on the (end) product of language use, i. e. a final interpretation, which is under investigation in most production studies (Swingley, Pinto, and Fernald, 1999). Clark and Hecht (1983) have also argued that researchers should focus more on the process than on the product when investigating child language and that production and comprehension should both be considered if one wants to draw a complete picture of language acquisition.

> [...] understanding syntax acquisition from the point of view of perception is essential to the task of fully describing this acquisition process. (Soderstrom, 2002, pg. 51)

Therefore, testing the receptive side of verb inflection processing should provide us with a more complete picture of children's knowledge of grammatical morphemes. Receptive studies may shed some light on the question of why children take quite some time to produce verbal agreement morphemes at an adult level and what kind of linguistic rule might underlie children's early verb agreement processing.

4.2 Methods for testing early receptive grammar

One cannot just ask an infant or even a child what he or she thinks about the structure of a sentence and whether she considers a sentence as well-formed or not (as syntacticians often do with colleagues or undergraduate students). Sometimes it's even hard to ask a child what he or she thinks a sentence actually means. This is why the development of 'new' techniques in child language research has had such an impact on theories of language acquisition in the last 30 years. Such techniques can be roughly divided into those that measure children's preference for one structure over another, like the Headturn Preference Procedure (HPP) and those that measure children's ability to choose which one of two or more stimuli match an auditory stimulus, like the Intermodal Preferential Looking Procedure (IPLP).

The first method is used to measure how much time children orient to flashing lights positioned to their left and right, while they are presented with auditory stimuli that differ in a certain respect (e. g. Nelson, Jusczyk, Mandel, Myers, Turk, and Gerken, 1995). Preference either relates to something children bring to the lab with them, e. g. a preference for child-directed speech over adult directed speech (Fernald, 1985), or to something that they have just been familiarized to, e. g. some newly learned words (Jusczyk and Aslin, 1995). This method serves to investigate how infants perceive the syntactic or phonological properties of a certain string of speech sounds, without relating to semantics. Thus, sensitivity to syntactic well-formedness can be examined even in preverbal infants.

An alternative method to measure young children's preference is used by Sundara et al. (2011). In the 'central fixation auditory preference procedure' (Pinto, Fernald, McRoberts, and Cole, 1998) children are presented with a visual display along with usually unrelated auditory stimuli (although Sundara and colleagues presented cartoons that provided a referential context for each sentence). The dependent measure is the length of time a child fixates on the visual display when hearing a certain kind of stimuli (e. g. a grammatical sentences) versus another kind of stimuli (e. g. an ungrammatical sentences). Longer listening times in one condition over the other indicate the existence of a preference and with this, that the child can distinguish between the two verbal conditions. Although the classical HPP-method of investigating children's preferences beween two auditory stimuli seems to be more widely used than the central fixation procedure, the latter is nevertheless widely accepted (e. g. Shi and Werker, 2001) and probably more suitable to test slightly older children than possible with the HPP (Fernald and McRoberts, 1996).

If one wants to study what children actually understand, the IPLP can be employed (e. g. Golinkoff, Hirsh-Pasek, Cauley, and Gordon, 1987; Hirsh-Pasek and Golinkoff, 1996; Trueswell, 2008). The dependent variable is the amount of time children look towards a picture or a moving visual display that matches the auditory stimulus and compares this to the duration of looks towards a distractor picture. Since this method examines comprehension of linguistic structures, it will be presented in more detail in Chapter 5.1.

4.3 Early sensitivity to verb inflection

4.3.1 Early sensitivity to verb inflection in English

Soderstrom (2002) has examined English-learning children's sensitivity to the presence of 3^{rd} person singular *-s* with the goal of reaching a better and more complete understanding of the optional infinitive (OI) stage in children. She reasoned that the properties of the OI-stage found in the productive domain should be detectable in the receptive domain as well, "if the OI-phenomenon truly reflects the nature of infant's early grammatical knowledge" (Soderstrom, 2002, pg. 40). Such properties, based on Wexler's work, are: the use of finite and non-finite verbs in finite sentence contexts, the very rare use of incorrect verb inflections or inflected verbs in inappropriate contexts and the ability to distinguish between finite and non-finite forms with respect to the position of the verb in the sentence. If children would (1) distinguish between finite and non-finite verbs, (2) reject incorrect forms but (3) still accept finite and non-finite forms in finite contexts, one could assume that OIs are allowed in the productive and receptive domain. Such a finding would support the theory that OIs assumptions actually reflect children's early grammatical knowledge and that OIs were not due to processing constraints or the input. If children on the other hand would reject OIs in favor of the adult form, one would need to explain the relationship between the productive and receptive domains (e. g. reception would then reflect adult-like grammatical knowledge while production would be influenced by performance factors), or one would need to abandon the idea of OIs as a grammatical phenomenon.

Employing the HPP-technique, Soderstrom and colleagues presented 19-month old English-learning infants with two sets of passages, one with agreement morphology (3^{rd} person singular *-s*) on the verb and the other one containing agreement violations, see 4.1 and 4.2 for examples (Soderstrom, 2002; Soderstrom, Wexler, and Jusczyk, 2002)

(4.1) At the bakery, a boy bakes bread.

(4.2) (*)At the bakery, a boy bake bread.

While in the adult grammar, only 4.1 constitutes a valid grammatical sentence, an OI-grammar would allow both kinds of sentences (indicated by (*)). It was found that 19-month old, but not 16-month old, children preferred the grammatical passages over the ungrammatical ones. This finding was in line with other HPP-experiments in showing preference of the grammatical form over the ungrammatical form (e. g. Santelmann and Jusczyk, 1998), but it did not support the predictions for a receptive OI-stage.

Since OI-theory additionally makes predictions about the placement of finite and non-finite verbs relative to negation, a weaker version of the receptive OI-theory was tested in a second experiment. Here, sentences like in 4.3 and 4.4 were presented, with 4.3 being ungrammatical in adult and OI-grammars (since the finite verb is in the wrong position) and 4.4 being only ungrammatical in the adult grammar, but still possible in an OI-grammar, which allows non-finite forms (but only if they stay in the position below the negation).

(4.3) *At the bakery, a boy not bakes bread.

(4.4) (*)At the bakery, a boy not bake bread.

In this experiment, 19-month old infants again showed a preference for the passages containing the verb inflection *-s* over those passages without verbal agreement. This finding does not support predictions made by the OI-theory. Soderstrom suggested that children simply did not recognize the negation and thus just preferred a structure that seemed to be grammatical to them, because it contained an inflected verb. An additional experiment in which the negation *not* was substituted with the pseudo word *nep* confirmed this suggestion. Again, children preferred the passages that contained agreement morphology over those that lacked verb morphology. To finally check whether children simply preferred the passages containing a *-s*-morpheme because of acoustical reasons, a further experiment contrasted sentences like 4.5 and 4.6.

(4.5) *At the bakery, a boy does bakes bread.

(4.6) ?At the bakery, a boy does bake bread.

In this experiment, no preference was found, thus the suggestion that all earlier results by Soderstrom were simply based on an acoustical preference of passages containing more *-s*-morphemes could not be supported.

On the one hand, the reported findings contradict the idea of an OI-stage in the receptive domain in English-learning children. On the other hand, they can be regarded as a "first evidence that infants as young as 19 months are sensitive to English verbal agreement morphology" (Soderstrom, 2002, pg. 49). All the infants tested in Soderstrom's study were well below the age of productive mastery of verb inflection in English (see Chapter 3). Therefore, the results are in line with the idea that infants are sensitive to functional morphemes long before they produce them (e. g. Gerken and McIntosh, 1993; Shady, 1996; Höhle et al., 2004).

Since *-s* serves not only as the present tense singular verb inflection in English, but also as a regular plural marker on nouns, Soderstrom (2002) further examined whether 19-months old infants were sensitive to the mutual distribution of this inflectional affix. Two groups of infants were tested. Group 1 was presented with passages that contained singular grammatical sentences like 4.7 and plural grammatical sentences like 4.8 on the one hand and passages that contained ungrammatical utterances without any inflection like 4.9 on the other hand ('no-s condition'). Group 2 was again presented with the passages containing grammatical sentences as in 4.7 and 4.8, but these were contrasted with ungrammatical sentences in which both inflections were present, like in 4.10 ('2-s condition').

(4.7) The boy bakes bread.

(4.8) The boys bake bread.

(4.9) *The boy bake bread.

(4.10) *The boys bakes bread.

Results showed that children preferred grammatical passages (singular and plural) compared to the passages that did not contain any inflection. Interestingly, no preference for the grammatical passages was found compared to the passages that contained two inflections.[41] Summarizing, infants preferred grammatical sentences containing a singular subject and a verb carrying *-s* inflection over sentences in which the verb did not carry an *-s* inflection. But infants as well preferred sentences containing an *-s* inflected verb, even the verb was preceded by a negation, which is ungrammatical in adult grammar. No preference was found when one *-s* inflection on the main verb was compared to sentences with *-s* inflection on two verbs, *do* as well as the main verb. Finally, children distinguished grammatical sentences that either contained an *-s* inflection on the subject noun *or* the verb from sentences without any verb inflection, but failed to detect a difference between grammatical sentences with one inflection (on noun or verb) and ungrammatical sentences with *-s* inflection on subject noun *and* verb. Especially the very last finding actually calls the notion of grammatical knowledge of inflection in question. As Soderstrom puts it: "Under most linguistic theories of syntax, it would be difficult to explain why the infant's grammar would include doubly-inflected sentences, but not uninflected ones" (Soderstrom, 2002, pg. 74). This is especially puzzling because the reversed pattern is found in production, where we find sentences without inflection but not doubly-inflected ones.

Generally, the data from Soderstrom's dissertation shows that English-learning infants seem to be sensitive to the presence of inflection in the sentences, although they seem to be indifferent as to where this inflection occurs or how much of it occurs (whether a sentence contains one or two *-s*-inflected elements). This would mean that infants are sensitive to *-s* inflection, but still lack a sophisticated knowledge of the agreement relations that subject and verb in a sentence enter into. With regard to the rules that might underlie children's early processing of verb in-

[41] Comparable to the experiment in which both the auxiliary *do* and the verb carried the verb inflection.

flections (see 2.5 and 2.6), Soderstrom's findings can best be explained under the assumption of a non adult-like rule, a claim that is formulated by Johnson et al. (2005) as well.

The result pattern fits the claim that young children are sensitive to the distributional patterns of the input, which has been found in studies on artificial language learning in infants (e. g. Saffran, Aslin, and Newport, 1996; Gómez and Gerken, 1999; Gómez, 2007). Recent research has found that child directed speech contains an abundance of distributional regularities (e. g. Mintz, 2006; Weisleder and Waxman, 2010) which are very likely to aid children in the process language acquisition (e. g. Höhle et al., 2004). Further evidence for infants' ability to track non-adjacent dependencies between function words comes from van Heugten and Shi (2010), who found that French-learning 17-month-old infants detect remote determiner-auxiliary co-occurences that span phonological phrase boundaries, and from Höhle et al. (2006), who found that 19-months old German-learning infants recognize the dependency relation between the auxiliary *hat* and the past particple verbs, like *geheult* (see as well Santelmann and Jusczyk (1998) for English-learning children). Thus, infants seem to be very apt to detect and process dependencies between functional elements in the language they hear, and English-learning infants' sensitivity to verb inflections seems to be a further aspect of this ability. Nevertheless, this statistical sensitivity apparent at 19 months must at some point feed into what will become the adult linguistic system (Soderstrom, 2002).

Soderstrom et al. (2007) further examined children's sensitivity to the distribution of the *-s* marker in English sentences in relation to other function and content words in a series of experiments. The English-learning 16-month old children were presented with sentences like 4.11 and 4.12.

(4.11) They used to sing in the chairs on the porch.

(4.12) *They used to chairs in the sing on the porch.

Children preferred the passages with correctly positioned functional elements over those in which the inflections were misplaced. Critically, this was only found when the inflected content words were adjacent to other function words like auxiliaries and pronouns, which clarified the syntactic class of the content word. Thus, the authors conclude that there is a relation between inflectional morphology and function words that might play a crucial role in the formation of early grammatical knowledge. Interestingly, Soderstrom and colleagues also found that word familiarity affects infant's ability to detect grammatical violations, since the infants showed reliable preferences when tested on familiar content words, but not when tested with pseudo content words. This last mentioned finding suggests that children, are better at applying a (non adult-like) rule that helps them to detect statistical regularities in instances they have encountered before. This might serve as an indicator that children build up their grammatical knowledge from individual memorized constructions and that application of rules is much easier in known constructions.

Sundara et al. (2011) tested as well whether English-speaking children (22-months and 27-months old) were able to detect the presence or absence of 3^{rd} person singular *-s*, but they used a modified version of the 'central fixation auditory preference procedure' (Pinto et al., 1998). The main aim of the study was to investigate whether sentence position affects the perception of the grammatical morpheme, as was found for production (Song et al., 2009; Sundara et al., 2011). This was found to be the case. Younger children, aged 22 months, preferred the grammatical passages with a 3^{rd} person singular *-s* attached to the verb, but only when the verb was in sentence-final position (e. g. *There he sleeps*) and not when the verb was positioned sentence-medially (e. g. *He eats now*). The older children, aged 27 months, preferred the ungrammatical passages, but again only when the verb was presented sentence-finally but not sentence-medially. The findings are interpreted as evidence that articulatory complexity alone cannot account for children's enhanced production performance of 3^{rd} person singular *-s* in sentence-final compared to sentence-medial position. The author interpret their findings as evidence that perceptual factors contribute to production of the verb inflectional morpheme as well, at least in certain prosodic contexts.

The findings by Sundara and colleagues are noteworthy for two reasons: first of all, they are again a good example of how important and informative receptive studies can be when production data can be explained by various theoretical accounts (in this case, articulatory versus perceptual underlying reasons). Second, and more important for the present study, should be noted that the 22- and the 27-month old children were *not* able to distinguish between sentences with and without 3^{rd} person verb inflection when the verb was presented in sentence-medial position. Thus, even though the children were older than the ones tested by Soderstrom and colleagues, they were not able to show the same grammatical knowledge. Importantly, all the verbs in Soderstrom's passages were presented sentence-medially and the sentences were longer and much more variable than the ones used by Sundara and colleagues (Soderstrom, 2002, pg. 133-134). Thus, it is unclear why (and how) 18-month old English-learning infants preferred grammatical sentences over ungrammatical ones and 22- and even 27-month old English-speaking children failed to do so when the verb was in sentence-medial position. The different methods could be one explanation, if one assumes that the referential pictures used in the visual fixation paradigm encouraged the children to process the test sentences more referentially. One could suggest that the 18-month old children based their preference solely on superficial properties of the input (and reacted to the presence of one functional *-s*-morpheme in relation to other elements in the sentence, i. e. the subject), while the older children tried to parse language more semantically (especially when encouraged by the testing method and material), and thus failed to detect the grammatical morpheme. If children only had a rule that worked properly based on superficial properties of the input (see 2.6), a referential input that 'induced' semantic processing might have hindered children from applying this rule properly.

4.3.2 Early sensitivity to verb inflection in Dutch

Dutch served as one of the other languages in which early sensitivity to verb inflection and grammatical well-formedness had been tested. Polisenska (2010) examined 18- to 19-month-old Dutch-learning children's early receptive knowledge of verb inflection, first of all to test Wexler's VEKI-account and secondly to expand Soderstrom's findings to another language with a richer inflectional system than the English one.[42] To disentangle underlying phonological and grammatical explanations for possibly found preferences, she introduced four condition conditions in her experiment. In condition 1 and 2, sentences contained a 3^{rd} person singular subject with an agreeing verb containing *-t*-inflection or with a disagreeing verb containing the 3^{rd} person plural inflection *-en*. In condition 3 and 4, sentences contained a 3^{rd} person plural subject, again either with an agreeing verb, in this case carrying *-en*-inflection or with a disagreeing verb, in this case carrying the *-t*-inflection (see 4.13 to 4.16 for examples).[43]

(4.13)	*De*	*wind*	*waait*	*door*	*het*	*bos.*
	The	wind-SG	blow-3SG	through	the	forest

'The wind blows through the forest.'

(4.14)	**De*	*wind*	*waaien*	*door*	*het*	*bos.*
	The	wind-SG	blow-3PL	through	the	forest

'The wind blow through the forest.'

(4.15)	*De*	*liedjes*	*klinken*	*mooi.*
	The	song-PL	sound-3PL	beautiful.

'The songs sound beautiful.'

[42] According to Wexler (1994), children have full knowledge of agreement inflection at the age of 18 months.

[43] Regular noun plural in Dutch is either formed by adding a *-s* or and *-en* to the noun. Polisenska used only nominal plural subjects that were suffixed by *-s*, "to keep the category of nouns and verbs morphologically distinct" (Polisenska, 2010, pg. 115).

(4.16) **De liedjes klinkt mooi.*
The song-PL sound-3SG beautiful.
'The songs sounds beautiful.'

The results of the classically designed HPP-experiment revealed the following: no preference was found when all grammatical and ungrammatical passages (thus collapsed across subject number) were compared. Additionally, no preference was found when passages containing an *-en*-inflection and passages containing a *-t*-inflection were compared, thus no preference for either of the phonological forms was found. When only passages containing a singular subject were analyzed, children showed a preference for the ungrammatical, thus the *-en*-inflected verbs, over the grammatical, *-t* inflected, ones. When only passages containing a plural subject were analyzed, no preference was found. The results are interpreted as evidence that Dutch-learning children aged 18- to 19-months detect violations in finite verbal inflections and are thus sensitive to verbal inflection. Since no overall preference for *-en*-affixes was found, a phonological account of the data is ruled out by the author. The findings by Polisenska finally do not support the assumption of an early OI-stage in children's grammatical development, but are rather interpreted by the author as being support for the VEKI-hypothesis.

It might be the case though that Dutch-learning children based their preference merely on distributional properties of the input instead of on abstract grammatical knowledge about subject-verb agreement, which is claimed by the VEKI-hypothesis. Simply put, the children could have preferred the ungrammatical passages in the singular subject condition because they were surprised that a subject noun without an *-s*-affix was combined with an *-en*-inflected verb. If children had formed a rule like in 2.6 or rather an equivalent rule for Dutch, something like 'V+/-en/ $\rightarrow$ V+/-t/ following NP+/Ø/', they would be expected to behave exactly as they did. They would 'know' that nouns without a final *-s* should be followed by a verb with a final *-t*, and they would be surprised if this was not the case. The lack of preference in the plural noun conditions can then be explained by assuming that the children have not acquired a rule for NPs with -s yet. Thus, I want to assert that Polisenska's findings do not necessarily imply that children have acquired an adult-like rule regarding verb inflection. This assumption nevertheless credits the child with sophisticated knowledge about distributional patterns of linguistic elements in their speech input.

4.3.3 Early sensitivity to verb inflection in French

Nazzi and colleagues examined French infants' sensitivity to grammatical non-adjacent dependencies involving subject-verb agreement in a series of HPP-experiments. Number was overtly (audible) marked on the determiner of the subject DP and the agreeing verb, therefore, the dependency spanned two syllables (see 4.17 to 4.20 for examples). French presents an interesting test case, because number marking on the verb is highly irregular in this language. While other studies investigating the detection of non-adjacent dependencies between function words and morphemes usually tested regular dependencies between two elements (e. g. Soderstrom et al., 2007; Höhle et al., 2006), this study focused on children's early knowledge of irregular dependencies.

(4.17)	*Le*	*garçon*	*fait*	*le*	*vippe.*
	The-SG	boy	make-3SG	a	vippe.

The boy makes a vippe.

(4.18)	*Les*	*garçons*	*font*	*le*	*vippe*
	The-PL	boy(s)	make-3PL	a	vippe.

The boys make a vippe.

(4.19)	**Les*	*garçons*	*fait*	*le*	*vippe.*
	The-PL	boy(s)	make-3SG	a	vippe.

The boys makes a vippe.

(4.20)	**Le*	*garçon*	*font*	*le*	*vippe*
	The-SG	boy	make-3PL	a	vippe.

The boy makes a vippe.

Nazzi and colleagues used high-frequent and familiar verbs, but with very variable phonological plural markings (e. g. vowel change, vowel change plus consonant addition, *-s* addition, vowel lengthening plus *-s*-addition). The 18- and 24-months old children (but not 14-months olds) preferred the grammatical passages over the ungrammatical ones when tested with familiar content words. When the experiment was replicated with pseudo-verbs, no preference was found. Nevertheless, the findings are remarkable. Despite the fact that the encoding of the plural number on the verbs varied across nine distinct phonological realizations and that the dependency between determiner and verb inflection spanned across two syllables and two structurally distinct phrases (subject-phrase and verb-phrase), 18-months olds showed sensitivity to the grammatical dependency. One possible explanation suggested by the authors was that children applied phonological rules, such as 'singular verbal forms co-occuring with *le* N end in a vowel and plural verbal forms co-occuring with *les* N end in a consonant' (Nazzi et al., 2011, pg. 131), but this was, according to the authors, ruled out with the second experiment that used pseudo-verbs and controlled for this phonological rule application.

The findings are in line with studies showing that 18-months olds are capable of tracking non-adjacent dependencies (e. g. Höhle et al., 2006; Santelmann and Jusczyk, 1998; Van Heugten and Shi, 2009), but they also show that children seem to be able to do this despite great morphophonological varation. Notably, input frequency does not seem to play a critical role in the French children's acquisition of the subject-verb dependency, since an input analysis showed that the constructions under investigation were actually quite rare in infant's input (Nazzi et al., 2011). The good performance despite very little input leaves the authors puzzled. They claim that infants must have generalized from subject-verb agreement instances in which the dependency spanned across longer stretches of speech (which were found in the input, while shorter dependencies were not). The fact that most verbs used were not even considered to be part of the children's lexicon, as revealed by a parental questionnaire, makes the findings even harder to interpret.

4.3.4 Summary of early sensitivity studies

Summarizing, young children's receptive ability to detect inflectional morphemes in spoken language have been tested in only a few studies so far. Using the HPP, Soderstrom (2002), Soderstrom et al. (2007) as well as Nazzi et al. (2011) found that 17- to 19-month old infants preferred passages with subject-verb agreement over those with agreement violations.[44] However, it is unclear to which properties of the speech children in such HPP experiments actually attended to and what kind of rules they based their preference on. The possibility that the children process the number information on the subject and verb and additionally check the matching of these morpho-syntactic features on a grammatical basis (in form of an adult-like rule, see 2.6) is not very likely. It is rather considered to be likely that children's preference for grammatical structures is based on their knowledge about distributional properties of verb forms and nominal properties.

The assumption that children under two years of age process grammatical features and are able to infer the semantic implications of these features that are involved in adult processing of subject-verb agreement relation, is not justified based on the HPP data. Additionally, recent research pertaining to relevant knowledge underlying the number contrast renders adult-like processing unlikely. Carey and colleagues investigated the development of the conceptual number distinction between 'one' and 'more than one' and how it depends on morphological number marking. They have investigated this topic in preferential looking experiments (Kouider et al., 2006) and in the manual search paradigm (Barner, Thalwitz, Wood, Yang, and Carey, 2007; Wood et al., 2009). They found conclusive evidence that children learn to make the conceptual number distinction between singular and plural between 22 and 24 months of

[44] Although Soderstrom did not find this pattern in every condition, because children did not prefer agreeing passages over those that contained two inflectional affixes, one on the noun and one on the verb.

age. They additionally found that the conceptual, non-linguistic distinction between singular and plural does not depend on the number morphology of the language, because Mandarin-learning children developed the ability to distinguish number at the same time in development as English-speaking children did, even though Mandarin does not code number linguistically (Li, Ogura, Barner, Yang, and Carey, 2009).

Additional evidence against a grammatical basis for the preference for agreeing structures found in very young children comes from literature on noun plural acquisition. Acquisition of noun plurals seems to precede acquisition of verb (plural) morphology, at least in English (e. g. Brown, 1973; de Villiers and de Villiers, 1973a), although it takes children quite a while to reach adult levels (e. g. Ettlinger and Zapf, 2011). Some researchers have even claimed that noun plural morphology is a prerequisite for verb inflection production (e. g. Pawlowska et al., 2008), or that the comprehension of verb agreement morphemes depends on the subject features, because children initially rely on the number information provided by the subject (Leonard, Miller, and Owen, 2000; Keeney and Wolfe, 1972).[45] Interestingly, 24-month-old children do not seem to be able to comprehend plurality when this is signaled by the noun phrase alone (Wood et al., 2009; Kouider et al., 2006). It is therefore unclear how children aged 18- and 19-months are supposed to be able to classify a subject noun correctly as singular or plural (which would be necessary assuming the use of an adult-like rule as presented in 2.5). Thus, matching the grammatical and semantic subject number information with the number information of the verb inflection should not be possible for one and a half year olds.[46]

[45]This relates to the question of whether verb inflectional information can ever be interpreted semantically but itself or whether verb inflection only marks the syntactic agreement relation between subject and verb, and listeners always have to process subject features to infer the number of the sentence subject.

[46]Two IPLP-experiments reported in Soderstrom (2002) provide further evidence that English-speaking toddlers aged 19- and 23-months do not process the number information provided by the verb (and noun plural) inflection. Since these experiments tap children's use of inflection in sentence compre-

Thus, the preference for the grammatical over the ungrammatical form in the children tested by Soderstrom and colleagues and Nazzi and colleagues can not rest on early semantic knowledge of verb inflections. Still, children must have the ability to track the dependencies between subject and verb morphology, otherwise they would not have been able to prefer the grammatical passages. Nazzi et al. (2011) suggest that 18-months old children are able to build form-based categories and that they have the ability to generalize. Thus, in the case of French, infants could have learned two morphological realizations of singular and plural verbs and that these go with either *le* or *les*. This is very similar to a non adult-like rule as proposed in 2.6 and would entail no comprehension of singular-plural distinction, but if children were able to generalize from these two instances, they could well distinguish grammatical and ungrammatical passages in an HPP-experiment. To enable semantic comprehension, further rules (such as '*les*+NP refers to plural / more than one entity') would need to be in place, or children would have to learn the adult-like rule containing abstract grammatical features. Support for Nazzi's claim that children build form-based categories and generalize from these is found in artificial language learning studies (e. g. Gerken, Wilson, and Lewis, 2005). These studies showed that children were able to form proto-categories without any referential cues. This still means that very young children have some abstract knowledge of the linguistic patterns in their native language, but that the knowledge is different from adult-like knowledge and does not necessarily allows for semantic processing.

4.4 Grammatical judgements of subject-verb agreement

4.4.1 Development of metalinguistic judgement ability

Asking children for a judgement about the grammaticality of a sentence is not an easy endeavor. Children very often evalu-

hension, they will be presented in Chapter 5.

ate the semantic aspects of an utterance and they have a hard time focussing on the form rather than on the meaning. Still, it is possible to extract grammaticality judgements even from young children (e. g. McDaniel and Cairns, 1996; de Villiers and de Villiers, 1972; Smith and Tager-Flusberg, 1982). Judging the grammaticality, or well-formedness, of a sentence requires metalinguistic abilities. The development of such abilities during childhood itself has been of interest to researchers, although it is still not quite clear what exact processes trigger a shift in linguistic focus from the implied meanings of an utterance to the form of the utterance (Sutter and Johnson, 1990).

Bialystok (1986) claimed that the development of two skills is necessary for the growth of metalinguistic awareness: the control of linguistic processing in order to select specific linguistic information and the analysis of linguistic knowledge. The first relates to the ability to select and focus on a certain part of the linguistic material while ignoring the surrounding language environment. In the case of subject-verb agreement, this would mean the necessary step of identifying the sentence subject and the verb, or at least the function morpheme attached to the verb inflection (if it is a regular lexical verb and we assume non-lexical morphological processing). The analysis of linguistic knowledge, however, can only be achieved when the relevant structure is known explicitly. Again applied to the case of subject-verb agreement, this would mean for English-speaking children the explicit knowledge of the number of the sentence subject and that the verb morpheme *-s* marks the 3^{rd} person singular.

Previous research on grammaticality judgements in children has shown that age is the most reliable predictor of metalinguistic awareness, although language proficiency as measured by receptive vocabulary and sentence comprehension has a strong predictive value as well (e. g. Sutter and Johnson, 1990; Smith and Tager-Flusberg, 1982). A similar influence is found for working memory (McDonald, 2008b), possibly being related to the ability to select and shortly store parts of the linguistic material. The nature of the deviation from grammaticality that is to judged is highly relevant, with word order violations

(e. g. *man the* or *running is*) being easier to detect than for example the absence of a functional morpheme.[47] Finally, the form of the ungrammatical structure must be taken into account.[48] It makes a difference whether the ungrammatical form represents an existing, but incorrect, morphological form taken from the same morphological paradigm or whether it is made up from a non-existing form. In the following, I present previous findings on the metalinguistic judgement of subject-verb agreement and verb inflection. These findings relate to Experiment 1 of the present study, in which children's ability to detect subject-verb agreement violations is investigated using an eye tracking task. It will become evident that detection subject-verb agreement violations seems to be difficult for children when examined with a grammaticality task. Some explanations suggested by other researchers are presented at the end of this chapter.

4.4.2 Grammatical judgement of subject-verb agreement in English

McDonald (2008b) examined the role of age, working memory and phonological ability on the judgement of ten grammatical constructions in six to eleven year old children and adult control participants. Third person singular agreement was one of the grammatical structures in the test battery (others were for example word order, missing auxiliaries and determiners, regular and irregular plural). Working memory span and phonological ability for each age group were assessed separately, the former using a 'size measurement task' and the latter a 'gating task'

[47] Although this might well depend on the language under investigation, because French-speaking children had more problems detecting word order violations, presumably because French has freer word order within the nominal phrase (Kail, 2004).

[48] Smith and Tager-Flusberg (1982) for example found good performance in three and four year olds when asked to judge morphological endings. Collapsed over all morphemes, 22 % of the three year olds judged 9 out of 10 ' morpheme structrures' correctly, as did even 83 % of the four year old children. But this good performance can at least partially be explained by the incorrect alternative presented, which was (at least in the case of 3rd person singular inflection, totally ungrammatical (*he swim-s* vs. *he skate-st*).

(see McDonald (2008b) for precise descriptions on the tasks and the procedure). Children were divided into three groups (six to seven year olds, seven to nine and a half year olds and nine and a half to eleven year olds) with 22 or 23 children in each group. Additionally, 19 adults were tested as control participants.

Overall results showed that the ability to correctly judge sentences as grammatical or ungrammatical increased with age. The youngest group displayed lower performance than all the others, while the two older child groups still performed less accurately than the adult controls did. Thus, grammatical judgement of various structures does not seem to be at an adult level even by the ages of ten or eleven. However, a regression analysis revealed that age, working memory and phonological ability made significant contributions to the model. Therefore, working memory capacity and phonological ability played a role in the metalinguistic task performance above and beyond that played by age.

Third person singular inflection was found to be one of the last structures mastered. Even the oldest group of children (9;6 to 11 years) did not show adult mastery on this function morpheme (along with regular and irregular past tense). These difficulties cannot be attributed to an overall difficulty with bound verb morphology, because progressive *-ing* was only of medium difficultly. Additionally, performance on 3^{rd} person singular inflection was significantly influenced by working memory and phonological ability. Verbs were additionally controlled for phonological form, to determine whether they displayed a sub-syllabic realization of the 3^{rd} person singular inflection *-s* (e. g. *he eats*) or a syllabic realization of the *-s*-inflection (e. g. *he catches*). In the former case, phonological ability was found to be a significant predictor of performance, while in the latter case, no predictive value of phonological ability was found. This was interpreted as another proof that phonological ability influenced grammaticality judgements in children.

The influence of working memory on 3^{rd} person singular judgements might be explained by the fact that participants were required to keep multiple sentence parts in mind and had to check whether the constituents matched according to (person and) number value. The finding that working memory and phonological ability significantly influenced the judgement performance strongly suggests that factors other than grammatical knowledge per se are important in grammaticality judgement tasks (McDonald, 2008b).

Wulfeck, Bates, Krupa-Kwiatkowski, and Saltzman (2004) took the examination of metalinguistic judgements a little further. They tested not only the rate of correct and incorrect judgements, but additionally collected information on how long it took participants to make a judgement. Reaction time was considered to reveal processing costs. Wulfeck and colleagues wanted to gather information about sentence processing, rather than relying only on an off-line judgement about the sentence's well-formedness. Participants were tested on word order violations, agreement violations and omissions (see 4.21, 4.22 and 4.23 for examples) involving different parts of speech, determiners and auxiliaries. Agreement errors always involved errors of number.[49]

(4.21) *She selling is * books at the fair.*

(4.22) *She are * selling books at the fair.*

(4.23) *She selling * books at the fair.*

Judgements and reaction times were made via button-press. Thirty-four normally developing children in three age groups (7 to 8-years, 9 to 10-years and 11 to 12-years) were tested.[50]

Results revealed overall sensitivity to ungrammaticalities, as all normally developing participants performed above chance. Still, grammatical sensitivity was again found to increase with age. Interestingly, only the oldest group of children (eleven to twelve year olds) approached the sensitivity levels reported for

[49] But note that the subject-verb agreement match was tested on auxiliaries, not on inflected lexical verbs.

[50] Additionally, children with SLI and focal brain lesions were tested. Their results are not presented here, as we are mainly interested normally developing children's ability to detect subject-verb agreement errors. It shall only be noted that SLI-children displayed inferior performance in all conditions, but were particularly impaired in their ability to detect agreement violations.

college-aged subjects (Blackwell and Bates, 1995). Additionally, the position of the ungrammaticality in the sentence had been controlled and had a significant effect on the judgement performance. Children were better at detecting ungrammaticalities that occurred later in a sentence compared to earlier in the sentence. This finding is in line with results from Song et al. (2009) and Sundara et al. (2011), who both report enhanced processing and production of verb inflections in sentence-final compared to sentence-medial position. The authors attribute this position effect to the accumulation of linguistic information, which could be exploited by the children and supposedly helped them in detecting ungrammaticalities. Furthermore, children were better at detecting word order violations than agreement violations, with omissions falling right in between. Finally, accuracy scores and reaction times did not always yield the same results. Children were less sensitive to agreement violations, but they were just as fast at detecting them, when they did.[51] This rapid detection of agreement errors was interpreted as a further evidence showing that children processed agreement information in real-time. Very similar results had been obtained in an earlier study Wulfeck (1993). This one additionally documented that subject verb agreement was harder to detect than plural agreement between determiner and noun for six to seven year olds. Thus, it is not only processing of an agreement relation in general that seems to be challenging for children, but the type of agreement and the constituents involved play a significant role.[52]

Wulfeck and colleagues interpret their findings within the *Competition Model* of language processing and language acquisition (e. g. Bates and MacWhinney, 1989; MacWhinney, Bates, and Kliegl, 1984). Because this model will be referred to repeatedly throughout this work, it will here be presented in more detail.

[51] Note that only reaction times from correctly judged sentences entered analysis.

[52] But determiner noun agreement is lexical and local, while (English) subject-verb agreement is "less lexical and local", because subject and verb can be separated by long phrases (Lukyanenko and Fisher, 2010).

4.4.3 The Competition Model

The competition model is an interactive-activation framework especially designed to explain quantitative and qualitative performance variations within and across languages and over the course of language acquisition (e. g. Bates, Devescovi, and D'Amico, 1999; von Berger, Wulfeck, Bates, and Fink, 1996). In this model, listeners are thought to rely on a variety of *cues* for sentence processing, for example to determine the agent of a sentence. This example of 'deciding who has done it', i. e. children's use of linguistic evidence to perform role assignment, is a very frequently used test case for the competition model (MacWhinney et al., 1984; von Berger et al., 1996). The sentence stimuli presented in such studies represent converging or competing semantic cues like 'animacy' (e. g. contrast between animate and inanimate objects), syntactic cues like 'word order' (e. g. canonical SVO word order or non-canonical OVS word order), and morphological cues like 'agreement' (e. g. subject-verb agreement with the first or second noun). For example, in English noun-verb-noun sentences, the first noun is more likely to be the agent than the second noun (e. g. *The boy hit the girl.* vs. *The girl was hit by the boy.*) and animate nouns are more likely to be the agent than inanimate ones (e. g. *The boy pushed the ball.* vs. *The ball pushed the boy.*). Performance is then compared on sentences that have converging cues (e. g. animate first noun and inanimate second noun) and sentences that have competing cues (e. g. inanimate first noun and animate second noun). When additionally varying the type of cue, researchers can determine which cues listeners of a particular language rely on most.[53]

[53]Certain combinations of cues result in ungrammatical or (as phrased by the authors) 'semi-gramamtical' test sentences (e. g. *The cat the girl chased*). The proponents of the competition model make two claims to justify the use of ungrammatical sentences. First, an utterance like *the cat the girl chased* is considered to be a well-formed phrase-structure in English that can be expressed as part of a sentence containing a relative clause (e. g. in a sentence like *The cat the girl chased was really fast.*) (e. g. Bates et al., 1999). Second, they claim that if they obtain similar results within the same language with sentences that all grammatical and those that are only sometimes grammat-

An alternative test method is the grammaticality judgement task for various sentence structures and violation types, as performed by Wulfeck et al. (2004). One of the main findings is that listeners of English rely primarily on word order as a cue to agent marking (e. g. MacWhinney et al., 1984; Bates et al., 1999). This reflects the properties of the language, since word order tends to be very rigidly preserved in English. Thus, word order is thought to be a highly *valid* cue in sentence interpretation for English-speaking adults.

The concept of 'cue validity' is at the heart of the Competition Model. It refers to the information value of a given phonological, morphological, lexical, or syntactic form within a particular language.[54] The other important concept is that of 'cue cost', "which refers to the amount and type of processing associated with the activation and deployment of a given linguistic form, when cue validity is held constant" (Bates, Wulfeck, and MacWhinney, 1991, pg. 127). An example is the amount of memory that is required to store linguistic information and compute agreement relations within an sentence. Cue cost is also thought to be dependent on extra-linguistic factors, such as when a phonetically subtle cue is presented in a noisy environment or when further attentional demands for the listener are raised, thus when it is harder for the listener to access and use a cue. Simply put, the less salient a cue is, the more costly it is in processing (e. g. Leech, Aydelott, Symons, Carnevale, and Dick, 2007; von Berger et al., 1996; Wulfeck et al., 2004).

Subject-verb agreement cues are found to be of relatively low validity in English (since they are much less available than for example word order cues to determine the agent role), but they are also low in perceptual salience (e. g. Bates et al., 1991; Dick, Wulfeck, Krupa-Kwiatkowski, and Bates, 2004). Italian

ical, the same processing mechanisms should be assumed for grammatical and ungrammatical sentences. Therefore, they should be able to generalize effects found for ungrammatical structures as well (Gibson, 1992).

[54] Cue validity is often analyzed as two subcomponents: 1. 'cue availability', which refers to how often a cue is available when a certain sentence role has to be assigned and 2. 'cue reliability', which refers the question of how often use of this cue leads to the right answer (e. g. von Berger et al., 1996).

listeners, for example, are found to rely much more on subject-verb agreement than on word order in sentence processing (e. g. Bates et al., 1999), one of the findings that promotes the cross-linguistic differences in cue hierarchy and with this, cross-linguistic differences in sentence processing strategies. A different pattern is again found for German listeners, who rely much more heavily on agreement and animacy in sentence comprehension (MacWhinney et al., 1984).

Summing up, the simple logic behind the competition model is that sentences are interpreted by comparing and exploiting the linguistic cues that are present in them. Sentence comprehension is thus viewed as a statistical task, "where the development of 'rule-like' behavior occurs through a process of establishing the validity and reliability of competing linguistic cues" (Leech et al., 2007, pg. 795). This concept may be applied to the language acquisition process as well. It is assumed that the order in which form-function mappings are acquired will reflect the relative strength of the mapping, where the the most valid cue is acquired first.[55] Cue cost has an additional influence, since cues that are more costly are acquired later, even though they might be of relatively high validity (e. g. Dick et al., 2004). One of the advantages of the competition model is that it accounts for variation in the linguistic performance in adults and changes that are observed in the language processing abilities in children over time. This contrasts with many generative theories of language processing and language acquisition that focus rather on linguistic competence. In the latter kind of proposal, variation in grammatical performance is usually explained with performance factors such as working memory capacities or lexical retrieval issues (e. g. Clahsen and Felser, 2006).

Because the competition model emphasizes quantitative variation, it is very appealing when children's variable performance on different structures has to be explained (such as is found

[55] It has for example been found that English-learning children rely on word-order beginning by 2 years of age, although this increases markedly until the age of 5 (Bates et al., 1984), although others have found that English-learning children first, around the age of 3, rely on semantic strategies and only four year olds make primary use of word order (e. g. Bever, 1970).

by Wulfeck and colleagues in the grammatical judgement task, reported above Wulfeck et al., 2004). Since word order is a strong cue for English listeners, violations of word order are much easier detected by English-speaking children. Subject-verb agreement, on the other hand, is a relatively weak cue (e. g. MacWhinney et al., 1984; Dick et al., 2004), therefore children have much more difficulty detecting violations of subject-verb agreement (Wulfeck, 1993; Wulfeck et al., 2004; McDonald, 2008b).

4.4.4 Grammatical judgement of subject-verb agreement in Dutch

Dutch-speaking children have as well been tested on their sensitivity to subject-verb agreement using grammaticality judgement tasks.[56] Rispens and Been (2007) investigated the effects of phonological awareness and non-word repetition on the sensitivity to subject-verb agreement in Dutch-speaking children aged eight and a half. They found that normally developing children outperformed those with SLI and those with developmental dyslexia. Still, across all three groups of children, non-word repetition (thus, phonological working memory) was correlated with morphosyntactic sensitivity. It should be noted that the eight and a half year old normally developing children in this study performed at ceiling level in the grammatical judgement task, in contrast to the English-speaking children tested by Wulfeck et al. (2004) and McDonald (2008b). Rispens, Roeleven, and Koster (2004) presented another set of data which showed that Dutch-speaking children aged eight to nine performed at ceiling when asked to judge subject-verb agreement. The difference between the studies might be due to language-specific morphosyntactic properties of English and Dutch, in the sense that verb inflection and subject-verb agreement is a

[56] I am not aware of any published study examining German-speaking children's ability to classify sentences as grammatical or ungrammatical according to morphological endings. But Wenzlaff and Clahsen (2004) have tested German-speaking Broca aphasics and unimpaired adult controls on this subject.

stronger cue in Dutch than it is in English (Kilborn and Cooreman, 1987), or to some specific property of the test stimuli.[57] Nevertheless, the SLI- and dyslexic children tested by Rispens and colleagues performed at chance-levels, again indicating significant problems with subject-verb agreement in these populations.[58]

4.4.5 Summary of grammaticality judgement studies

Summarizing, judging the agreement between subject and verb seems to be a very challenging task for children. Even during the early school years, children perform significantly worse than adults on detecting ungrammaticalities arising by subject-verb agreement violations. Performance has been repeatedly found to correlate with phonological working memory (Rispens and Been, 2007), general working memory (McDonald, 2008b) and phonological ability. Interestingly, subject-verb agreement detection has been found to be difficult for adults under working memory load (Blackwell and Bates, 1995; McDonald, 2006).

One explanation put forward by Wulfeck and colleagues relates to the low informative value of verb inflection as a cue to sentence processing. This explanation is mainly based on the finding that word order violations are easier and earlier (agewise) detected in children than subject-verb agreement violations. English verb inflection is assumed to be a cue of little validity, because it is usually a redundant due to number marking on the noun phrase, and of high cost, because the inflectional affix is not easy to perceive due to its phonological form.

[57] Rispens and colleagues provided children with three types of ungrammatical variations: person-mismatch (1. 3^{rd} person singular subject and verb inflected for 1^{st} person singular) or two variants of number mismatch (2. 3^{rd} person singular subject and verb inflected for 3^{rd} person plural or 3. 3^{rd} person plural subject and verb inflected for 3^{rd} person singular). To differentiate between the first contrast and the third contrast is not possible in English, due to the impoverished morphological system.

[58] There was no task-specific problem found in the SLI- and dyslexic children, because they all performed well on judging the grammatical well-formedness of control sentences (Rispens et al., 2004; Rispens and Been, 2007).

Therefore, it is acquired rather late by typically developing children (later than word order regularities) and consistently more difficult to use in sentence processing for English SLI-children (Wulfeck et al., 2004).

McDonald (2008b) emphasizes that working memory load plays a role in grammatical judgement mastery and relates the specific problems found with verb morphology to this. According to her findings, working memory plays a significant role only in those constructions that require participants to keep in mind multiple sentence parts and check if information across these parts is consistent, which is certainly the case for subject-verb agreement.

Another possible explanation targets the amount of acoustic information of the verb morphology tested in subject-verb agreement violations. This could explain why children were better at detecting violations of the present progressive *-ing* than violations of 3^{rd} person singular *-s*, and it would explain the connection found between phonological ability and grammatical judgements of singular verb inflection (McDonald, 2008b), or non-word repetition and judgements of subject-verb agreement (Rispens and Been, 2007).

One aspect that leaves one puzzled is the huge asymmetry found between the studies on early sensitivity to verb agreement morphology (Soderstrom, 2002; Soderstrom et al., 2007; Nazzi et al., 2011) and the very late detection of subject-verb agreement as found in the grammaticality judgement tasks. How is it that 16- and 19-month old infants seem to be sensitive to subject-verb agreement, but normally developing school-agel children are hardly able to tell whether a verb form matches a sentence subject? It is unclear whether task-effects can be the whole story. One and a half year old infants might be at an advantage, because they can focus more on the pure linguistic structure without worrying to much about meaning, especially in an HPP-setting. This might enable them to employ a non adult-like rule they have derived from the input and that can best be applied to superficial properties of sentences. Older children might be at a disadvantage because they consciously need to block out the sentence meaning to concentrate on the grammatical judgement task. A similar point is made by Naigles (2002) and Soderstrom (2008), who attribute young children's impressive sensitivity to distributional regularities in the input to the fact that they can analyze the incoming speech without having to pay attention to the meaning, at least in an HPP task. Recent findings by Sundara et al. (2011) provide further evidence that a referential context, which might prompt children to focus rather on the meaning of an utterance than on the form, prevents even slightly older children from detecting subject-verb agreement violations in simple sentences. Thus, a combination of factors might account for the differences found between infants' ability to detect agreement violations and older children's inability to do so: the task demands, the required metalinguistic abilities, and the different ways of analyzing the linguistic input.

5 Comprehension of verb inflection

In this chapter, I will present research on children's use of verb inflections in sentence comprehension. The critical point of the studies presented here is that children actually have to infer meaning from the information provided by the verb inflections. First, I will explain commonly used methods to assess children's language comprehension. Then I will present previous studies that have examined children's use of verb inflection to infer the number of the sentence subject (in English, Spanish and French), most of which reveal very late comprehension abilities. The chapter will close with a review of factors that seem to influence comprehension of verb inflection.

Concerning the receptive side of language, researchers are not only interested in what children know about the well-formedness of utterances, but also in when children are able to recover the intended meaning of an utterance (i. e. to understand what a particular sentence actually means). Without the process of comprehension, and its coordination with language production, speakers would be unable to use language to communicate and to infer intentions (Clark and Hecht, 1983). The impression of generations of parents and care takers has been that young children understand more than they produce. This long favored view has been confirmed in many observational studies and in experimental research (e. g. Fraser, Bellugi, and Brown, 1963; Naigles, 2004). Children are usually found to understand more content words than they produce at a given point in development (e. g. Goldin-Meadow et al., 1976; Reznick, 1990) and to understand sentence constructions that are not yet present in their language production (e. g. de Villiers and de Villiers, 1973b). Comprehension of function words has also been found to precede their use in productive speech (e. g. Gerken et al., 1990; Shady and Gerken, 1999). But interestingly, the findings in the literature on children's sentence comprehension are far from being consistent. Recent research shows that children often misinterpret sentences and do not show adult-like comprehension performance, for example for pronouns, restrictive modifiers or relative clauses for quite some years (e. g. Trueswell, Sekerina, Hill, and Logrip, 1999; Sekerina et al., 2004; Hurewitz, Brown-Schmidt, Thorpe, Gleitman, and Trueswell, 2000). Relying on verb inflection seems during sentence comprehension seems to be challenging for children as well.

5.1 Methods for testing comprehension

It is important to evaluate and choose the appropriate method when doing research on language comprehension, especially when working with children. There has been a wealth of publications dealing with methods in language acquisition research, see e. g. Blom and Unsworth (2010), McDaniel, McKee, and Cairns (1996) or Crain and Thornton (2003). The impact of methodology can be seen in both language modalities, and was already explained in Chapter 3.

5.1.1 Picture selection task

The most common method seems to be the *picture selection task* or *picture pointing*. In such a task, participants are presented with an array of pictures (usually two or four) and an auditorily presented word or sentence. Participants are asked to point to the picture that best matches the verbal stimulus. This task is very often used in language assessment batteries testing for language impairments (e. g. Fox, 2006). It has been used to assess comprehension abilities for nearly all types of linguistic constructions, no matter whether they pertain to phonological, lexical, semantic, morphological or syntactic knowledge. The selection of the distractor picture is critical. If one wants to examine children's knowledge of vowel length as a phonological property, one might present the pictures of a sheep and a ship along with the verbally presented noun 'sheep'. If one rather wants to investigate children's lexical knowledge, one should contrast the pictures of a sheep and another animate being, creating a semantic contrast. Of course morphosyntactic knowledge can easily be assessed using the classical picture selection task as well, for example by presenting a picture on which Figure A is performing some causal action on Figure B (e. g. *Big Bird is gorping Cookie Monster*) and a second picture on which Figure A and B are performing the same non-causal action (e. g. *Big Bird is gorping with Cookie Monster*). Pointings to the correct picture after hearing a sentence like 'Hey look, Big Bird is gorping with Cookie Monster' shows that children can use the syntactic context within which the verb occurs to determine transitivity, i. e. whether or not a verb takes a grammatical object (e. g. Gleitman, 1990).

Gerken and McIntosh (1993) and Shady and Gerken (1999) used the picture selection task in a slightly different way, namely to examine children's sensitivity to syntactic well-formedness. They presented children with pre-recorded sentences in which a target noun was preceded by a grammatical article (*the*), no grammatical morpheme, an ungrammatical auxiliary (*was*) or a pseudo word (*gub*). Children were asked to point to pictures in a picture book. It was found that children were significantly better at selecting the correct picture after hearing sentences containing a grammatical determiner compared to sentences containing an ungrammatical determiner. They actually found an interesting decrease in performance, with the no morpheme condition being second easiest, the ungrammatical auxiliary being third easiest and the pseudo-word condition being the hardest. Thus, picture selection can as well be used to assess children's knowledge of sentence (or at least phrase-internal) well-formedness. Gerken and Shady (1996) provide a detailed description and discuss necessary considerations concerning the picture selection task. Studies examining the comprehension of verb inflection using the picture selection task will be presented in Chapter 5.2. This method is also employed in the present study, testing German children's use of verb inflection in sentence comprehension.

5.1.2 Act-out task and manual search task

Another common method is the *act-out task*. Here, the experimenter reads (or plays a pre-recorded version of) a sentence to a participant, who then acts out the interpretation of the sentence with a set of figures and props that are provided. Of course, the participant's interpretation of the sentence is what the researcher is mainly interested in. Act-out tasks have most commonly been used to assess morphosyntactic knowledge in children, e. g. the comprehension of relative clauses and sentences with pronouns and anaphors (e. g. Chien and Wexler, 1990; Friedman and Novrogodsky, 2004). A detailed discussion on the act-out task can be found in Goodluck (1996) and Crain and Thornton (2003).

A similar method that is used especially when number comprehension and conceptual number knowledge is assessed is the *manual-search paradigm*. In this task, the experimenter puts either one or multiple (usually three or four) balls in a box and hands the box to the child. The child, who saw how many balls were put in the box, is allowed to retrieve one ball. In the multiple-ball condition, the remaining balls are then removed surreptitiously by the experimenter. At this point, the actual test phase starts, because for the next 10 seconds, the experimenter measures the amount of time the child searches for another ball in the box. Correct number comprehension is inferred when the child searches longer in the multiple-ball condition than in the one-ball condition. Number discrimination, number comprehension and quantifier acquisition in children (non-verbal and verbal) have been successfully tested with this paradigm (e. g. Barner et al., 2007; Li et al., 2009; Wood et al., 2009).

5.1.3 Preferential looking paradigm

An alternative method that taps into children's comprehension of linguistic structures is the *intermodal preferential looking paradigm* (IPLP) or *visual world paradigm* (VWP). This method has the advantage that it can focus either on the *product* of children's lexical and syntactic comprehension, e. g. the final interpretation of a certain syntactic structure (Naigles, Bavin, and

Smith, 2005) or on the *process* of comprehension, e. g. the continuous processing in word recognition and sentence comprehension (Swingley et al., 1999; Fernald, Swingley, and Pinto, 2001). Process and product of comprehension can also be investigated in a combined effort, for example when investigators track children's eye gaze towards a variety of possible referents while a sentence is presented (on-line measure targeting the process) and additionally ask the children to perform a an task with one of the referents after the sentence is finished (off-line measure targeting product of comprehension process) (Trueswell et al., 1999; Hurewitz et al., 2000).

In the IPLP, a child is seated in front of two monitors (or one monitor with a display of two pictures) and a linguistic stimulus is presented that matches only one of the displays shown on the screen. The infant's 'task' is simply to look at the screen and listen to the verbally presented material. The control of the distractor stimulus is as important as in the picture selection task, with the critical feature that is under investigation needing to be reflected in the choice of target and distractor picture. Additionally, the brightness, size and 'interestingness' of the pictures should be matched. (Otherwise, if the target picture were much more appealing simply for visual reasons, preference for this target picture could not be related to linguistic processing.) The inclusion of a baseline phase is another very important way to control for any initial visually based preferences (e. g. Legendre, Barriere, Goyet, and Nazzi, 2010; Mani and Plunkett, 2010a). An extensive explanation of the IPLP method and it's more processing-related extension *looking while listening* (LWL) is found for example in Hirsh-Pasek and Golinkoff (1996) and Fernald, Zangl, Portillo, and Marchman (2008).

Researchers have examined almost all levels of linguistic processing in children using the IPLP, including phonological processing (e. g. Mani and Plunkett, 2010b; Swingley and Aslin, 2000), word learning and word comprehension (e. g. Ma, Golinkoff, Houston, and Hirsh-Pasek, 2011; Meints, Plunkett, and Harris, 1999), word order comprehension and syntactic processing (e. g. Golinkoff et al., 1987; Seidl, Hollich, and Jusczyk, 2003), the processing of function words (e. g. Lew-

Williams and Fernald, 2007; Van Heugten and Shi, 2009), the comprehension of pronouns (e. g. Bergmann et al., 2011) and number comprehension (e. g. Kouider et al., 2006).

In the VWP, children are often presented with real objects rather than pictures and their eye gaze is tracked while they are listening to sentences and then performing an action, like touching a specific object or changing the position of an object (Trueswell et al., 1999; Hurewitz et al., 2000). This methods taps into children's real-time sentence processing (like the LWL), but it enables the researcher to compare a final interpretation and the way to get there, for example the alternatives that were taken into account (e. g. Snedeker and Trueswell, 2004; Choi and Trueswell, 2010; Sekerina et al., 2004). See for example Trueswell (2008) for an description and evaluation of the method.

In addition to measuring comprehension, IPLP has also been used as a technique to investigate children's sensitivity to phonological or syntactic well-formedness. When presenting children with correctly or mispronounced object labels, Swingley and Aslin (2000) found that children were better and faster at finding a target picture on a visual display when the word labeling the object was pronounced correctly. From this, they concluded that children's representations of familiar words are phonetically well specified, because otherwise the children should not be influenced by slight changes in the words phonological structure (see Mani and Plunkett (2010b) for similar findings). Zangl and Fernald (2007) and Kedar et al. (2006) have worked along the same lines to examine children's sensitivity to syntactic well-formedness. Because these studies have worked as guidelines for Experiment 1, they will now be presented in more detail.

Kedar and colleagues aimed to examine young children's ability to detect ungrammaticalities that were caused by the manipulation of one functional element in a sentence: a nominal determiner. Therefore, they directly followed Gerken and McIntosh (1993), who had tested this using a picture selection task in a group of two year olds. Additionally, Kedar and colleagues were interested in whether young children would use a determiner to establish a referent for a noun phrase. Therefore, they presented 18- and 24-months old English-learning children with two pictures per trial along with a verbally presented sentence that contained either the grammatical determiner *the* (e. g. *Can you see the ball?*), another known function word *and* (e. g. *Can you see and ball?*), a nonce determiner *el* (e. g. *Can you see el ball?*) or no determiner (e. g. *Can you see ball?*). The 24-month old children directed more first looks to the target picture in the grammatical condition compared to the other function word and the nonce determiner condition. The younger children showed a tendency to look more to the target after hearing a grammatical determiner, but this was less consistent. Additionally, children of both groups were found to orient faster towards the matching picture after hearing the correct determiner *the* than after hearing a sentence with a nonce determiner (*el*), another function word (*and*) or no function word. Interestingly, no effect of the determiner manipulation was found on the proportions of children's overall looks towards target in the test phase. The findings were interpreted as providing evidence that 24-month-olds not only detect an ungrammaticality that is caused by a missing or wrong function word, but use a grammatical determiner to establish reference. It has to be kept in mind that 24-month-olds usually omit determiners in their own production. Therefore, the findings are another testament to the ability of young children to detect and use function morphemes in the acquisition and the comprehension of language.

The design and the findings of the experiments reported by Zangl and Fernald (2007) were very similar. The first experiment revealed that 18-month old English learning infants were faster and more accurate in identifying a familiar noun when the sentence contained a grammatical determiner (*the car*) than when it contained a nonce determiner (*po car*). No effect was seen when grammatical trials were compared to trials in which no determiner preceded the noun (*Where's car?*). Interestingly, children aged 36 months, who already produced determiners consistently in their own speech, did not seem to be affected, since they were as fast and as accurate (considering their proportion of looks) in finding the target picture. A second experiment was conducted to examine children's determiner processing when pseudo words were used. Therefore, 34-month old children had to learn new words prior to testing and were then tested on familiar and unfamiliar nouns, presented either with a grammatical or a nonce determiners. As it turned out, older children showed the same effects of processing disruption that the the younger ones had displayed when tested on familiar nouns, but only in the conditions testing the identification of the newly learned words. Zangl and Fernald (2007) interpret their findings as indicating that linguistically less advanced children rely on surface regularities and lexical familiarity in sentence processing and therefore experience disruption when regularities are not met. The older and linguistically more advanced child can more easily ignore ungrammatical determiners and still find a target picture quickly and reliably as long as no additional demands burden sentence processing, as in the case of the comprehension of newly learned words.

Thus, children's looking behavior has been shown to be a good indicator of their sentence processing and how much this is disrupted when a sentence structure violates their expectations. Additionally, methods that rely solely on infant's looks have, according to the literature, various advantages. First, it is possible to test much younger infants on their comprehension and processing of various linguistic structures (e. g. Yoshida et al., 2009; Fernald et al., 2001). Second, it provides more measures than a mere correct/incorrect distinction, thereby providing information about language processing in children (e. g. Fernald et al., 2008). Third, it does not require an overt response from children, which can be a particular advantage when testing younger children or more demanding linguistic structures.

5.2 Experiments on the comprehension of verb inflection

Relative to the wealth of studies that examined the production of verb inflection, studies regarding the comprehension of these grammatical morphemes are very sparse. But studies that investigated whether children can make use of the information provided by the inflectional morphemes for the identification of subject number actually reflect a more direct approach to the question of when children are sensitive to the different morpho-syntactic categories involved in verb inflection. In the following, I will present studies on the comprehension of verb inflection in English, Spanish and French. To date there is no study examining verb inflection comprehension in German.[59]

5.2.1 Comprehension of verb inflection in English

Johnson et al. (2005) conducted a picture selection task in which verb inflection was the only cue to subject number. To achieve this, they used verbs that began with an *s*-consonant cluster (e. g. *swim, see*), which was coarticulated with the plural *-s* on the noun (see 5.1 and 5.2 for examples). Of course only regularly inflected nouns were used as subjects. Each sentence was presented with two pictures that either showed one or two actors performing the action denoted by the verb. Three to six year old English-speaking children were tested.[60]

(5.1) The duck swims on the pond.

(5.2) The ducks swim on the pond.

Only the five and six year-old children, but not the three and four year olds, performed above chance level. Importantly, all children had passed practice trials in which number was either marked by forms of auxiliary BE (*is*, *are*) or by overt plural marking on the subject noun. Thus, low performance could not be attributed to the task or the number contrast itself. But even the older children performed well below 100 % correct across all

[59] Expect the work presented in this thesis and the related papers (Brandt-Kobele and Höhle, 2010, 2011).

[60] The three year olds were purposely chosen to overlap with the age range in which 3^{rd} person singular *-s* production is typically still variable. The four, five and six year olds were expected to reliably produce the morpheme under investigation in all obligatory contexts, especially as all children came from European-American families and spoke main stream American English (MAE).

conditions and therefore significantly below adult ceiling performance. In the plural condition even the six year olds did not continue to display above chance-level performance. Percentage of accuracy is listed in Table 5.1.

Age group	singular	plural
three year olds	52.38 (*22.34*)	41.9 (*21.82*)
four year olds	64.44 (*21.68*)	46.67 (*22.34*)
five year olds	78.67* (*27.74*)	61.33 (*35.83*)
six year olds	78.89* (*27.84*)	53.33 (*34.30*)

Table 5.1: Means (*Standard deviations*) by age for percent accuracy of 3^{rd} person singular *-s* comprehension (* significantly better than chance), taken from Johnson et al. (2005).

Johnson et al. (2005) concluded that, especially for the three and four year old children, the English 3^{rd} person singular *-s* inflection is not a transparent marker for subject number. They discussed two possible explanations for children's late and rather low comprehension performance. First, they suggested that children might have acquired the rule of 3^{rd} person singular *-s* insertion without recognizing that this rule involves the number contrast. Since this morpheme does not exclusively and reliably represent singularity (it also marks verbs for present tense and 'verb-ness' (de Villiers and Johnson, 2007) and 1^{st} and 2^{nd} person singular verbs are not marked by *-s*), the authors claim that children could have acquired a number insensitive learning device following the rule (stated in 2.6, repeated below).

V+/Ø/ → V+/s/ following he, she, it, NP+/Ø/

The authors claim that such kind of rule would rather be expected in very young children, for example in the age group tested by Soderstrom (2002). Especially the notion of 'NP+/Ø/' (meaning NP without *s*) is hardly a reliable linguistic rule, as argued by Johnson and colleagues.[61] The authors additionally call on the problem that singularity as marked by *-s* is part of a

[61] In this thesis, I want to assert that children might very well operate with this kind of rules prior to learning adult-like abstract rules which incorporate

more complex feature, namely 3^{rd} person singular, suggesting that this one part of the feature complex may not be available to children in comprehension. Because of this need to access only one part of the feature complex, the authors assume a metalinguistic dimension to the comprehension task, and metalinguistic knowledge is usually testable after the age of four years (see Chapter 4.4).

As a second explanation, Johnson et al. (2005) call on Chomsky's notion of agreement. In the Minimalist Program it is argued that verbs move to INFL to check person and number features, but that those do not survive to the representation of Logical Form (LF), once they have been checked (Chomsky, 1995). According to this, agreement morphemes would be purely grammatical, but not meaning bearing.[62] This explanation presumes that abstract explanatory frameworks for grammatical representation actually reflect the psychological reality of sentence processing, a view that is not shared by everybody in the linguistic community (e. g. Bates, 1998). Why the uninterpretable inflectional features on the verb couldn't still be used by the performance systems to infer the interpretable features of the noun (and thus the semantic number of the subject) is not obvious. This explanation therefore makes one wonder how adults should be able to interpret the number of the sentence subject when only verb inflectional information is present. This scenario is certainly not very common in English (Brown, 1973), since almost all nouns are marked for number (except *deer*, *fish*, *sheep* etc.), but perfectly common in pro-drop languages like Spanish, which allow subject-less sentences.

Overall, Johnson et al. (2005) hypothesized that the 'poverty' of the English present tense agreement system, in which the 3^{rd} person singular is the only morphological form that is overtly marked, may be relevant for the children's ignorance concerning the information provided by the inflectional ending. John-

the notion of grammatical features. This view is not shared by Johnson et al. (2005).

[62] Clahsen and colleagues use a very similar argument to explain the low production rate of verb inflections in SLI-children (Clahsen, Bartke, and Göllner, 1997; Clahsen, 2008).

son and colleagues do not explicitly call on the notion of 'cues' as used in the Competition Model, but their explanation is very similar to the one offered by Wulfeck (1993). A very similar explanation is put forward by Leonard, Caselli, Betolini, McGregor, and Sabbadini (1992) to account for the problems SLI-children have regarding English morphology.

Very similar findings were obtained by Leonard and colleagues. They investigated this issue in English-speaking SLI-children and normally developing controls (Leonard et al., 2000). The SLI-children aged 4;3 to 5;7 performed at chance-level in the picture selection task tapping the comprehension of 3*rd* person singular *-s*. The normally developing control children, aged 4;0 to 5;7 performed correctly 71 % of the time, which was significantly better than chance, but still far from perfect. Interestingly, the control children produced the 3*rd* person singular verb inflection correctly in over 90 % of the obligatory contexts found in their spontaneous speech. A second experiment contrasted the availability of number cues in the verb inflection only with redundant number cues in the sentence subject plus the verb inflection. The younger control group in this second experiment, aged 4;0 to 5;2, performed at chance-level when 3*rd* person singular inflection was the only cue to subject number, but better when subject and verb were number marked. The older control group, aged 4;6 to 7;2, performed above chance in both conditions.[63]

Thus, Leonard and colleagues found late comprehension of when verb inflection morphology should cue subject number, just as Johnson et al. (2005). Especially puzzling is the fact that comprehension of verb morphology seems to lag behind production, both in normally developing children as well as in SLI-children. The authors call this a 'paradox' and try to explain this pattern of production being superior to comprehension with the 'level of knowledge' required for the production and comprehension tasks. More precisely, they claim that 'the sense that *runs* is a better fit with the subject *girl* whereas *run* is a better

[63]Test sentences with copula verbs elicited better performance in all experiments, groups and conditions, thus the late comprehension cannot simply be attributed to task effects.

match with *girls* might develop prior to knowledge that *runs* 'means' singular and *run* 'means' plural.' (Leonard et al., 2000, pg. 479). The authors do not clarify what kind of knowledge this 'sense' relies on, and whether this should be considered grammatical or not.

This explanation put forward by Leonard and colleagues is very similar to the claim that children have a non adult-like rule that can be applied in production but does not help children in semantic comprehension. According to this rule, children add an *-s* affix to the verb stem because the NP girl does not have an *-s* marker. This gives rise to correct production (when the noun does not end with an *-s* and does not have irregular plural marking) but does not allow the children to infer the number value of the sentence subject (see 2.6). The 'knowledge' that a singular inflected verb refers to a singular entity and a plural inflected verb refers to multiple entities (as referents of the subject) might develop later in the form of the adult-like rule (see 2.5) or additional rules that state that a NP with an *-s* refers to plural entities.

To date, a handful of preferential looking studies have investigated children's use of verb morphology in sentence comprehension. Lukyanenko and Fisher (2010) presented English-speaking three year olds with pictures showing either one or two familiar objects and 'informative' sentences in which a number marked auxiliary occurred before the critical noun phase (e. g. *Where* are *the good cookies?*). Additionally, children heard 'uninformative' sentences without such an auxiliary (e. g. *Can you find the good cookies?*). Children shifted their eye gaze faster from the distractor to the target picture in the informative compared to the uninformative trials (although this effect stems solely from the plural trials). Therefore, English-speaking three year olds were found to be able to expect a plural noun when hearing the auxiliary 'are'. A very similar procedure was used by Grüter and Fernald (2011) to test 24-months old English-speaking children. They found as well that children were able to use a plural marked auxiliary (and/or a demonstrative) to create an expectancy for the upcoming noun. No effects were found for the singular marked trials. Overall, English-speaking children around ages two and three seem to be able to use the number information on an auxiliary verb in sentence comprehension. At least they seem to build up the expectation that a plural marked noun will be presented which refers to a picture with multiple entities.

Notably, these two IPLP-studies test children's processing of auxiliaries. Auxiliaries are certainly inflected, but it is very likely that these high-frequency verb forms are stored and processed differently than are English main verbs with regard to inflectional information (e. g. Wilson, 2003; Phillips, 1995). Soderstrom (2002) investigated whether English-speaking toddlers were able to infer number information from *-s*-inflected verbs and nouns in two IPLP-studies. The 17- and 23-months old children were presented with pictures showing one or two objects and heard sentences like *The balls roll* or *The ball rolls*. If children were able to use the verb inflection information in sentence comprehension, they should look more to the 2-object picture when hearing a sentence with a plural inflected noun and more to the 1-object picture when hearing a sentence with a singular inflected verb. None of this was found. Children overall looked more to the 2-object picture, but auditory input did not have an influence on their eye gaze. The second experiment was designed to overcome flaws in the procedure, but again, no evidence was found that children understood the connection between inflection morphology and sentence meaning. Soderstrom concluded that children at the end of their second year do not have a semantic understanding of the number marker *-s*, "let alone an understanding of the difference between nominal and verbal *-s* inflection" (Soderstrom, 2002, pg. 120).

One critical aspect of verb inflection, at least in non-agglutinating languages like English, German and Spanish, is that it does not only code number information but is part of a feature complex that additionally codes tense, person and aspect (Johnson et al., 2005). Two studies have examined whether children were able to use the tense and aspect information provided by the English 3^{rd} person singular *-s*. De Villiers and Johnson (2007) investigated whether four to six year old children were able to infer a generic reading in contrast to a past tense reading from the *-s*-inflection attached to the verb (e. g. *Who just cut the bread?* versus *Who just cuts the bread?*). They presented a simple story along with pictures which provided the information necessary to answer the questions. They found that four, five, and six year old children were not able to answer the questions correctly, thus to differentiate between the past tense reading of a verb (e. g. *he cut*) and the generic reading of the same verb (e. g. *he cuts*). This was true both for children acquiring Mainstream American English (MAE) and for children acquiring African American English (AAE), in which verbal *-s* is mostly omitted.

Similarly, Beyer and Hudson-Kam (2009) tested whether English-speaking children were able to use the verb inflection *-s* as a marker for present tense. They compared this to the comprehension of *-ed* as a marker for past tense. Six and seven year old children were tested with a picture selection task and the eye-tracking paradigm. In the picture selection task, children of both age groups performed well on sentences containing the past tense inflection *-ed*, with more than 80 % correct. For sentences containing the present tense marker *-s*, a significant effect of age was found, since the six year olds showed chance-performance and only the seven year olds showed adequate comprehension (about 85 % correct). Thus, when tested with a classical picture selection task, the six year old children seemed to be able to use *-ed* as a morphological marker for tense, but not *-s*.

To examine the possibility that six year olds were sensitive to the presence of verbal *-s*, but that such sensitivity could not be tapped using a picture selection task, the authors additionally conducted an eye-tracking task using the same verbal and visual material. They found that six and seven year old children showed very similar eye-gaze patterns. The critical difference was that the six year olds were slower at finding the target picture in sentences containing a verbal *-s* than the seven year olds were. Interestingly, the authors analyzed correctly pointed trials separately from incorrectly pointed ones. With this procedure, they found correct comprehension of *-s* in the six year olds in the correctly answered trials (but about 600 ms later than the seven year olds). In the incorrectly answered trials, the children mostly stuck to the picture that they had looked at from the very beginning of the trial and seemed unable to revise their initial decision. Overall, the authors conclude that six year old English-speaking children are sensitive to the presence of the verb inflection *-s* and that they can use this as a marker for present tense, but that picture selection task is not an adequate measure to detect this sensitivity. Beyer and Hudson-Kam (2009) interpret their findings in the framework of the Competition Model and attribute the six year olds' difficulties in verbal *-s* comprehension (as evident in the picture selection task) to processing difficulties. They claim that the verb inflection *-s* is not a strong cue for children in sentence processing and that offline tasks underestimate children's developing knowledge.[64]

5.2.2 Comprehension of verb inflection in Spanish

The assumption that the impoverished English agreement system is the decisive factor for English-speaking children's difficulties when using verb inflection in sentence comprehension was tested in a language with a rich morphological system, namely Spanish. Pérez-Leroux (2005) conducted a study with Spanish-speaking children aged three to six years. In contrast to English, Spanish has a rich and robust set of verb inflections with specific inflectional endings for all person and number forms of a verb.[65] The children in this study came from the Dominican Republic. Dominican Spanish has a very high rate of deletion of final *-s* on plural nouns, therefore plural marking on nouns is a very unreliable cue to number. This means that number (of the sentence subject) is primarily recoverable from the verb rather than from the nominal. According to Pérez-Leroux

[64]The same explanation was suggested by Wulfeck (1993) for grammaticality judgement tasks and by Leonard et al. (1992) for SLI-children.

[65]Spanish present tense verb morphology for e. g. 'hablar' (*to speak*): 1^{st} SG: habl-o, 2^{nd} SG: habl-as, 3^{rd} SG: habl-a, 1^{st} PL: habl-amos, 2^{nd} PL: habl-ais, 3^{rd} PL: habl-an (for -er and -ir verbs, only the first vowel is changed, but not in 1^{st} SG).

(2005), this makes this Spanish dialect an ideal test case for the interpretation of verb inflection morphology and an ideal contrast to English. Perez-Leroux adopted the material and the procedure used by Johnson et al. (2005), so children were tested with a picture selection task. Perez-Leroux avoided the confounding role of nominal inflection by using subject-less sentences, which are grammatical in Spanish, due to its property of being a pro-drop language. In these sentences the verb ending was the only cue to number (see 5.3 and 5.4 for examples). She included sentences with a full subject DP as a control condition (see 5.5 and 5.6 for examples).

(5.3)		*Nada*	*en*	*el*	*charco.*
	(The duck)	swim-3SG	on	the	pond

The duck swims on the pond.

(5.4)		*Nadan*	*en*	*el*	*charco.*
	(The ducks)	swim-3PL	on	the	pond

The ducks swim on the pond.

(5.5)	*El*	*pato*	*nada*	*en*	*el*	*charco.*
	The	duck	swim-3SG	on	the	pond

The duck swims on the pond.

(5.6)	*Los*	*patos*	*nadan*	*en*	*el*	*charco.*
	The	ducks	swim-3PL	on	the	pond

The ducks swim on the pond.

The comprehension results from the Spanish-speaking children (on the test sentences like 5.3 and 5.3) were remarkably similar to those from the English-speaking children. This was despite the fact that Spanish has a robust agreement system and that the dialect under investigation has a robust phonologically reliable agreement system only on the verb and not on the noun. While three and four year old children's responses did not differ from chance performance in either number condition, the five and six year old children reached performance levels better than chance at least in the plural condition (see Table 5.2). In this sense, the data does not support the assumption that the 'poverty' of the

Age group	singular	plural
three and four year olds	52	45
five and six year olds	50	67*

Table 5.2: Mean percent correct for singular and plural pro drop sentences (* significantly better than chance), taken from Pérez-Leroux (2005).

English verb morphology is the reason for the findings by Johnson and colleagues.

Pérez-Leroux (2005) proposed an alternative explanation for the apparent inability of English- and Spanish-speaking children to use verb morphology to infer the number of the sentence subject. If children for example applied a generic reading to the test sentences (e. g. *Ducks swim on a pond* as a generic description of what animals of this kind regularly do), both pictures fit the utterance equally well. Perez-Leroux further claimed that generic readings are the semantic default (e. g. Gelman, Goetz, Sarnecka, and Flues, 2008).

The Spanish comprehension data, as the English one, alludes to a significant gap between comprehension and production, with production preceding comprehension. Perez-Leroux explains this gap with a language acquisition theory that assumes two 'layers' of acquisition, of which one is distribution and the other one is syntax-semantics mapping, with the second layer developing later. Thus, 'a child could attain surface distributions that appear grammatical, but maintain pockets of semantic underspecification' (Pérez-Leroux, 2005, pg. 10). Concerning the acquisition of number, this would mean that would have acquired the morphology and syntactic distribution of number marking, but would lack the complete knowledge of where number marking is actually interpretable in the language (and not uninterpretable, as in generic expressions or dependent plurals). The fact that generic readings exist and that they make number information uninterpretable is thought to be one of the main reasons of why subject-verb agreement is in such a 'pocket of semantic underspecification'.

Miller and Schmitt (2009) provided indirect support for the late comprehension findings by Pérez-Leroux (2005). They tested Spanish-speaking children from Chile to investigate the

'Variability Delay Hypothesis' (Miller, 2007). While noun plural marking is variable in Chilean-Spanish, verb morphology is applied consistently. Therefore, Spanish-speaking children from Chile were expected to comprehend verb morphology earlier than noun morphology marking singular and plural. Thus, children aged 4;5 to 6;0 (mean age 5;1) were tested in a picture selection task. Results showed better than chance-performance in comprehension (77 % correct plural responses), but performance was still well below adult control subjects. Thus, Chilean children were able to use verbal agreement in comprehension tasks before they could use plural morphology in the noun phrase, thereby providing evidence for Miller's hypothesis. One has to keep in mind though that children were already five years old and still performed worse than adults, thereby supporting the findings by Pérez-Leroux (2005) that inflection morphology on Spanish main verbs is comprehended late.

With regard to the different kind of grammatical knowledge that might underly children's processing of verb inflections, the English as well as the Spanish data rather seem to support the notion that children rely on a non-adult like rule until the age of five. A similar point is made by Pérez-Leroux (2005) although she rather attributes the problem to the syntax-semantic interface. Leonard et al. (2000) claim that children might have a 'sense' that *-s* better fits NPs without *-s*, which could be stated in a rule like presented in 2.6, but lack the knowledge that these forms relate to the semantic distinction between one and more than one. If one claims that children nevertheless have adult-like knowledge of verb inflections early on, severe performance limitations have to the assumed to explain how children might not be able to use this knowledge in comprehension experiments until the age of five or six.

5.2.3 Comprehension of verb inflection in Italian

Very recent research has examined whether Italian-speaking children were able to use verb morphology in sentence comprehension. Dispaldro and Benelli (2012) refrained from using a picture selection or preferential looking task because they claim that a picture selection task underestimates children's ability to comprehend singular (on definite articles). They are mainly concerned that children have difficulties understanding the context of the task in an adult-like way, more precisely that they do not understand that only one of the two pictures they see (differing in the number of items depicted on them) is considered a 'correct' choice. Children might 'mean' only one item/actor on the plural picture when the hear a singular sentence or they might 'mean' all items/actors on both pictures when they hear a plural sentence. Both 'strategies' would not be adult-like and lead to failing or guessing in a comprehension task. Therefore, Dispaldro and Benelli (2012) employ a different task to test children's grammatical knowledge that singular forms refer to 'one' entity and plural forms refer to 'more than one'. The procedure is the following: a child sit in front of two plates. On one of these plates, only one object, a teddy, is placed [X], while on the other one, two identical objects are placed [XX] (these are identical to the one object on the other plate as well). Then the child hears a sentence like 'Show me how (they) dance' (*Fammi vedere come ballano*). Italian, like Spanish, allows for subject-less sentences, which were used as test sentences. Correct comprehension in the singular trials was inferred when a child made only one bear dance, either the one from the [X] plate or one from the [XX] plate. Correct comprehension in the plural trials was inferred when the child made two bears dance, either both taken from the [XX] plate or one taken from the [XX] and one from the [X] plate. Alternatively, a child could pick up all three bears. Three, four and six year olds as well as adults were tested. The mean percentages of correct responses on singular and plural trials are shown in Figure 5.3. The authors interpret their findings as evidence that Italian children understand that singular verb forms refer to 'one' when they are about four years of age, whereas they understand that plural verb forms refer to 'more than one' when they are older than six years of age.[66] The authors further discuss the ap-

[66]Dispaldro and Benelli (2012) additionally report the percentages of partici-

Age group	singular	plural
three year olds	86 (*30*)	60 (*34*)
four year olds	95* (*8*)	68 (*18*)
six year olds	95* (*10*)	92* (*12*)
Adults	100* (*0*)	97* (*7*)

Table 5.3: Mean percent (and *standard deviations*) correct comprehension for singular and plural pro drop sentences (*significantly better than chance, according to the authors), taken from Dispaldro and Benelli (2012).

parent production-comprehension asymmetry that results, under the assumption Italian-speaking children "produce singular and plural morphemes correctly from age of three (Caprin and Guasti, 2009)". The asymmetry is interpreted as further evidence that children produce grammatical morphemes in the first phase of their grammar acquisition without a deeper knowledge of the semantic properties.

5.2.4 Comprehension of 'verb inflection' in French

The studies presented above found only very late comprehension of verb inflection as a means of number marking (and present tense marking as well as generic reading) in normally developing children. Legendre and colleagues on the other hand have found rather early comprehension of agreement information as a cue to subject number (Legendre et al., 2010).[67] Since 90 % of French verbs are not phonologically marked for number (3^{rd} person singular and 3^{rd} person plural inflected verbs are near-homophones), the number marking actually appears on the pronoun, if this takes the form of a subject clitic (*il* versus *ils*). But this number distinction is only audible when

pants who correctly comprehended the 3^{rd} person verb inflections in the given age groups. To 'understand correctly', participants had to perform correctly on either 5/6 or 6/6 trials per number condition. From these comparisons they conclude that children are only able to comprehend plural inflection correctly around age six.

[67] The authors of the cited studies claim that they examine children's comprehension of verb inflection (Legendre et al., 2010; Barriere, Goyet, Nazzi, Kresh, and Legendre, 2011), but they actually seem to be testing comprehension of pronoun marking.

verbs with a vowel onset follow the pronoun. Such sentences were used by Legendre and colleagues to investigate French children's comprehension of subject-verb agreement (see 5.7 and 5.8 as examples). Importantly, they used pseudo-words as object labels in the test sentences and pseudo-objects in the visual stimuli, to neutralize any conceptual number information on the object (e. g. for a sentence *The boy kisses the doll*, a picture of one boy kissing a doll and a picture of two boys kissing a doll would be felicitious). Legendre and colleagues first tested 24- and 30-months old French-learning children using the IPLP method.

(5.7) *Il embrasse le voube.*
He kiss-3SG the voube.
He kisses the voube.

(5.8) *Ils embrassent le taque.*
They kiss-3PL the tak.
They kiss the tak.

They found that 30-months old children looked more to the matching picture in the test compared to the baseline phase in the singular and the plural number condition. For the 24-months olds, no effect was found. Because these results clearly challenge the findings by Johnson et al. (2005) and Pérez-Leroux (2005), Legendre and colleagues did a control experiment in which they employed a picture selection task. Here, the 30-months old children again pointed significantly more often to the matching picture than would be expected by chance (although it has to be noted that they pointed to the correct picture to an overall extend of 61 %, which should still be considered far from adult performance). The findings were interpreted by the authors as evidence that 30-months olds have knowledge of subject-verb agreement and that this knowledge can be applied independent of the demands of task.

In a follow-up experiment, the comprehension of French agreement morphology was extended to pseudo-verbs (Barriere et al., 2011). In the first experiment reported above, children were only tested on verbs that they knew according to parental report (Legendre et al., 2010). To investigate whether this allowed the children to show better than chance performance, another group of 30-month-olds was tested on sentences containing only pseudo-verbs. Again all pseudo-verbs began with a vowel to make the number information on the pronoun audible via 'liaison' between the subject clitic and the verb (e. g. *ils arrouvent*). Only a picture selection task using dynamic scenes was employed. The children pointed to the matching video 61 % of the time, which is significantly above chance. The authors took this as evidence that "French-speaking 30-month-olds are able to pay attention to a single verbal number agreement cue marked on a pseudoverb and to match it to the appropriate number of agents of an action" (Barriere et al., 2011, pg. 44). According to the authors, there is no difference found between the number conditions. But the difference between the singular and plural condition revealed a trend for significance (p=.07), and single comparisons to 50 % chance-level revealed that children showed better than chance performance in the plural trials, but performed at chance in the singular trials. The authors nevertheless concluded that the children "perform equally well in the singular and plural conditions" (Barriere et al., 2011, pg. 45). This conclusion should, in my view, be taken with caution, considering the statistical analyses.

Importantly, it should be noted that Legendre, Barriere and colleagues do not truly test the knowledge of verb inflection, since this morphological information is not phonologically present in the verbs used. They rather tested children's knowledge of pronouns and their number distinction, which was only audible in combination with a particular kind of verbs, namely those starting with a vowel (and through co-articulation effects, could be (mis)interpreted by the child as though the *s* in the plural form were actually a prefix of the verb). Additionally, the authors reported that more than 80 % of the children tested in their study comprehend the singular and plural pronoun according to parental report.[68] Therefore, it should not come as a surprise

[68] Additional differences apply to the stimuli. Legendre et al. (2010) used dynamic scenes in which one or two boys actually acted instead of static pic-

that the same children could make use of the pronoun number information in an experimental setting (although parental report and experimental data often differs from each other, see for example Houston-Price, Mather, and Sakkalou (2007)).

Another aspect that makes the studies of Legendre, Barriere and colleagues hard to compare to the English- and Spanish-studies on verb inflection comprehension is the procedure employed. More precisely, it is unclear whether the feedback provided to the French-speaking children had an effect on the comprehension results. In both studies, children were presented with an eye-catching video on the *matching* screen at the end of each trial. The authors did not provide any data that would prove or discard the idea of a 'learning curve' during the experiment. Considering that each child was tested on eight trials and that the critical pronoun-verb structure was presented three times during each trial, one can imagine that the children learned the connection between the form of the pronoun (*il* versus *ils*) and the matching visual scene during the experiment. Studies on artificial grammar learning and word learning in infants have revealed that even much younger children are very good at forming abstract categories and linking new word forms to specific referents and scenes (e. g. Gómez and Gerken, 1999; Gerken and Bollt, 2008; Plunkett and Schafer, 1998; Woodward, Markman, and Fitzsimmons, 1994). The French-learning 30-months olds could very well have coupled the presence of a 'liaison'-s with scenes in which two actors were performing an action. The fact that the effects found are rather small (roughly 60 % correct), raises the possibility that children performed at chance-level at the beginning of the experiment, but improved to better than chance performance by the end of the experiment. Unfortunately, it is not possible to discard this possibility based on the reported results.[69]

tures.

[69] Legendre et al. (2010) cite a study by Kouider and colleagues who used the same kind of procedure, including visual feedback, to justify their choice of introducing feedback information (Kouider et al., 2006). But Kouider and colleagues provide statistical evidence using a linear regression on trial number that no 'learning' happens over the course of testing, which is not

Thus, since the studies by Legendre, Barriere and colleagues do not investigate the use of verb inflection *affixes* in sentence comprehension, and since they provide feedback to children of which the effect is unkown, it is in my view not a compelling evidence against late comprehension of verb inflection morphology as reported by Johnson et al. (2005), Leonard et al. (2000), Pérez-Leroux (2005) and Miller and Schmitt (2009).

Overall, studies that have examined the use of verb inflection as a cue in sentence comprehension have found that children are only able to use this cue rather late in development, most notably much later than when the same morphemes appear in their production. Comparisons between English and Spanish show that the richness of the morphological paradigm does not seem to play a critical role. Some researchers have examined production and comprehension of verb inflections with the same material in the same group of children, as it seems ideal approach.[70] These studies will be presented in Chapter 6. Before this, it will be discussed whether comprehension of verb inflection seems to be subject to linguistic and non-linguistic influences, as found for verb inflection production and the production of other grammatical morphemes.

5.3 Factors influencing comprehension of verb inflection

While there is little research on the comprehension of verb inflection, the systematic exploration of factors influencing verb inflection comprehension is even sparser. Johnson (2005) suggested that variable input is a factor that causes a delay in the comprehension of grammatical morphology. Children acquiring the African American dialect of English (AAE), in which 3^{rd} person singular verbs are frequently not marked overtly in connected speech, were not found to be able to use the 3^{rd} per-

provided by Legendre, Barriere and colleagues.

[70]Leonard et al. (2000) have only very roughly scanned spontaneous speech for the presence of correctly inflected verbs, without reporting detailed data on this matter.

son singular *-s* in sentence comprehension until age 7 or even later. The present tense singular *-s* was neither a cue to subject number (Johnson, 2005) nor a cue to present tense or 'verbness' for children acquiring the AAE dialect (de Villiers and Johnson, 2007).

In contrast, Miller (2012) did not find that variable input affected comprehension of verb inflection morphology. She tested English-speaking children's comprehension and production of agreement marking on the auxiliary *do*, which can be variable in certain English dialects when produced in combination with negation. Since Miller tested one group of children in both modalities with very similar material, the results are presented in Chapter 6.1 in more detail. But when comparing Millers findings on agreement markers on the auxiliary *do* to those of Johnson et al. (2005) on agreement markers on main verbs, there seems to be no difference between the two kinds of verbs. Thus, English-speaking children at age 4 and half are at chance-level when interpreting the 3^{rd} person singular *-s*, no matter if this agreement morpheme is attached to a main verb or to an auxiliary.

Sundara et al. (2011) found that the position of the verb in the sentence had an influence in children's ability to detect the presence or absence of agreement marking. Children aged 22 and 27 months showed sensitivity to the presence or absence of verb inflection morphology only in sentence final position, not in sentence medial position. It has to be noted though that this study did not actually tap comprehension of verb inflection, but rather sensitivity to morphosyntactic well-formedness.

Thus, at this point there is only very limited and partly contradictory information on factors that might influence comprehension of verb inflection. Factors that have been shown to be relevant in production should be evaluated for their role in comprehension as well. These are questions for further research.

6 Comparing comprehension and production

Some studies have documented that children produce verb inflections correctly rather early in development. Other studies have investigated children's use of verb inflections in comprehension and found that children are much later able to infer the number of the sentence subject by relying on verb inflection information. A production-comprehension asymmetry seems to be in place. In this chapter, I first review studies that have examined production and comprehension of verb inflection in one group of children. Then I present other areas of language acquisition for which the pattern 'production-preceding-comprehension' has been found and try to find underlying similarities between those linguistic entities under discussion. Finally, I discuss how the relation between comprehension and production in children can be accounted for by different theoretical assumptions, most importantly those pertaining to the early knowledge of verb inflection.

6.1 Testing comprehension and production of verb inflection in one group of children

Testing a particular linguistic structure in the productive and receptive modality should yield a more complete picture of children's knowledge regarding the structure under investigation. Keeney and Wolfe (1972) investigated spontaneous speech, elicited imitations and comprehension using a picture selection task to describe the acquisition of agreement in English. The motivation for this thorough assessment of English children's agreement morpheme acquisition was an earlier study by Keeney and Smith (1971), in which they had found only chance-level comprehension for sentences with pseudo-nouns and regularly inflected verbs by four year old English-speaking children. Keeney and Wolfe (1972) found that the three to almost five year old children used the correct verb inflection morphemes in 94 % of the obligatory contexts found in spontaneous speech.

In sentence imitation, children were asked to repeat sentences with correct and missing subject-verb agreement. Children repeated verbatim 80 % of the grammatical sentences, but only 49 % of the ungrammatical ones. Of the ungrammatical sentences that were not imitated verbatim, 93 % were changed into grammatical sentences, thus subject-verb agreement was established. This led the authors to conclude that "the subject-verb agreement rule is part of the language competence of the children tested" (Keeney and Wolfe, 1972, pg. 702). Interestingly, and contrary to expectations, children were far more likely to change the number information on the subject to agree with the verb than vice versa. The authors had expected that children would rely more heavily on the number information that was presented on the subject noun, because noun plural was found to develop earlier (see for example Leonard et al. (2000) and Hsieh et al. (1999) on this matter) and because number is semantically meaningful on the noun. This pattern might be seen as an indicator that the rule that children employed in sentence production was less based on semantic information regarding the subject but rather on phonological properties of the nominal (namely that the noun should lack a final *-s* when the verb is marked by an *-s*).

In the picture selection task finally, children pointed to the correct picture in 54 % of the cases, thus showing chance-level comprehension. Interestingly, even the addition of an inflected subject noun to the test sentences did not improve children's responses very much, yielding only 59 % correct responses. In addition to the picture selection response as a measure of comprehension, children were asked to give a verbal response. Thus, they were presented with the verb *sings* and had to respond verbally with either *one bird* or *two birds*. Children responded correctly in 58 % of the cases. According to the authors, this exceeds chance-level performance, and is thus interpreted as showing correct comprehension of verb inflection, contrary to the results gathered in the picture selection task. Keeney and Wolfe (1972) suggested that both the verbal comprehension test and the production and imitation tests tap only 'verbal knowledge', while in the picture selection task a child has to comprehend the meaning of the verb inflection and then identify a pictorial referent. In this case, reliance of a purely verbal rule is considered to be not sufficient, giving rise to the production-preceding-comprehension pattern.

Another early study that compared production and comprehension abilities in one group of children using the same material was conducted by Fraser et al. (1963). Among other grammatical contrasts, Fraser and colleagues examined the singular-plural distinction as marked by inflection (see 6.1 and 6.2 for examples) or by the auxiliaries *is* and *are*. To avoid doubly marking the number information, only nouns without overt plural marking (e. g. *sheep* or *deer*) were used (see 6.3 and 6.4 for examples). They tested three to three and a half year old children's comprehension using a picture selection task with a pair of pictures differing in the number of animals shown, and their production by asking the children to name the same pictures.

(6.1) The boy draws.

(6.2) The boys draw.

(6.3) The deer is running.

(6.4) The deer are running.

Comprehension was found to be ahead of production in the sense that the children had higher scores on the comprehension than on the production task. But this conclusion was challenged by Johnson et al. (2005), whose close inspection of the data revealed that the children's performance in the comprehension task was only 50 % correct, which is not different from chance-level performance.[71] Note that the findings allow to infer that

[71] Johnson et al. (2005) rather ascribe the poor comprehension performance to difficulties with the subject nouns used. They report anecdotal evidence

the children tested produced less than half of the sentences in the production task correctly, thus their ability to produce verb inflections cannot be considered to be adult-like.[72]

Miller (2012) confirmed and extended the findings on production and comprehension of verb inflection in one group of English-speaking children. The main aim of the study was to investigate whether the acquisition of grammatical morphology is affected by variability in children's input. English-speaking children from working-class (WC) backgrounds and middle-class backgrounds (MC) were tested on their production and comprehension of the 3^{rd} person singular verb inflection *-s*. Critically, the study focused on the processing of the verb inflection when attached to the auxiliary *do*. The English *do* patterns syntactically like modals, as it is fronted in question formation, but, unlike to modal *can*, it agrees with the subject. *Do* is interesting because it sometimes lacks agreement when it contracts with negation (e. g. *He don't live here.*), but only in working-class English speakers.[73] Speakers of both dialects (WC and MC-English) consistently produce agreement morphemes on main verbs. Therefore, WC children receive variable input for agreement marking on *do* while MC children receive consistent input. This pattern makes agreement marking on *do* an interesting test case to observe the influence of inconsistent or variable input.

from a small pilot study in which they had used the irregular homophone plural forms, to ask children for example *Show me the picture where the deer run/runs.* Even five year olds seemed to be very confused by the plural forms, reflected in questions like 'Do you mean the one with the deers?'. Thus, the findings by Fraser et al. (1963) can not be considered to have revealed an asynchronous relationship between children's production and comprehension of grammatical morphemes.

[72] Fraser et al. (1963) report only collapsed data with the aim of determining the order of difficulty of linguistic problems for children, so that the precise rate of correctly solved subject-verb agreement sentences is not mentioned. Similarly, mistake patterns are not reported, so it is unclear what part of inflection (noun or verb inflection) was the source of difficulty for the children when asked to produce sentences like 6.1 and 6.2.

[73] In working-class English speakers and in African American English, dropping the *-s* agreement affix can be viewed as a grammatical option in colloquial speech, since speakers of these communities very often, but not always, do this (Miller, 2012).

To examine production, children were prompted to produce questions that either required overt agreement marking because the sentence subject was 3^{rd} person singular (e. g. *Does your dad write with glitter glue?*) or did not require agreement marking (e. g. *Do you write with glitter glue?).* The four year old WC children and the four year old MC children produced questions without necessary agreement marking (DO-questions) to an adult-like extend correct. Group differences were found for the questions that required an overt inflection marker (DOES-questions). MC children produced significantly more agreement marking that WC children, as was predicted if variable input affects acquisition.

Miller examined whether this pattern would hold in comprehension as well, i. e. whether WC children were delayed or less able to use *does* in comprehension than their MC peers. She further wanted to test whether three, four, and five year old children truly have difficulty in the comprehension of agreement morphology as found by Johnson et al. (2005) and whether this would hold true not only for main verbs but also for the agreeing auxiliary *do*. The material was very similiar to the one used by Johnson and colleagues, except that agreement was marked on the auxiliary instead of the main verb (e. g. *Where does the duck swim? Where do the ducks swim?*). Additionally, control sentences were used in which the number of the subject noun was not disguised (e. g. *Where do the birds fly?*). Both groups of children performed above chance in the control sentence condition, although the MC children performed better than the WC children. In the test sentence condition, both groups of children performed at chance level. This was true for both number conditions (DO and DOES). Thus, the comprehension task did not provide evidence that input variability affects acquisition. The findings are, however, consistent with Johnson et al. (2005), who predicted that agreement is difficult in comprehension, regardless of its form.

Hence, when comparing the findings from the production and the sentence comprehension studies regarding verb inflection, a puzzling picture emerges. The production studies suggest that at the latest by age 4, English children have mastered the subject-verb agreement system, since they produce verb inflection morphemes on an adult level. This indicates that they can process the number information of the noun that is relevant for selecting the correct verb form. On the other hand, children of the same age show no evidence of being able to use this same kind of information in sentence interpretation. This suggests that children's abilities to handle number information related to subject-verb agreement develop in an asynchronous fashion in the domains of production and comprehension. This paradoxical situation is even found when the same group of children is tested with the same set of materials, as shown by the early studies of Keeney and Wolfe (1972) and Fraser et al. (1963) and the very recent study by Miller (2012).[74]

6.2 Comprehension-production asymmetries in other linguistic domains

The observation that children understand linguistic structures that they are not able to produce yet, is very common and therefore not particularly surprising (Clark and Hecht, 1983).[75] Young children's receptive lexicon for example is found to be more elaborate than their productive one (e. g. Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick, and Reilly, 1993; Dapretto and Bjork, 2000; Goldin-Meadow et al., 1976). And while children only consistently use grammatical morphemes in their speech production around their third birthday (e. g. Brown, 1973), they already recognize them in the speech stream and use them for example for word segmentation and classifica-

[74]Miller (2012) even showed the asymmetry at the level of individual children.

[75]But see Smolensky (1996) for the claim that even comprehension-preceding-production asymmetries pose problems for theories that assume one underlying grammar for comprehension and production and how this dilemma could be resolved within an OT-framework.

tion during their first and second year of life (e. g. Shi and Lepage, 2008; Höhle et al., 2004). Similarly, children performed best when they heard sentences containing correct function words in the appropriate position in a sentence, even though they were not producing those fundtion words in their own speech at the time of testing (e. g. Gerken and McIntosh, 1993; Shady and Gerken, 1999). IPLP-studies have further shown correct comprehension of word order and wh-questions during the one-word-stage, again indicating that comprehension abilities are in place earlier than production abilites (e. g. Seidl et al., 2003; Gertner, Fisher, and Eisengart, 2006; Hirsh-Pasek and Golinkoff, 1996). Regarding the general discussion about the level of abstractness that should be attributed children's early linguistic knowledge is, evidence for early abstract knowledge is easier found in comprehension studies. Production studies more often point to early lexically based knowledge (e. g. Fisher, 2002; Boyd and Goldberg, 2011). Thus, comprehension tasks most commonly reveal greater linguistic knowledge than production tasks at any given age.

But recent studies have also provided evidence for the reverse pattern, according to which children produce forms that they do not yet understand correctly. One example of such a counter-intuitive pattern is English pronouns (e. g. Hendriks and Spenader, 2006; Sekerina et al., 2004). The German focus particle *auch* ('also') (e. g. Hüttner, Drenhaus, van de Vijver, and Weissenborn, 2004; Höhle et al., 2009), restrictive modifiers (e. g. Hurewitz et al., 2000), French pronouns (e. g. Legendre and Smolensky, 2012) and sentence negation (e. g. Wojtecka, Koch, Grimm, and Schulz, 2011) are other domains in which production seems to precede comprehension in the course of language acquisition. In the following, we will take a closer look at the available data for pronouns, focus particles and restrictive modifiers.

English pronouns are perhaps the best known instance of a production-preceding-comprehension pattern. While children produce pronouns (*him, her*) from the ages of two or three on, adult-like comprehension is usually not found before the age of six (Chien and Wexler, 1990; Hendriks and Spenader, 2006;

Sekerina et al., 2004). This holds for pronouns in main clauses, in which the pronoun is not coreferential with a preceding NP (e. g. *Tigger is washing him*), as well as for pronouns within prepositional phrases (e. g. *Tigger put the box behind him*). In the former case, children often (mis)interpret the personal pronoun as a reflexive one. In the latter case, the sentence is ambiguous in adult English, as the pronoun can refer to the sentence internal referent or to an external referent mentioned in the discourse. Children younger than six do not grasp the two meanings that these sentences have in adult English (Sekerina et al., 2004). The pattern of better performance in the production than in the comprehension of pronouns is even found in children older than six (de Villiers, Cahilane, and Altreuter, 2004).

A comparable cross-modal asymmetry has as well been found for the acquisition of focus particles. The interpretation of sentences containing focus particles is highly dependent on the prosodic structure of the sentence. If the word 'carrot' is stressed in a sentence like *Tigger only gave a CARROT to Piglet* the listener should infer that Piglet received a carrot from Tigger and nothing else. If on the other hand the word 'Piglet' receives primary stress (*Tigger only gave a carrot to PIGLET*), the listener should infer that Piglet received a carrot and that no one else did (Szendroi, 2004). In production experiments asking for descriptions of picture pairs that only differ in one single attribute, children below age four produce contrastive stress in an adequate way (Hornby and Haas, 1970), but children at that same age seem not to be able to make use of contrastive stress when interpreting sentences with a focus particle like 'only' (e. g. Szendroi, 2004; Gualmini, Maciukaite, and Crain, 2002). Interestingly, children perform better when discourse information in the preceding context adds to the prosodic information (Gualmini et al., 2002). Thus, prosodic information alone does not determine children's comprehension of sentences containing the focus particle 'only' at five or six, even though they use this information productively by the age of four. Similar problems for German children arise when processing sentences containing the focus particle *auch* ('also'). Again, interpretation of such sentences is highly dependent on prosodic information. If the focus particle is accented as in the sentence *Tobi hat AUCH eine Puppe* ('Tobi has ALSO a doll'), the interpretation calls for a subject-alternative set, i. e. there must be someone else besides Tobi who also has a doll. If the focus particle is not accented (*Tobi hat auch eine PUPPE*, 'Tobi has also a DOLL'), the interpretation calls for an object-alternative set to make the sentence felicitous, i. e. there has to be something else that Tobi owns. Even though children were found to produce accented and unaccented *auch* as early as the age of two or three (Nederstigt, 2003), they still did not perform in an adult-like manner in a picture selection task at the age of five to six (Hüttner et al., 2004). The use of restrictive modifiers is another area in which this production-comprehension asymmetry has been observed. Hurewitz et al. (2000) aimed to further explore children's inability to comprehend a prepositional phrase like 'on the napkin' in sentences like *Put the frog on the napkin in the box* as a restrictive modifier. As found by Trueswell et al. (1999), children aged four to five years always interpret such an ambiguous PP as a destination of the verb put, even if the referential scene is calling for a modifier interpretation. Hurewitz and colleagues presented children with arrays of objects that always contained two animals of the same kind (e. g. two frogs) as well as short stories about these animals. The four to five year old children produced restrictive modifier phrases in 87 % of all cases, when they were asked which particular animal performed a specific action mentioned in the short story. But in the act-out comprehension task that immediately followed the production task, the same children interpreted a PP as a restrictive modifier in only 22 % of all cases. Thus, while English-speaking children aged four to five years use restrictive modifiers productively to disambiguate the members of a set, they fail to realize that a PP is used to determine a specific referent during language comprehension (Hurewitz et al., 2000).

All of these cases in which production seems to be ahead of comprehension have led to intense discussions about what such cross-modal asymmetries might reveal about the child's underlying linguistic system. One possible way of dealing with them is to assume the existence of dissociations in the grammatical systems, which underly the performance in language production and comprehension. This does not necessarily imply different grammars for comprehension and production, but rather assumes that comprehension lags behind in those cases where the speaker's perspective has to be taken into account (Hendriks and Spenader, 2006). A second approach renders extra-grammatical factors, e. g. children's computational restrictions and processing limitations responsible for comprehension failures in areas where production seems to be ahead (e. g. Grodzinsky and Reinhart, 1993). A similar approach considers the asymmetry to be an artifact of methodological problems in traditional comprehension tasks (Bloom, Barss, Nicol, and Conway, 1994; Bergmann et al., 2011; Conroy, Takahashi, Lidz, and Phillips, 2009).

Most interestingly, within-modality asymmetries and task-dependent effects have to date been found for three of the four mentioned linguistic domains in which production seems to precede comprehension: pronouns, focus particles and verb inflections, but not prepositional modifiers. Studies investigating these linguistic phonemona while comparing children's comprehension abilities using different experimental methods showed adult-like patterns with some techniques and performance levels inferior to those of adults with other techniques.

For example Sekerina et al. (2004) examined the processing of pronouns in adults and children using pointing reactions, reaction time, and eye gaze as dependent variables. Adults' pointing reactions indicated that English pronouns are referentially ambiguous. In sentences such as *The boy put the box behind him*, the pronoun was interpreted as referring to the sentence-internal referent (the boy) in 80 % of all cases, while it was interpreted as referring to an external referent (a previously mentioned man) in about 20 % of all cases. Adults' eye-tracking data replicated these findings by showing that participants detect the referential ambiguity of pronouns online. The five to six year old children, on the other hand, almost always choose the sentence-internal referent when tested with the picture selection task. Thus, based on the pointing reactions, one would have to assume that English-speaking children aged five to six do not construct a sentence-external interpretation, and thus are not sensitive to the referential ambiguity of English pronouns. But children's looking behavior revealed a pattern quite similar to the one found in adults. Their eye gaze patterns showed that they were indeed taking an external referent into consideration. This was not evident in the offline pointing reactions. Recent work by Bergmann et al. (2011) as well showed that Dutch-speaking children showed appopriate comprehension of pronouns when sentence comprehension was investigated using the eye-tracking paradigm but not when children were asked to point to the referent of the pronoun. An extensive overview on studies examining pronoun comprehension comes from Conroy et al. (2009), who as well concluded that asymmetric acquisition patterns are best described as methodological artifacts. Comparable findings for sentences with focus particles come from (Höhle et al., 2009), who examined the comprehension of accented and unaccented *auch* ('also') employing the eye-tracking paradigm. As mentioned earlier, adult-like comprehension of sentences with this focus particle has not been found before the age of five or even seven years. When Höhle and colleagues tested German children's comprehension of the focus particle *auch*, they found fixation patterns reflecting correct processing of these particles in three year old children.[76]

The complexity of the linguistic material mentioned in this survey is noteworthy. It seems to be the case that late comprehension, sometimes even later than production, is usually found for linguistic structures that involve some kind of ambiguity. English pronouns are clearly ambiguous, since they can refer to sentence-external or sentence-internal referents, depending on syntactic environment and sentence prosody (Bloom et al., 1994; Sekerina et al., 2004). Depending on the sentence structure, restrictive modifiers are either only temporarily ambiguous, with their syntactic role becoming clear upon the arrival of another constituent in the sentence (e. g. *Put the frog on the napkin in the box*), or they are completely ambiguous

[76]See Berger and Höhle (2012) for additional evidence that young children can make use of focus particles in sentence comprehension when this information is highly relevant for completing the task.

and meaning is constrained by referential context and/or lexical bias (e. g. *Tickle the frog with the feather* or *Choose the cow with the stick*) (Snedeker and Trueswell, 2004; Trueswell et al., 1999; Hurewitz et al., 2000). Sentences containing focus particles like *nur* 'only' and *auch* 'also' are ambiguous as well, because the focus particles can be associated with various constituents in the sentence (e. g. Höhle et al., 2009).

Additionally, all of the linguistic entities for which production-comprehension asymmetries seem to arise and which are prone to within-modalitiy asymmetries (pronouns, focus particles and partly restrictive modifiers) *restrict the meaning of another constituent* in the sentence or even the preceding discourse AND entail the processing of various strands of linguistic and non-linguistic information. Pronouns undisputably refer to another referent mentioned in the sentence or discourse (e. g. Sekerina et al., 2004), and children need to learn during the course of language acquisition which referent this might or might not be. joo Song and Fisher (2005) have found that young children's comprehension of sentences containing pronouns is affected not only by lexical and syntactic knowledge but also by the prominence of a referent in the preceeding discourse. Therefore, various strands of information (lexical, syntactic, discourse) have to be integrated in order to interpret the pronoun correctly. The same applies to the interpretation of restrictive modifiers. The PP 'on the napkin' (in a sentence like *Put the frog on the napkin in the box*) modifies either the verb 'put', thus giving details about the putting-action, or the NP 'the frog', thus specifying which frog is mentioned. The latter reading is much more likely given two frogs in the referential scene or discourse because then it is felicious to further restrict which frog a speaker refers to. Thus, children need to integrate linguistic information on lexical bias and discourse-related information provided by the referential scene in order to comprehend such sentences in an adult-like fashion. Focus particles, finally, associate with the focused element in a sentence and their interpretation as well depends on lexical, syntactic, prosodic and discourse information (e. g. Höhle et al., 2009; Berger and Höhle, 2012).

Verb inflections might as well fit in this pattern, as their presence is dependent upon another element in the sentence (the subject) and since there is, given the existence of syncretism, a certain amount of ambiguity involved.[77] It remains to the seen whether the proposed asymmetry between verb inflections holds true for German even when tested in a less demanding paradigm (like restrictive modifiers) or whether ealier comprehension can be found using eye tracking (like pronouns and focus particles). But when discussing children's use of verb inflections in comprehension, similarities to other linguistic structures mentioned above should not be neglected.

6.3 The comprehension-production asymmetry

The question remains as to whether subject-verb agreement production in fact precedes the comprehension of number information provided by verb inflection in sentence interpretation. This relates the more general question of whether the acquisition of verb inflection is an aspect of language acquisition in which the development of comprehension truly lags behind the development of productive skills (and why this might be so) or whether the findings available to date should be viewed instead as task-dependent methodological artefacts.

Previous explanations for the pattern of production preceding comprehension in language acquisition can roughly be divided into two broad categories. One type of theories renders grammatical factors as responsible (e. g. Hendriks and Spenader, 2006) and the other type attributes the pattern completely to extra-grammatical factors, like processing limitations or task difficulties (Bergmann et al., 2011; Conroy et al., 2009). Nevertheless, theories of both sides generally assume *one* underlying grammatical system for comprehension and production, "because assuming two different grammars that develop at a different rate is highly unattractive" (Hendriks and Koster, 2010,

[77] Certain sentences, such as *The ducks swim on the pond* (see example 5.1) and *The deer is running* (see example 6.3), are also be locally ambiguous until the inflectional morpheme is encountered.

pg. 1889). It would be unclear how children finally end up with one adult grammatical system and how grammatical knowledge acquired via one modality would end up being accessible for the other modality (see also Smolensky (1996) on this matter).

The accounts that render extra-grammatical factors (like task demands or processing difficulties) as responsible for production-preceding-comprehension patterns usually conclude that children possess the relevant knowledge of grammar, but that errors children make in comprehension (and production) relate to performance factors (e. g. Conroy et al., 2009; Bloom et al., 1994). Alternatively, researchers have argued that production-comprehension asymmetries can be "generated on the basis of grammar only" (Hendriks and Koster, 2010, pg. 1892). Hendriks and colleagues have done so in an Optimality Theory (OT) framework (OT). They argue that 'bi-directional optimization' is present in adult grammars, but over-burdens the computational processing capacities in children. Such bi-directional optimization is necessary because the input in production is meaning and the input in comprehension is form (the constraints in OT that are applied during the process of optimization are output oriented) and thus the direction of optimization is reversed in production and comprehension (e. g. Hendriks and Spenader, 2006; Hendriks and Koster, 2010; Smolensky, 1996). In the OT-account of production-comprehension asymmetries in children, linguistic knowledge is not applied in an asymmetric way, such that children can use a rule in production but not in comprehension, but it is assumed that linguistic knowledge (in the form of constraints) is applied in a symmetric way, therefore in comprehension and production alike. The critical factor to explain production-preceding-comprehension asymmetries are thought to be the result of direction-sensitivity of the constraints (and thus the grammar).[78]

[78] Legendre and Smolensky (2012) on the other hand argue that uni-directional optimization is sufficient and that constraint re-ranking can explain patterns in which production precedes comprehension.

This distinction between (1) that children possess relevant (adult-like) linguistic knowledge that cannot be applied due to performance factors and (2) that children's grammatical knowledge is (in some way) not sufficient to allow for adult-like production and comprehension performance relates to the acquisition of verb inflection and the two broad approaches (nativist vs. constructivist approach) presented in Chapter 2.2. Remind that proponents of the strong continuity approach credit children with adult-like knowledge of subject-verb agreement and verb inflection as soon as they produce inflected verbs in their own speech (e. g. Hyams, 1999; Wexler, 1998). Such an adult-like rule could be formulated as was in 2.5 (repeated below for convenience as 6.5). This rule, specifying person and number features of the subject and the verb, would constitute children's competence and should (performance matters aside) be equally applicable in language production and comprehension. Alternatively, one can assume that children possess a non-adult like rule that rather relies on form properties of the input (formulated in 2.6, repeated below as 6.6) and does not relate to semantic meaning. Such a rule could would still constitute children's grammatical competence and give rise to many correct verb inflections in production but no adult-like comprehension, irrespective of performance factors.[79]

(6.5) $V_{INF} \rightarrow V_{3SG}$ iff SubjNP_{3SG}

(6.6) V+/Ø/ $\rightarrow$ V+/s/ following he, she, it, NP+/Ø/

The logical consequences of a production-comprehension asymmetry for the two approaches to verb inflection acquisition are quite different. Such an asymmetric pattern is puzzling

[79] Assuming that child grammar could, at some point in development, be comprised of rules such as presented in 6.6 is certainly very different from the OT-approach put forward by Hendriks and colleagues, since they assume that a child's grammar is made up of the same constraints as the adult grammar, but ranked differently and possibly without the knowledge of bidirectional optimization (e. g. Hendriks and Spenader, 2006; Hendriks and Koster, 2010). The present suggestion regarding children's early knowledge of verb inflection does not credit children with adult-like grammatical knowledge. But both accounts assume that the reasons for across-modality asymmetries lie within the child's grammar.

under the assumption that children have one underlying grammatical system for production and comprehension in which they have stored an adult-like rule regarding verb inflection. As stated, production and comprehension data that exist so far indicates a production-comprehension asymmetry with regard to children's use of verb inflections. Even more surprising is the direction of the asymmetry with production preceding comprehension. Usually, comprehension is seen as the 'easier' task, which should precede production in acquisition based on logical and empirical grounds (e. g. Clark and Hecht, 1983). A theory that assumes that adult-like knowledge of verb inflection is available in children (in form of a rule like 6.5) needs to explain how comprehension performance can be more severely 'impaired' than production performance (if the production-comprehension asymmetry really holds true and can not be ascribed to methodological factors). If one on the other hand assumes that children's early verb inflection productions are based on non adult-like rules that relate to the form of verb and nominal (as in 6.6), rather than to their semantic content, a production preceding comprehension pattern could easily be explained. Children's productions could be based on a non adult-like rule, giving rise to superficially good production performance. But such a rule would not be helpful in sentence comprehension. To do this, children either need an adult-like rule that relates to meaning or they need an additional rule, which links specific noun form to semantic (number) meaning. Comprehension studies could provide an answer on when children begin to form an adult-like rule regarding verb inflection (or form additional rules relating noun form to meaning).

Without already taking a stand on the theories mentioned above, it has to be noted that the comprehension studies on verb inflection cited so far all employed some variant of a picture selection task, which requires explicit decisions to be made by the children tested.[80] This places demands on non-linguistic

[80] Except the study by Legendre et al. (2010) and the experiments reported in dissertation by Melanie Soderstrom (Soderstrom, 2002), which used the IPLP technique as well. But because the former studies rather the comprehension of pronouns, it is not considered as counter evidence to the finding

cognitive skills and one may hypothesize that children's actual language comprehension abilities are obscured by such non-linguistic factors (Höhle et al., 2009). In the framework of the competition model, one could argue that the verb inflection information as a cue to subject number is of low validity and high cost (e. g. Bates et al., 1991; Dick et al., 2004), and therefore children might not be able to use it when tested in a paradigm that poses higher demands on them (e. g. Leech et al., 2007).

When searching for a method for assessing children's early interpretation of grammatical morphemes, which puts only low task demands on children and does not require an explicit choice among a set of pictures, IPLP suggests itself. As explained in Chapter 5.1, this method has been widely used to assess young children's lexical and syntactic comprehension abilities (e. g. Hirsh-Pasek and Golinkoff, 1996; Naigles, 2002; Golinkoff et al., 1987) as well as their ability to detect phonological and syntactic well-formedness (e. g. Kedar et al., 2006). Recent research has provided evidence that earlier comprehension of difficult linguistic structures can be assessed when IPLP measures are employed (e. g. Sekerina et al., 2004; Bergmann et al., 2011; Höhle et al., 2009). Whether earlier comprehension of verb inflection can be found when it is assessed using the preferential looking paradigm has theoretical implications regarding the kind of grammatical knowledge that children should be credited with at a certain point in development.

that verb comprehension is found to be late when tested with a picture selection task. The latter study did not find any evidence for the early use of verb inflection in sentence comprehension in English-speaking toddlers, thus indirectly supporting the notion of 'late comprehension'.

7 Summary and Research Questions

The main aim of the current work is to investigate German children's knowledge of verb inflection morphology in the receptive modality. Such receptive knowledge can be divided into (1) knowing that a certain morphological form has to be present in a particular syntactic context, and (2) knowing how this morphological form relates to a semantic meaning (of another constituent) in sentence context. The former relates to subject-verb agreement, i. e. the linguistic phenomenon that forces the verb to carry the same person and number features as the sentence subject, (evident in the inflectional forms of subject and verb). The latter relates to the connection between the verb form and the meaning of the agreeing sentence subject, i. e. whether some 'semantic information' can be inferred from an inflectional morpheme.

As outlined in Chapter 3, young children's spontaneous speech is often interpreted as indicating early grammatical knowledge of the agreement relation subject and verb enter into. This assumption partly relies on the observation that children rarely make agreement mistakes. Under this view, held by proponents of the nativist approach to language acquisition (e. g. Hyams, 1999; Wexler, 1998), children are attributed with early adult-like knowledge of verb inflection morphology and subject-verb agreement. Children's use of non adult-like forms (e. g. *he go*) are explained by performance factors, such as processing demands (e. g. Phillips, 1995). This account has been challenged by proponents of a constructivist approach to language acquisition (e. g. Pine et al., 1998; Wilson, 2003). This alternative account claims that children's very early verb inflections in spontaneous speech reflect knowledge of memorized chunks of speech and/or non adult-like rules that children have abstracted from the stored constructions. Children's use of non adult-like verb forms (e. g. *he go*) is explained by properties of the input and children's 'wrong' recalling of chunks (e. g. 'Can *he go* to the store?') (e. g. Theakston et al., 2003). This view calls on findings indicating that children's early verb inflection productions are limited to particular lexical items, on strong frequency and familiarity effects that can be observed and on the fact children have difficulties using verb inflections productively with pseudo verbs (e. g. Pizutto and Caselli, 1992; Rubino and Pine, 1998; Theakston et al., 2003). Based on production data alone, it is difficult to decide which of the theoretical assumptions is more appropriate for describing children's early morphological knowledge.[81]

To help decide this matter we turn to receptive studies. These have either focused on children's sensitivity to the presence of agreement morphology (see Chapter 4.3) or the use of verb inflections in sentence comprehension, testing whether children understand the connection of the verb form and the meaning of the sentence subject (or tense and aspect of a sentence, see Chapter 5.2 or 6.1). Previous studies have found early sensitivity to the presence of verb inflection in English-, Dutch- and French-learning children below age two.[82] But critically, the results of such HPP studies do not indicate that young children base their preferences for grammatical structures on adult-like grammatical rules (e. g. Soderstrom, 2002).[83] Infants rather seem to be attuned to the distributional regularities between verb inflection and properties (position and form) of the subject noun phrase (e. g. Soderstrom et al., 2007; Soderstrom, 2008; Nazzi et al., 2011).

Previous studies that examined whether children can make use of verb inflection to infer the semantic number of the subject (whether the subject refers to one actor or more than one actors) have found surprisingly late comprehension. Children below the age of five of six do not seem to be able to use verb inflection as a cue to subject number (e. g. Johnson et al., 2005;

[81] Frequency and familiarity effects as well as limited productivity in the use of verb inflections in elicited production tasks is ascribed to performance factors in a nativist framework.

[82] But such sensitivity has not been found for German-learning infants yet.

[83] Although early (adult-like) grammatical knowledge is claimed by Polisenska (2010).

Pérez-Leroux, 2005). If this findings is set in comparison with children's use of verb inflections in the productive modality, a production-comprehension asymmetry becomes apparent. If one takes a nativist view, claiming that children possess an adult-like grammatical rule to subject verb agreement early on, any 'failure' to use verb morphology appropriately has to be ascribed to performance factors. This approach is faced with the challenge of explaining how performance in comprehension can be 'more impaired' than performance in production. Usually, comprehension of a given structure is considered to be easier than its production and over the course of language acquisition, comprehension is said to precede production, a claim based on empirical and theoretical grounds (e. g. Clark and Hecht, 1983; Bates, 1993). If one on the other hand takes a constructivist view, across-modality asymmetries can be explained with the kind of non adult-like knowledge children base their processing of verb inflection on. Children might possess a grammatical rule that allows them to produce seemingly correct verb forms in most linguistic contexts (see 6.6), and allows them further to distinguish grammatical structures from ungrammatical one based on distributional regularities, but that same rule might not relate to meaning and therefore does not help children to infer semantic implications of verb inflections. To decide on the questions of whether a true production-comprehension asymmetry is in place and what kind of theory is more adequate to explain children's early knowledge of subject-verb agreement, it is helpful to examine whether young children's failure to show adult-like comprehension, as found for example by Johnson et al. (2005) and Pérez-Leroux (2005), whether earlier comprehension can be found when employing a different methodology, thus if comprehension performance is related to task demands.

This thesis aims to shed further light on children's receptive abilities regarding verb inflection. I will try to answer the following questions:

1. Are German-speaking children aged three to four years able to understand the connection between the *form* of a verb inflection and the *meaning* of a sentence subject in a preferential looking task?

2. Is the ability to make use of verb inflections during sentence comprehension (to infer the number of a sentence subject) influenced by methodological factors, i. e. by task demands?

3. Are German-speaking children aged three to six years able to detect the dependency between the *form* of the subject and the *form* an inflected verb has to take, i. e. are they sensitive to subject-verb agreement violations?

Experiment 1 (Chapter 8) aimed to determine whether German-speaking children aged three, four and five years showed sensitivity to subject-verb agreement violations when processing simple SVO-sentences. It was measured whether children showed 'disruption' in finding a visual referent when the target noun was presented in an ungrammatical (with subject-verb agreement violation) as opposed to a grammatical sentence (with correct subject verb agreement). A further goal of Experiment 1 was to test whether an IPLP setting could be used to assess children's ability to detect grammatical morphemes (other than determiners, as shown by Kedar et al. (2006), Van Heugten and Johnson (2011) and Zangl and Fernald (2007)).

The other experiments investigated German children's use of verb inflection morphology in sentence comprehension. Experiment 2 (Chapter 9) was designed to test whether German-speaking children were able to use the verb inflection information in sentence processing as a cue to subject number in an a pure IPLP design. Experiment 3 (Chapter 10) extended the research question to another experimental methodology, namely picture selection. This was done to determine whether the earlier comprehension found in Experiment 2 as compared to other studies reported in the literature should be attributed to task-demands or to cross-linguistic differences. Finally, experiment 4 (Chapter 11) was conducted to validate the findings from the picture selection task in Experiment 3 using a slightly different experimental procedure. The comprehension experiments were administered to shed further light on the proposed production-comprehension asymmetry in the acquisition of verb inflection and the rules that might underlie children's processing of verb inflections.

Part II

Empirical Investigations

8 Experiment 1: Sensitivity to SV-agreement

In this first experiment I investigate whether German-speaking children are sensitive to subject-verb agreement violations in auditorily presented sentences. Children aged three-, four-, and five-years as well as adults were tested in a preferential looking paradigm. It was investigated whether children's ability to find a target picture was affected by the grammaticality of the sentence in which the target noun was presented.

8.1 Rationale and research question

The present experiment aimed to examine German-speaking children's ability to detect the agreement relation between subject and verb and violations thereof. The choice of an appropriate method proved to be critical. Adult's processing of subject-verb agreement is often tested with on-line measures that are not suitable to use with pre-schoolers, as they require reading abilities (e. g. Pearlmutter, Garnsey, and Bock, 1999; Nicol et al., 1997) or the ability to perform metalinguistic grammaticality judgements (McDonald, 2008b; McDaniel and Cairns, 1996). Very young children on the other hand are tested with non-referential methods assessing preferential listening times (e. g. Soderstrom et al., 2007; Nazzi et al., 2011; Van Heugten and Johnson, 2010), but this method is as well unsuitable for the age range under investigation here (e. g. Nelson et al., 1995).

An alternative measure that has proven useful in assessing young children's syntactic competence is *preferential looking* or *eye tracking* (Hirsh-Pasek and Golinkoff, 1996; Trueswell, 2008). For the most part, this method is used to assess children's language comprehension, but studies by Kedar et al. (2006), Zangl and Fernald (2007) and Van Heugten and Johnson (2011) have shown that children's eye gaze can be used to measure their sensitivity to the grammaticality of a sentence as well. Therefore, the method of tracking children's eye gaze when being presented with grammatical and ungrammatical sentences was adopted to measure whether they detect subject-verb agreement violations.

The following preferential looking experiment was conducted to examine if German-speaking children aged 3, 4 and 5 years show sensitivity to agreement violations in auditorily presented sentences. Participants were presented with sentences in which the subject and the inflected verb were either agreeing (grammatical condition) or not (ungrammatical condition). The critical agreement feature in this study was number, so verbs were either inflected for 3^{rd} person singular (*-t*) or 3^{rd} person plural (*-en*). Furthermore, to examine whether the number of the sentence subject plays a role in the detection of SV agreement violations, sentence subjects were presented in either singular or plural.

If German children are sensitive to SV agreement violations, we expected them to show evidence for a 'disruption' in their sentence processing. Such a disruption could be reflected in children's looking behavior, as indicated by other studies using the eye tracking paradigm to measure grammatical knowledge in children (Kedar et al., 2006; Zangl and Fernald, 2007; Van Heugten and Johnson, 2011). Therefore, we measured children's looks to two object pictures, one being the referent of the object noun mentioned in the test sentence, serving as the target picture (e. g. a ball), and the other one representing an object not mentioned at all, serving as distractor (e. g. a shoe). The object noun in the test sentence was presented directly after the inflected verb, which was the locus of ungrammaticality in the agreement violation condition. Additionally, looks to a third picture, depicting the referent of the sentence subject with the appropriate number information, were measured directly after the critical agreement information had been presented.

So, if German children are sensitive to SV agreement violations, we expected a different looking behavior in grammatical compared to ungrammatical trials. Such differences could manifest themselves in various ways, such as in the proportion of looks to target picture, the latency to fixate on the target picture or the summed duration of the fixations on the target picture. Since we additionally presented children with a picture showing the sentence subject with the appropriate number information, we inspected the percentage of looks to the subject picture in all conditions as well.

Finally, this experiment was supposed to examine whether eye tracking is a viable method to examine sensitivity to subtle ungrammaticalities in an age group whose grammatical knowledge is not easy to test (three to six years).

8.2 Method

8.2.1 Participants

Fourty-seven children of three age groups (three year olds, four year olds and five year olds) as well as 27 adults participated in this experiment. All children were mono-lingual speakers of German recruited from the Berlin/Potsdam-area and had no known language deficits and were not born prematurely.

In the group of three year olds, 15 children, seven of whom were girls, were tested. The age range for this group reached from 3;1 years to 3;4 years of age. In the group of four year olds, 16 children were tested, nine of them girls. The age ranged from 3;9 to 3;11 years. Thus, this group should more precisely be called the 'late three year olds', but for ease and claritiy, I will refer to them as the four year olds throughout this work. Finally, in the group of five year olds, again 16 children were tested, ranging from 4;11 years to 5:4 years of age. In this group, six girls participated. Additionally, 27 students from the University of Potsdam were tested and served as an adult control group. In this group, 25 women and two men in the age range between 19 and 34 years participated. The children's parents gave informed constent. They were further reimbursed on their travel costs to the lab and children received a small toy after completion of the experiment. The adult participants, all students, received either class credits or a monetary reimbursement.

8.2.2 Material

Verbal material To create the simple SVO-sentences used in this experiment, eight animate nouns, eight transitive verbs as well as 16 inanimate nouns were selected. The animate nouns served as subjects in the test sentences and were chosen to be known to young children and to be easily depicted in the visual display. The parental inventory checklist in German ELFRA (Grimm and Doil, 2000) was used as a guideline to select the nouns. All of them were produced by at least 75 % of the children who served as the standardizing population of the ELFRA at the time of the 2nd birthday. Therefore we concluded that it is very likely that all the nouns are understood by German children aged three years and above. Additional restrictions were that a.) all nouns serving as sentence subjects needed to be of either masculine or neuter grammatical gender and b.) needed to have distinct phonological forms in the singular and plural form. This was done to ensure that the number distinction on the noun was as easy to detect as possible.[84] Thus, the following seven animal characters plus one human character were chosen (presented here in singular and plural form): *der Hund, die Hunde* ('the dog(s)'), *das Schwein, die Schweine* ('the pig(s)'), *der Bär, die Bären* ('the bear(s)'), *der Hase, die Hasen* ('the rabbit(s)'), *das Pferd, die Pferde* ('the horse(s)'), *das Schaf, die Schafe* ('the sheep') and *der Vogel, die Vögel* ('the bird(s)'), *das Baby, die Babys* ('the baby(s)'). Four of these nouns are of masculine and four of neuter gender.

Eight transitive verbs were chosen and each one was combined with one of the subject nouns. The verbs were selected according to the following criteria: (1) to be bisyllabic when inflected for 3^{rd} person singular as well as 3^{rd} person plural, (2) to be as frequent as possible, (3) to be known to young children, and (4) to be equally likely to occur with several objects (thus, highly frequent verb object combinations like *drinking*

[84] The singular feminine definite determiner (*die*) and the plural definite determiner for all genders (*die*) are identical in the nominative and accusative cases in German, therefore feminine nouns were avoided as sentence subjects.

coffee or *reading the newspaper* were avoided). Since the ELFRA only provides a very small set of verbs, this checklist could not serve as an indicator of children's knowledge of the verbs. The following verbs were used: *angeln* ('to fish'), *basteln* ('to craft'), *futtern* ('to eat'), *knabbern* ('to nibble'), *liefern* ('to deliver'), *öffnen* ('to open'), *schütteln* ('to shake (sth.)'), *zeichnen* ('to draw'). Importantly, all verbs used are inflected regularly in the present tense, with the inflectional morpheme *-t* suffixed to the verb stem in 3^{rd} person singular and the inflectional morpheme *-(e)n* suffixed to the stem to mark 3^{rd} person plural.

Finally, eight pairs of inanimate nouns, thus 16 nouns, were selected to serve as object nouns in the test sentences. Each pair was combined with one of the subject-verb-combinations to form two SVO-sentences. The criteria for selecting the objects were the following: (1) the two objects of a pair had to be of similar size and similarly visual complexity, (2) the labels for the two objects had to have the same grammatical gender (but all grammatical genders, i. e. masculine, feminine and neuter were possible), (3) the ELFRA-score of the two object labels had to be as similar as possible, (4) the object nouns had to be equally likely to appear in combination with the verb used in the specific sentence (i. e. a dog might be similarly likely to bring a ball or a shoe, but not a ball and a truck), (5) the number of syllables in the object nouns had to be as similar as possible, and (6) the semantic category of all the object nouns was supposed to be the same or at least similar (i. e. both objects should for example be food, clothing or toys). The matching of the object nouns and their visual referents was considered critical, because the looks to the referents was the essential measure of the eye-tracking experiment and any bias to one of the object pictures that were presented for a sentences was tried to be avoided.[85] The ultimately selected pairs of objects were the following: *Hose* ('pants') and *Hut* ('hat'), *Mond* ('moon') and *Stern* ('star'),

[85] Although it has to be admitted that it was not possible to fulfill all criteria in every single pair (i. e. the German words for *sausage* and *banana* differ in two syllables and *door* and *can* cannot be considered to belong to the same semantic category). As many criteria were fulfilled as possible, while still using only words that are known to German children aged three and above.

Banane ('banana') and *Wurst* ('sausage'), *Käse* ('cheese') and *Kuchen* ('cake'), *Ball* ('ball') and *Schuh* ('shoe'), *Tür* ('door') and *Dose* ('can'), *Bürste* ('brush') and *Flasche* ('bottle'), *Tisch* ('table') and *Stuhl* ('chair').

Using all these words, 16 different test sentences were created. Additionally, practice sentences were created from two other subjects, one singular and one plural, (singular: *der Affe* ('the monkey'), plural: *die Kinder* ('the children'), two other verbs (*wollen* ('to want'), *suchen* ('to search ')) and four objects (*Apfel* ('apple'), *Keks* ('cookie'), *Buch* ('book'), *Messer* ('knife')). Note that the words used in the practice sentences did not follow all the criteria mentioned above. Practice trials were mainly included to familiarize the participants with the procedure of an experimental trial and were not included in the analysis, therefore strict criteria for word selection seemed to be negligible here. A list of all sentences is provided in the Appendix (see Appendix A.2).

All verbal stimuli were recorded by a female native speaker of German. The speaker read all grammatical sentence versions, thus singular and plural sentences (32 sentences), in a slow, friendly and child-directed manner at least three times. Each subject noun was additionally read in isolation, combined with the number word *ein* ('one') in the singular form (e. g. *ein Hund* 'a dog') or with the number word *zwei* ('two') in the plural form (e. g. *zwei Hunde* 'two dogs'). Moreover, verbal attention getters like *Hey! Schau mal!* ('Hey! Look at that!') were recorded. After recording, all sentences were spliced into subject (definite determiner plus subject noun), verb (the inflected verb) and object (definite determiner plus object noun) phrases and the best recording for each phrase was selected by the experimenter. Since only grammatical sentences were read and recorded, ungrammatical ones had to be created via cross-splicing. To ensure that cutting and splicing could not account for any possibly found grammaticality effect, all sentences, i. e. grammatical and ungrammatical ones, were created through cross-splicing (see Van Heugten and Johnson (2011)). The mean length of grammatical sentences was 2983 ms (2550-3719) and the mean length of ungrammatical sentences was 2974 ms (2622-3363).

Visual material For each subject-verb combination, a visual array of three pictures was created to be presented on the eye-tracking monitor (see Figure 8.1). Note that the same visual array was used for both sentences of an object pair. In such an array, the picture depicting the sentence subject was presented at the horizontal center at the top of the screen. The pictures that depicted the objects were placed on the lower part of the screen, one on the right and one on the left side of the screen.

All pictures were simple colored drawings taken from the internet and found to be adequate for children. The two pictures of an object pair were comparable enough in size and color intensity so as to be equally visually interesting. All subject and object drawings were presented within a 300 pixel- times 300 pixel-square on a light colored background. Besides those squares, the screen was black.

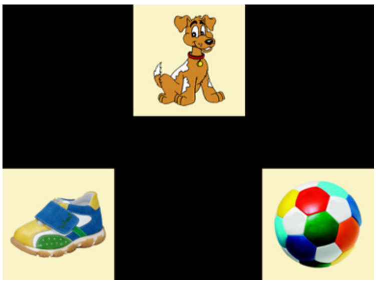


Figure 8.1: Example of visual stimuli array as was presented on the eye-tracking monitor.

Importantly, the subject picture changed according to the number of the sentence subject. Thus, if a singular sentence was presented, the picture of only one animate being (e. g. animal or baby) was presented in the top center square, while two animate beings were presented in the same square when a plural sentence was presented. This adjustment of the visual information was included to help the children even further in recognizing the number of the sentence subject. The side on which the referents of the objects nouns were presented in the array was counterbalanced across subjects. Finally, the visual arrays and the verbal sentences were combined as avi-files.

8.2.3 Design

Each of the 16 test sentences was used in four different conditions following the design of this experiment. A sentence was either presented as a singular or a plural sentence, this depending on the use of a singular or a plural subject noun. Additionally, each sentence was either presented in a grammatical or ungrammatical version, this depending on the present or absent number agreement of the subject noun and the verb inflection. See examples of these four conditions in Table 8.1.

	Subject number	Agreement	Example
1.	singular	agree	*Der Hund liefert einen Ball.* 'The dog delivers a ball '
2.	singular	non-agree	**Der Hund liefern einen Ball.* 'The dog deliver a ball '
3.	plural	agree	*Die Hunde liefern einen Ball.* 'The dogs deliver a ball '
4.	plural	non-agree	**Die Hunde liefert einen Ball.* 'The dogs delivers a ball '

Table 8.1: Experimental conditions (Experiment 1).

Eight versions of the experiment were created to counterbalance the side of the target picture, the side of the object picture (object A was on the left or the right), target object per number of sentence subject (i. e. in Version 1 object A was the target object when a singular subject was presented and object B served as target when a plural subject was presented, in Version 2 this was reversed) and grammaticality of the sentence (i. e. the agreement of the inflected verb with the number of the sentence subject).

In the end, each version consisted of 16 test trials, whereby each sentence subject was presented twice, once in the singular and once in the plural version. Furthermore, each subject was once presented with an agreeing verb (grammatical version) and once with a non-agreeing verb (ungrammatical version). For each sentence subject, object A served once as target and once as distractor. Finally, each sentence subject appeared once in the first 8 trials of the experiment and another time in the second 8 trials of the experiment (in case children's attention diminished so quickly that only the first half of the test trials should be considered for analysis).

Summing up, every participant encountered eight grammatical and eight ungrammatical sentences, eight sentences with singular subjects and eight sentences with plural subjects and eight trials were the target picture was on the right side as well as eight trials where the target object was positioned on the left side of the screen. In every age group, two participants were assigned to one of the eight experimental versions.[86]

Thus, every participant was presented with 16 test trials as well as four practice trials. The practice trials were the same for all participants, without such an elaborate counterbalancing. The subject *monkey* was only presented as a singular subject, once with an agreeing and once with a non-agreeing verb, thus in a grammatical and an ungrammatical sentence. The subject *children* was only presented as the plural subject, again once with an agreeing verb form (i. e. as a grammatical sentence) and once with a non-agreeing verb form, namely a verb inflected for 3^{rd} person singular (i. e. as an ungrammatical sentence). Side of object was not counterbalanced in the practice trials as their primary purpose was just to familiarize the participants with the trial procedure.

[86]All participants heard the same 16 sentences (composed of the 8 subject nouns, 8 verbs and 16 object nouns, but in different conditions and with target and 1-actor picture side counterbalanced across subjects.)

8.2.4 Procedure

Every experimental trial had the same following structure: At the very beginning, the three pictures appeared simultaneously on the screen without any initial verbal input. All three pictures remained on the screen and were thus visible during the whole duration of a trial. At 1000ms post trial onset, a verbal attention getter plus the name of the sentence subject with the appropriate number word was presented auditorily (*Schau! Ein Schwein!* 'Hey look! A pig!' or *Wow! Zwei Schweine!* 'Wow! Two pigs!'). At roughly 3000 ms post trial onset, the test sentence was presented (*Der Hund liefert einen Ball.* 'The dog delivers a ball'). Critically, every test sentence was presented such that the *onset of the object determiner* occured at exactly 5000ms post trial onset. This aligned the offset of the critical verb inflection information in every trials. Thus, the onset of the test sentences varied slightly between the test trials, as the duration of the subject-verb combinations, i. e. the first part of the test sentences, varied slightly. After the presentation of the test sentences, the pictures were visible for approximately four further seconds (again slightly varying in length because of the different duration of the object names), so that every trial lasted exactly 10 seconds. See Figure 8.2 for a schematic description of the course of an experimental trial. The next trial started after a 5 sec. pause.

All participants were tested using a remote Tobii® T120 eye tracking system in a dual computer set-up. Eye gaze position was tracked with a resolution of 60 Hertz on a 17-inch monitor. Trial presentation and data collection was controlled using the software Tobii-Studio®.[87]

During the test sessions, all participants were sitting in a reclined chair with their heads roughly 60 cm from the eye tracking monitor and camera. Three year old children were allowed to sit on their parent's lap if they felt more comfortable that way.

[87] The principle of the corneal reflection tracking technique employed by the Tobii eye tracker is that an infrared lightsource is directed at the eye and the reflection of the light on the cornea relative to the center of the pupil is measured. This is used to estimate where on the screen the gaze is fixated.

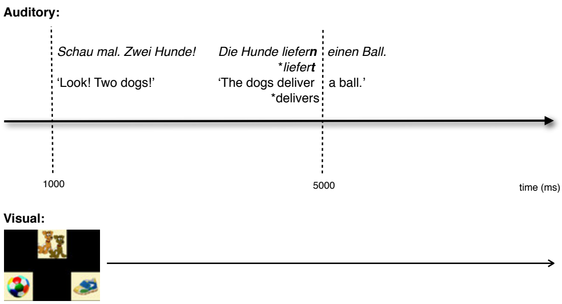


Figure 8.2: Schematic course of a trial, plural subject condition.

Parents were asked to look away from the screen and not to interact with their children. For the older children, parents were usually sitting behind them on an extra chair, while the experimenter sat in an adjacent room, not visible to the child, controlling the experiment (which amounted to starting the calibration phase and starting the experimental trials when the participants were attentive; all other visual and auditory presentation was controlled by the Tobii-Studio software). All participants were instructed to watch the pictures on the screen and to listen to the auditory stimuli. No further instruction was given and no reaction was asked of the participants.

When the participants were sitting comfortably and seemed attentive, a five-point calibration procedure was performed. Then the two practice trials and the 16 experimental trials were presented. After every fourth trial, a short clip with a moving comic character (e. g. Snoopy) was presented to redirect and keep the child's attention to the screen.

8.2.5 Data Analysis

Eye movements were measured and analyzed automatically using the Tobii-Software® employing the standard settings. The output data used for further analysis was one text file per participant providing information about the time course of the experiment, the accurate position of the eye gaze (an X-Y-coordinate representing the eye tracker screen) at each given time point as well as the number and duration of fixations according to the Tobii standard settings. To examine whether children and adults were sensitive to the agreement of subject and verb on a group level, the eye gaze data was aggregated and analyzed in a number of ways, always using the R software (The R Foundation for Statistical Computing, 2009).

To find out when and to what extent the participants looked at one of the presented pictures in relation to the verbal material, temporal and spatial areas of interest were defined. The light colored squares which contained the subject and object pictures on the screen of the eye tracker were defined as spatial areas of interest (AoI), each one 300 x 300 pixels in size. When the gaze of a participant fell onto one of these squares, the look counted as a gaze to the subject, target object or distractor object (depending on the object noun mentioned in the test sentence), respectively. It was further coded whether a look onto the right or left object on the screen was directed at the target or the distractor object, depending on the object mentioned in the specific test trial.

Additionally, temporal regions of interest, dubbed 'phases' were defined, based on the verbal information the participants heard during each trial. The first phase of interest (Phase 1) in this experiment started at the onset of the subject description, thus 1000 ms post trial onset, and lasted exactly 3000 ms. During this time frame, participants were expected to direct their eye gaze towards the picture showing the sentence subject as it is verbally mentioned with an accompanying attention getting phrase.

The second temporal region of interest (Phase 2) started at the onset of the object determiner, at 5000 ms post onset. This was directly after a participant had encountered the critical information which made a sentence grammatical or ungrammatical, namely the verb inflection (or lack thereof). Therefore, Phase 2 can be viewed as the actual 'critical temporal phase'of the experiment. Phase 2 again lasted for exactly 3000ms. It was expected that participants directed their gaze mostly to one of the object pictures, ideally the target picture, during this time frame.

All of the following analyses focus on the question of whether children displayed a different looking behavior in grammatical compared to ungrammatical trials, thereby exhibiting sensitivity to the agreement of subject number and verb inflection. The number of the sentence subject (singular versus plural) was included as a further variable to detect possible differences according to whether the participants processed singular or plural subjects within the test sentences.

During Phase 2, eye gaze to the target picture (i. e. the object mentioned in the sentence) was used as main dependent variable, as, based on previous experiments, child and adult participants were expected to show differences in their ability and speed to detect the target picture in relation to the grammaticality of the test sentences. Further, the eye gaze to the subject picture was used as secondary dependent variable, as we hypothesized that participants might display a different amount of 'control looks' to the picture depicting the number of agents in relation to the grammaticality of the test sentences or the verbal inflection they encountered.

To control for the attentiveness of the participants (on a group level) to the visual stimuli and the auditorily presented test sentences, the looks to the subject picture during Phase 1 were analyzed. Further, the proportion of looks to the target target picture during Phase 2 was analyzed (based only on looks to target and distractor) and compared to chance level. Here, higher proportion of looks to target than expected by chance was thought to serve as an indicator that participants actually processed the test sentences, or at least the object names.

8.3 Results

8.3.1 Dependent measures

To get a first impression of the participant's looking behavior, the eye gaze at every measured time stamp was calculated. The eye gaze (percentage of looks) to the target picture data was first averaged over 50 ms-bins in every trial, then averaged over all trials of a condition (four conditions: singular subject + agreeing verb, singular subject and non-agreeing verb, plural subject and agreeing verb, plural subject and non-agreeing verb) and then participants of a group. Thus, a 'supertrial' was created, which could be plottet to depict the **time course** of the looking behavior. The same was done for the looks that landed on the target picture and looks that landed on the subject picture separately. The percentage of looks to each of the pictures for Phase 2, the time frame from 5000 to 8000 ms post onset, is depicted in Figure 8.3(a) and in Figure 8.3(b). Visual inspection of these graphs allows to determine when possible effects of the test sentences grammaticality occur. But to be entered into statistical analysis (ANOVAs) to validate whether the participants' amount of looks to target and/or subject picture was actually affected by the properties of the test sentences, the eye gaze data had to be further aggregated.

Therefore, **proportions of looks** to the target picture and to the subject picture per condition and group were calculated. This constitutes a more relativized measure, whereby the time that participants spent looking at one picture (usually the target picture) is put in relation to the time they spent looking at all three pictures presented on the screen (looks to target (ms) / (looks to target (ms) + looks to distractor (ms) + looks to subject picture (ms)). Such a measure is used in many preferential looking experiments (e. g. Van Heugten and Johnson, 2011; Mani and Plunkett, 2010b). Usually, the looks to target only need to be divided by the looks to target and distractor, but since there are three pictures visible on the screen in this experiment, the formula needed to be slightly adjusted, to take into account all the looks that land on the subject picture during the testing phase. The proportion of looks was first calculated over the 3000 ms time window that constituted Phase 2 of the experiment, separately for each condition and group, once for the target picture (PLT), and once for the subject picture (PLS). Because the time course graphs indicated that especially the five year olds showed an effect of grammaticality in a smaller time window which occured in the first part of Phase 2, the proportion of looks was additionally calculated for this part of the trial (5300 to 6500 ms post onset). The time that needs to be included to account for eye movement initiation in children varies from study to study, but stays between roughly 100 ms and 500 ms (e. g. Swingley and Aslin, 2000; Swingley et al., 1999; Mani and Plunkett, 2010a; Van Heugten and Shi, 2009), which reflects the fact that it is still not known exactly how long it takes children to initiate an eye movement towards a picture after hearing auditory information. Estimated mean saccade latencies in 12 months olds are around 300 ms (Canfield, Smith, Brezsnyak, and Snow (1997), cited in Swingley et al. (1999)), which is the value used in the present study. Van Heugten and Johnson (2011) found higher proportions of looks to the target picture in trials containing a grammatical determiner compared to trials containing a nonce determiner. Thus in this experiment, we would expect children and adults to show a higher proportion of looks to the target in sentences with correct subject-verb agreement than in those that contained an agreement violation.

An alternative way to analyze eye tracking data is to take only such looks into account that are part of a fixation, see for example Poltrock (2011) and Bergmann et al. (2011). A fixation is defined to have a specific minimal duration during which a person's gaze has to stay in a certain spatial region of the eye tracking screen. In this analysis, a fixation has to have a minimal duration of 100 ms on a maximal radius of 30 pixels. Using a **fixation based approach** cancels out all looks which are supposed to be too short to actually provide the viewer with sufficient visual information. To measure the actual amount of time participants spent looking at the target picture in the different grammaticality and subject number conditions, the durations of fixations were first summed for each participant and trial and then averaged for each condition and group. This provides a looking time measure similar to that used by Golinkoff et al. (1987), Naigles and Gelman (1995) and Soderstrom (2002). Summing up the duration of all fixations per condition gives a measure of how long the participant fixated on each picture during this critical time frame and is thus preferable over an averaged fixation duration.[88] This way of analysis was additionally applied to the fixations that landed on the subject picture during Phase 2. As in the PLS-analysis, this is done to check whether participants possibly look more to the subject picture after encountering a mismatch between subject number and verb inflection.

The first fixation that occurs after critical information has been presented is considered especially revealing of the processing of that information. For example, Kedar et al. (2006) compared the number of fixations that were directed to the target picture in the grammatical and ungrammatical conditions

[88] Since it is possible that there is more than one fixation within a trial or the critical phase of a trial, it is preferable to add up the durations instead of averaging them. In principle, since a fixation has a minimal duration of 100 ms and the test phase in this experiment has a duration of 3000 ms, maximally 29 fixations could have occurred within a given test phase.

as well as the latency of these first fixations. Poltrock (2011) as well used children's first fixation after a critical event within the trials as dependent measure.[89] But the **latency of first fixation** was used as dependent measure, providing an equivalent to a reaction time measure (see 8.3.7). For this latency analysis, only the *first* fixation to the target and distractor picture *in the test phase* (Phase 2) was taken into account. These latencies were averaged over condition and group and entered into statistical analysis. With this analysis we wanted to explore whether the time to find the target picture was influenced by the grammaticality of the test sentences. A similar latency measure was employed for example by Mani and Plunkett (2010a), Fernald, Pinto, Swingley, Weinberg, and McRoberts (1998) and Lew-Williams and Fernald (2007).[90] The latency of the first fixation towards the subject picture was not analyzed as the expectations for this measure were unclear and not theoretically motivated. Regarding the fixations onto the target picture, longer latencies in the ungrammatical condition compared to the grammatical condition were expected, if participants were distracted by the sentences ungrammatility. Previous research has shown that children are slower at finding a target picture following ungrammatical structures compared to grammatical ones (e. g. Zangl and Fernald, 2007). Similarly, other methodologies that are only suitable to use with adults, like self-paced reading tasks, have revealed slower processing when encountering missing SV agreement compared to correct SV agreement (e. g. Pearlmutter et al., 1999; Freedman and Forster, 1985).

[89] The number of first fixations on the target picture per condition were calculated for the present experiment as well, but they are not reported since they did not reveal any effects of grammaticality or subject number.

[90] Although Fernald and colleagues only take into account trials in which the participants were looking at the distractor picture while the critical information was presented and then switched to the target picture. The latency of this switch is then used as a reaction time measure (e. g. Zangl and Fernald, 2007).

8.3.2 Looking behavior in Phase 1

To examine whether participants had paid attention to the number information provided by the subject picture, the mean percentage of looks to this picture was calculated for each condition and participant during the first phase (1000 ms to 4000 ms post onset). During this phase, participants were expected to direct their eye gaze towards the subject picture, with a possible effect of the number of the sentence subject, but without an effect of grammaticality, as no information on the grammaticality of the sentence had been presented yet. During Phase 2 on the other hand, participants were expected to direct their gaze primarily towards the target picture, with possible effects of grammaticality and subject number. Because of different expectations, looks during Phase 1 were calculated and are reported separately.[91] The mean percentage of looks per condition and group for Phase 1 is listed in Table 8.2.

		three year olds (N=15)	four year olds (N=16)	five year olds (N=16)	Adults (N=27)
singular	grammatical	0.506	0.400	0.404	0.739
	ungrammatical	0.450	0.409	0.385	0.679
plural	grammatical	0.576	0.648	0.497	0.773
	ungrammatical	0.599	0.516	0.509	0.751
mean		0.533	0.495	0.449	0.736

Table 8.2: Mean percentage of looks to subject picture during Phase 1.

The mean values in Table 8.2 show that adults looked to the subject picture for roughly three quarters of all time stamps that landed on screen in Phase 1. The children of all age groups looked to the subject picture for roughly half of the time.[92]

[91] When calculating the average percentage of looks to either picture, all eye gaze data that were measured as being *on screen* were taken into account. That means that the percentage values of looks to all three pictures do not necessarily add up to 100 % as participants might have spent some time looking towards the black areas of the screen, although this percentage of time should be rather small. Periods if time within this experiment or within a trial in which eye gaze was not recorded, did not enter analysis.

[92] Note that one four year old had to be discarded from analysis, because he did not provide data in all conditions.

The percentage of looks per group and condition were entered into a mixed-effect three-way analysis of variance (ANOVA) with Group (three year olds, four year olds, five year olds, adults) as between-subject factor and subject number (singular, plural) and grammaticality (grammatical, ungrammatical) as within-subject factors. This revealed a highly significant main effect for Group ($F(3,69)=14.052$, $p<0.001$) and for Subject number ($F(1,69)=45.309$, $p<0.000$). The interaction between Group and Subject number almost reached significance as well ($F(3,69)=2.712$, $p=0.052$).[93] To disentangle the almost significant interaction between Group and Subject number, separate 2*2 ANOVAs were carried out for every age group.

For the three year olds, the five year olds and the adults, the main effect for Subject number was significant without any further interactions (three year olds: $F(1,14)=6.177$, $p=0.026$; five year olds: $F(1,15)=16.325$, $p=0.001$; adults: $F(1,26)=7.581$, $p=0.016$). Therefore, participants in those groups looked more to the subject picture in plural trials than in singular trials. The four year olds showed a slightly different looking behavior. The main effect for Subject number reached significance ($F(1,14)=21.165$, $p<0.001$), but so did the main effect for Grammaticality ($F(1,14)=4.695$, $p=0.048$). Additionally, the Grammaticality X Subject number interaction approached significance ($F(1,14)=4.055$, $p=0.064$). To further investigate this interaction, a subsequent one-way ANOVA for the four year olds was performed, showing that Grammaticality did not have an effect in the singular trials ($F(1,14)<0.1$), but in the plural trials ($F(1,14)=9.365$, $p=0.008$). It is unclear why the four year olds showed an effect of grammaticality that early into the sentence, but this pattern needs to be kept in mind when analyzing the four year olds' looking behavior in Phase 2.

[93] Throughout this thesis, only significant statistical results will be reported in the text as well as those that warrant further statistical investigations. All other results, those that approach significance or have an F-value higher than 1.0, but fail to approach a significance level of 0.05, are reported in the Appendix (see A). This is done to improve readability and clarity of the result sections.

Further statistical analysis showed that the adult participants looked more to the subject picture during Phase 1 than the children of all three age groups (three year olds vs. adults: $F(1,40)=13{,}284$, $p<0.001$; four year olds vs. adults: $F(1,40)=20.079$, $p<0.001$; five year olds vs. adults: $F(1,41)=30.962$, $p<0.001$). No difference was found between the groups of children (all $p>0.1$).

The tendency of all groups to look more to the subject picture in plural trials might be due to the higher amount of visual information being presented in the pictures showing two characters, or to a higher kind of processing load associated with plural forms presented verbally. In general, the tendency of all children to look to the subject picture to a higher extent in plural trials should be kept in mind when analyzing the looking behavior in the second part of each trial, Phase 2, which is presented in the following. It can be concluded that three and five year olds as well as adults showed no difference in their looks to the subject picture in relation to the test sentences' grammaticality. This indicates that there is no acoustical information in the sentence up to the end of Phase 1 that makes it possible to differentiate between grammatical and ungrammatical sentences (which would explain the grammaticality effect found in the four year olds). Finally, it can be concluded that the participants did look to the subject picture, at least to around half of the measured time stamps. All of the following analyses will deal with the looking behavior in Phase 2, ranging from 5000 ms to 8000 ms post trial onset.

8.3.3 Detection of the target picture

To simply control whether participants had paid attention to the visual and verbal stimuli in Phase 2, the proportion of looks to target picture were calculated. Critically, only looks that landed either on the target or on the distractor picture were considered for this analysis, therefore, PLT-value was calculated by simply dividing the time participants spent looking towards the target picture by the time they spent looking towards target or distractor object picture. All looks that were directed towards the subject picture or off-screen were discarded from analysis. This measure should simply provide an estimate of whether participants were preferring the target over the distractor picture. It was expected that participants of all age groups would look more to the target picture after the object noun had been presented. The mean PLT-values per group are listed in Table 8.3. These PLT values were statistically compared to chance level 0.5 via t-tests. The t-values are listed in Table 8.3 as well.

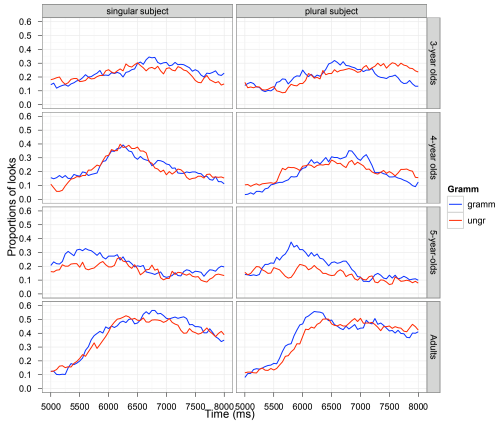
Group	mean PLT (*Std.dev*)	t-test vs. chance
three year olds	0.699 (*0.207*)	3.399**
four year olds	0.645 (*0.191*)	5.29***
five year olds	0.378 (*0.169*)	7.347***
Adults	0.875 (*0.133*)	23.091***

Table 8.3: Mean proportion of of looks to target (PLT) during Phase 2 (signif. codes: *** <0.001, ** <0.01, * <0.05, (*) <0.1).

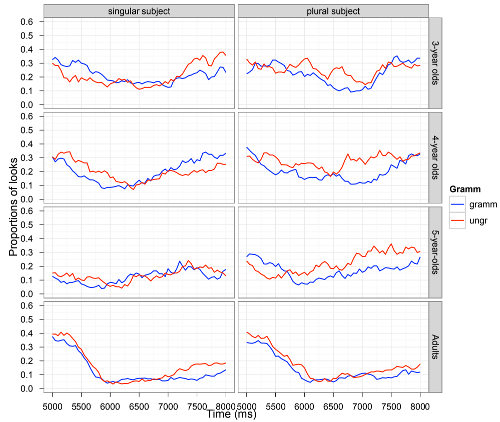
The looking data was collapsed over grammaticality and subject number condition. The results show that participants of all four groups looked more to the target picture than was expected by chance. This is interpreted as evidence that participants have paid attention to the experimental stimuli and comprehended at least the object noun properly.

8.3.4 Time course

The first analysis was intended to provide a rather rough, descriptive estimate of the looking behavior over the time course of a trial. For sake of clarity, only the looking behavior to the target picture and the subject picture in Phase 2 (from 5000 ms until 8000 ms post onset), is depicted. Looking behavior is shown separately for each subject number and grammaticality condition. The percentage of looks were averaged for each condition to create a 'supertrial' (see Figure 8.3(a) for looks to target picture and Figure 8.3(b) for looks to the subject picture).



(a) Looks to **target** picture



(b) Looks to **subject** picture

Figure 8.3: Time course of the percentage of looks to the target (left panel) and the subject picture (right panel) per number condition and group for Phase 2 (5000 to 8000 ms post onset).

As the graphs show, the looks to the target picture increased after the object word had been presntend in the sentence (i. e. at around 5000ms), but only in the group of adults, not so much in any of the groups of children. Statistical analysis have revealed however that the three year olds, the five year olds and the adults looked more to the target than to the distractor picture after the object noun had been mentioned (see Chapter 8.3.3). From this, at least lexical comprehension of the object noun had been deduced. Further, an early advantage in the grammatical compared to the ungrammatical trials is visible in the five year olds, but not so much in the other groups. Statistical comparisons have to be pursued to determine whether the rate of looks to the target picture were actually higher in the grammatical compared to the ungrammatical trials.

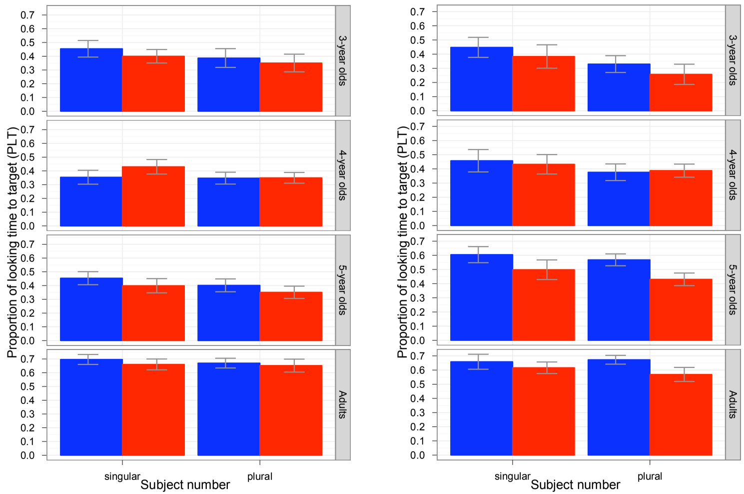
The graphs that represent the looks to the subject picture show a rather steep decline after ca. 5000 ms post onset, but only in the group of adults. Thus, adults' amount of looks to the subject picture decreased sharply after the target word, i. e. the object noun, had been mentioned. A difference between the grammaticality conditions might be found in the groups of four year olds and five year olds, especially so in the trials containing a plural subject. Again, whether such a difference in the percentage of looks holds statistically, has to be determined.

8.3.5 Proportion of looks analysis

To analyze the looking behavior in Phase 2, the proportion of looking time to the target picture and subject picture was calculated. The mean PLT-value (target picture) for each group and condition is depicted in Figure 8.4(a), while the mean PLS-value (subject picture) for each group and condition is depicted in Figure 8.5(a).[94] The statistical analyses of these PLT-values

[94] In this analysis, trials in which participants did not look at any picture during the critical Phase 2 could not be included. Because division by 0 is not possible, no PLT-value could be calculated for those trials. Therefore, it was necessary to exclude the data of two participants (one four year old child and one adult) completely from analysis because they did not deliver data in all four conditions.

are presented below, separately for the adult controls and the three groups of children.



(a) PLT 3000 ms

(b) PLT 1200 ms

Figure 8.4: Mean proportion of looks to the **target picture** for the 3000 ms lasting Phase 2 (left) and the 1200 ms lasting shorter time window (right). Blue bars show grammatical trials and red bars show ungrammatical trials (Errorbars: $+/-$ 1 SE).

Proportion of looks to target picture (3000 ms)

Adults Since the adults were tested as control group to investigate whether the methodological approach used yielded any interpretable results, their looking behavior is going to be analyzed and presented separately.[95] The 2x2 ANOVA comparing Subject number and Grammaticality did not yield any significant results. Adults looks to the target picture were not influenced by the grammaticality of the test sentence or by the number of the sentence subject.

Children The mean PLT-values per Group and condition were entered into a 2x2x3 ANOVA comparing Grammaticality, Subject number and Group. Only the main effect for Subject number reached significance, confirming higher PLTs for the singular than the plural trials ($F(1,45)=6.655$, $p=0.013$). No further effects were found. Nevertheless, because we were mainly interested in whether children of any group would display effects of grammaticality or subject number, further planned post-hoc ANOVAs were performed for each age group. No effects were found in the group of three year olds and in the five year olds. In the group of four year olds, the interaction between Subject number and Grammaticality reached significance ($F(1,14)=5.92$, $p=0.029$). Post-hoc tests comparing Grammaticality in the single number conditions separately did not yield any significant results. Therefore, when using proportion of looks as dependent variable, no sensitivity towards the grammaticality of the test sentences could be found in any group of participants.

Visual inspection of the participants' looking behavior (see Figure 8.3(a)) indicated that an effect of grammaticality on the looks to the target might only be found in a smaller time frame than the one constituted by Phase 2, which lasted for 3000 ms. Therefore, PLT- and PLS-values were again calculated, this time

[95] The same approach is taken in Experiment 2 and 4. Another advantage of this procedure is that it will keep Experiment 2, 3 and 4 easier to compare, because it does not add another level of comparison to the ANOVAs (adult controls were tested in Experiment 2 and 4, but in Experiment 3. Comparing the children to the adults in one statistical analysis could obscure the findings, especially considering that this comparison is not possible in all experiments).

only taking looking behavior into account that occurred between 5300 and 6500 ms post onset.[96] This smaller time window begins 300 ms later than the bigger one does, to account for eye movement initiation in children.[97] The duration of the time window was restricted to 1200 ms, a rather arbitrary value that was motivated by visual inspection of the time course graphs. Eye tracking studies with children vary in the duration of the 'test phase' for which looking data is analyzed (e. g. 1000 ms and 1600 ms in Poltrock (2011), 2000 ms in Van Heugten and Johnson (2011), 2500 ms in Mani and Plunkett (2010b)). Therefore, a 1200 ms long time window duration was defined.

Proportion of looks to target picture (1200 ms)

Adults The 2x2 ANOVA comparing effects of Grammaticality and Subject number on adults PLT-values for a smaller time window did not reveal any significant effects. As for the bigger time window, no influence of the test sentences' grammaticality is found on adult's looks to the target picture.

Children The '1200-time window' PLT-values of the children were as well entered into a 2x2x3 ANOVA comparing Grammaticality, Subject number and Group. In this analysis, a main effect of Group ($F(2,40)=3.331$, $p=0.045$), a main effect of Grammaticality ($F(1,40)=4.475$, $p=0.041$) and a main effect for Subject number ($F(1,40)=8.689$, $p=0.005$) are found. No interac-

[96] In this analysis, trials in which participants did not look at any picture during the critical time window starting at 5300 post onset and lasting 1200 ms (up to 6500 ms post onset) could not be included. Because division by 0 is not possible, no PLT-value could be calculated for those trials. Therefore, it was necessary to exclude the data of six participants (two four year old children, two five year old children and two adults) completely from analysis because they did not deliver data in all four conditions.

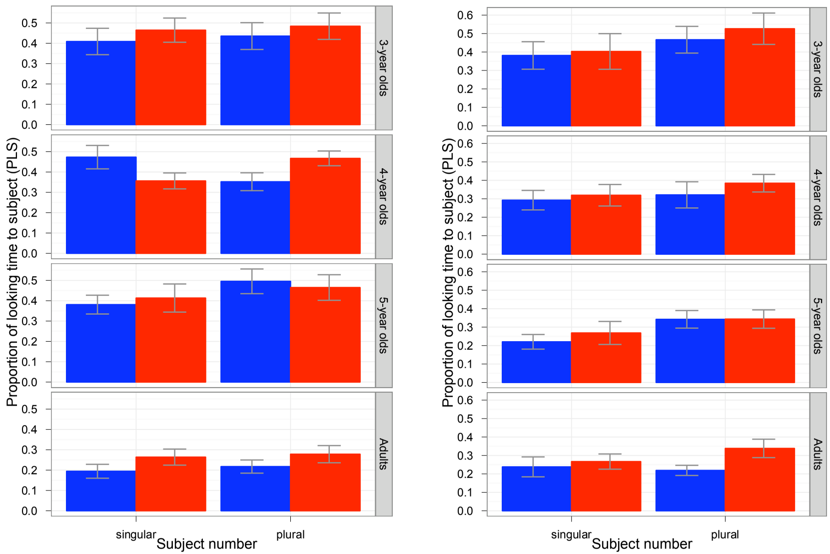
[97] Eye gaze studies with children very often do not analyze eye movements that occurred within the first 100 ms to 500 ms after the onset or offset of the linguistic stimulus under investigation (e. g. Swingley and Aslin, 2000). Van Heugten and Johnson (2011) discard eye gaze that occurred within the first 320 ms after target onset, and Zangl and Fernald (2007) discard the eye gaze that occurred within the first 367 ms after target onset.

tion is found to be significant. Still, to investigate the looking behavior of the single child groups in further detail, post-hoc 2x2 ANOVAs were performed. The data of the five year olds revealed a main effect for Grammaticality ($F(1,13)=5.386$, $p=0.037$), without any interactions. Therefore, the five year olds were looked more towards the target picture in the grammatical compared to the ungrammatical trials, but only when the PLT-values were calculated over a smaller time window, lasting 1200 ms instead of 3000 ms.[98]

Proportion of looks to subject picture (3000 ms)

Adults The mean PLS-values, the mean proportion of looks towards the subject picture, were as well entered into a 2x2 ANOVA, comparing Grammaticality and Subject number. The adults showed an influence of the test sentences' grammaticality on their PLS-values ($F(1,25)=8.323$, $p=0.008$) without any further interaction. They looked more to the subject picture in the ungrammatical compared to the grammatical trials, without any influence of subject number.

[98] The data of the three year olds revealed a main effect for Subject number ($F(1,14)=6.726$, $p=0.021$), since the youngest group looked more to the target picture in the singular compared to the plural trials.



(a) PLS 3000 ms

(b) PLS 1200 ms

Figure 8.5: Mean proportion of looks to the *subject picture* for the 3000 ms lasting Phase 2 (left) and the 1200 ms lasting shorter time window (right). Blue bars show grammatical trials and red bars show ungrammatical trials (Errorbars: $+/-$ 1 SE).

Children When the PLS-values of the children were entered into a 2x2x3 ANOVA, only the interaction between Group, Grammaticality and Subject number reached significance ($F(2,43)=3.454$, $p=0.041$).

To disentangle this three-way interaction, post-hoc 2x2 ANOVAs for the different age groups were performed and yielded the following results. For the three year olds, no effects were found to be significant (all F's <1.0). For the four year olds, the interaction between Grammaticality and Subject number was significant($F(1,14)=9.022$, $p=0.009$). The five year old children displayed a slight preference to look at the subject picture in the plural trials compared to the singular trials ($F(1,15)=3.576$, $p=0.078$), without any interaction. To further investigate the interaction found in the group of four year olds, the number conditions were analyzed separately. In the singular trials, the effect of grammaticality approached significance ($F(1,14)=3.919$, $p=0.068$), while in the plural trials, grammaticality did have a clear effect on the PLS-values ($F(1,14)=5.345$, $p=0.036$). Interestingly, the four year olds looked more to the subject picture in the grammatical trials when a singular subject was presented and in the ungrammatical trials, when a plural subject was presented. Put differently, the four year olds looked more towards the subject picture whenever they encountered a *-t*-inflected verb compared to an *-n*-inflected verb (although in the singular subject condition, this effect only approaches significance).

Proportions of looks to subject picture (1200 ms)

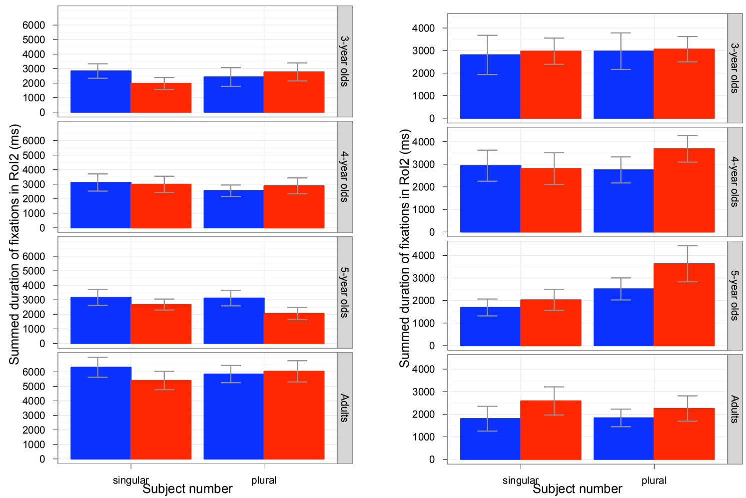
Adults The mean PLS-values calculated from the smaller 1200 time window were again entered into a separate 2x2 ANVOA. The data of adults did not reveal any significant effect or interaction. Thus, in this smaller, and most importantly earlier time window, no effect of grammaticality was found on the adult's looks to the subject picture. Therefore one can conclude that the effect found in Phase 2 (reported above) rests on looking behavior that occurs later than 6500 ms post onset.

Children The PLS-values for the children were as well entered into statistical analysis. The 2x2x3 ANOVA revealed a very marginal trend for Group (F(2,40)=2.53, p=0.092), but a significant main effect for Subject number (F(1,40)=11.39, p=0.002). No further effects or interactions were found. Still, the data of the groups of children were analyzed separately. This revealed main effects for Subject number in the three year olds (F(1,14)=6.503, p=0.022) and in the five year olds (F(1,13)=7.164, p=0.019), but no effects in the group of four year olds. Thus, the preference for the subject picture in the trials containing a -*t*-inflection, that was observed when all data from Phase 2 were use to calculate proportion values, must stem from looking behavior occurring later than 6500 ms post onset. Overall, no effect of grammaticality is found in the smaller, earlier time window, but three year olds and five year olds direct more looks to the subject picture in the plural compared to the singular trials.

8.3.6 Fixation duration analysis

The fixation duration on the target picture are shown in Figure 8.6(a) and those on the subject picture can be seen in Figure 8.6(b). The values were entered into four separate ANOVAs, one for adults and one for children, each for the target picture data and for the subject picture data.[99]

[99] To be able to perform statistical tests on fixation data, only participants who fixated the target picture at least once in all four conditions could be included. For the analysis of the target picture, this criterion applied to 52 participants. In the group of three-year olds, 10 participants remain. In the group of four-year olds, 12 remained. In the group of five-year olds, only nine participants remained. In the group of adults, 21 participants were included in the analysis. To analyze the summed durations of fixations onto the subject picture in RoI2, the following number of participants had to be excluded in each group: four in the three year olds, leaving 11; four in the four year olds, leaving 12; seven in the five year olds, leaving nine participants; and finally 15 in the group of adults, leaving 12 participants in the analysis. Altogether, only the data of 44 participants could be included in this analysis.



(a) Fixation duration on target picture (b) Fixation duration on subject picture

Figure 8.6: Summed fixation durations on the target picture (left) and the subject picture (right) during Phase 2. Blue bars show grammatical trials and red bars show ungrammatical trials (Errorbars: $+/-$ 1 SE).

Fixation duration on target picture (3000 ms)

Adults The 2x2 ANOVA comparing Grammaticality and Subject number for the adults did not reveal any significant effects. The adults therefore display very similar fixation durations in all conditions, as all *F*'s were below 1.0.

Children The fixation duration values were entered into a 2*2*3 ANOVA comparing Grammaticality, Subject number and Group for statistical analysis. The ANOVA did not revealed any significant effects. Still, to explore the results of the single age groups in greater detail, separate post-hoc ANOVAs were carried out, revealing the following. For the three year olds, the Grammaticality X Subject number interaction approached significance ($F(1,9)=3.826$, $p=0.082$). In the four year olds, no significant main effect or interaction was found (all F's <1.0). For the five year olds, as well no effect of grammaticality or subject number on the duration of fixations was found.

Nevertheless, the data were investigated further, to investigate whether grammaticality had an effect of the looking behavior in the separate number conditions. For the three year olds, an effect of grammaticality was found in the singular trials ($F(1,9)=8.114$, $p=0.019$), but not in the plural trials ($F<0.1$). For the four year olds, no such influence of grammaticality was found in either number condition (both F's<0.1). For the five year old children, no influence of grammaticality was found in the singular trials, but a trend emerged in the plural trials ($F(1,8)=3.79$, $p=0.087$). Figure 8.6(a) shows that these children do look longer towards the target picture in the grammatical compared to the ungrammatical trials, thereby reflecting the pattern found for the percentage of looks in this group, but in this fixation-based analysis only a subset of the data and children could be included.

Fixation durations on subject picture (3000 ms)

Adults The adults fixated the subject picture significantly longer in the ungrammatical than the grammatical trials ($F(1,11)=6.072$, $p=0.031$). This effect was true in both subject number conditions, as no interaction with subject number was found. This result is in line with the findings from the proportion of looks analysis, showing that adults indeed look more to the subject picture when number information on the subject and the verb mismatch.

Children Children's summed durations of the fixations on the subject picture were entered into a 2x2x3 ANOVA comparing Grammaticality, Subject number and Group. The main effect for Subject number reached significance ($F(1,29)=4.748$, $p=0.038$). Additionally, an almost significant interaction between Grammaticality and Subject number emerged ($F(1,29)=3.813$, $p=0.061$). The trend for an interaction warranted further statistical analyses, separately for every age group. In the groups of three year olds and four year olds no significant effects were found.

In the group of five year olds on the other hand, the summed durations of the fixations did differ in relation to the sentences' grammaticality and subject number. The children in this group showed a reliable main effect for Subject number ($F(1,8)=19.273$, $p=0.002$) and an approaching main effect for Grammaticality ($F(1,8)=4.256$, $p=0.073$). Further, the interaction between Grammaticality and Subject number showed a strong trend for significance ($F(1,8)=4.941$, $p=0.057$). Therefore, fixation duration were analyzed separately for each number condition for the five year olds. In the plural trials, a significant effect of grammaticality was found ($F(1,8)=7.782$, $p=0.024$). In the singular condition, no effect appeared ($F<1.0$). Thus, the five year olds looked more towards the subject picture in the ungrammatical trials, but only so in the plural number condition.

8.3.7 First fixation latency analysis

In the following analysis, the latency of first fixations towards the target picture will be examined. Only fixations that actually landed on the target picture were considered for analysis. This restriction was already applied before the first fixation of each trial was singled out. Thus, if a participant *first* fixated the distractor picture but then switched to fixate the target picture, only that second fixation - the first one *on* the target picture - was used for analysis. This might have an effect on the latency of the first fixation, because one or even more whole fixations on the 'wrong' pictures might have occurred before. Importantly, only fixations that occurred after 5300 ms post onset of each trial were considered for analysis (see PLT-analysis for the small time window). This left 300 ms for the participants to launch their eye gaze after the critical SV agreement information had been presented (e. g. Van Heugten and Johnson, 2011). The mean latencies of the first fixations that landed on the target picture are depicted in Figure 8.7.[100]

Adults For the adults, a 2x2 ANOVA was carried out to determine whether grammaticality or subject number had an effect on the time that they needed to look towards the target picture. It turned out that the adults were not influenced in their speed to recognize the target picture.

Children A 2x2x3 ANOVA was carried out to statistically analyze the mean latency of first fixation in the three groups of children. Results revealed no effects. Nevertheless, to explore the looking behavior of the single groups in more detail, post-hoc 2x2 ANOVAs for the single age groups were carried out. For the three year olds, no effects were found (all F's <1.0). The same was found for the four year old children (all F's<1.0).

The five year old children on the other hand were influenced by the grammaticality of the test sentence, since they

[100] Only those participants that delivered data in all conditions could be included in the analysis. This left 52 participants, more precisely the data of 10 three year olds, 12 four year olds, 9 five year olds and 21 adults.

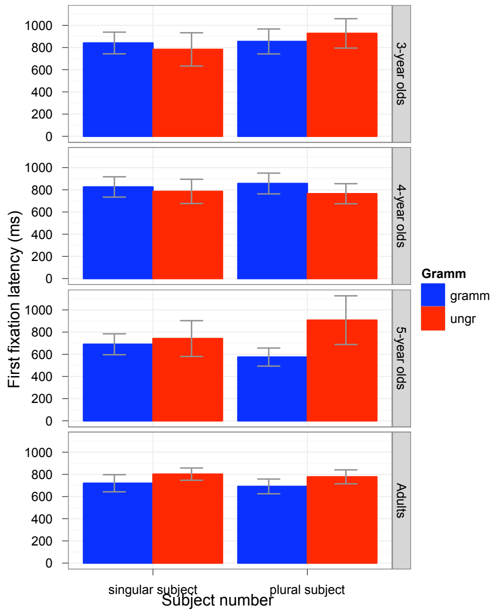


Figure 8.7: Mean latency of the first fixation per trial on the target picture. Only the data of 52 participants is represented in this plot. Blue bars show grammatical trials and red bars show ungrammatical trials (Error-bars: $+/-$ 1 SE).

were slower to direct a fixation onto the target picture in the ungrammatical compared to the grammatical trials ($F(1,8)=6.358$, $p=0.036$). No influence of Subject number was found to be significant. This result pattern, showing that five year olds were slower to find the target picture in the ungrammatical condition compared to the grammatical condition, is displayed in Figure 8.8. All other groups did not differ in their reaction time to look towards the target picture in relation to the grammaticality of the test sentence.

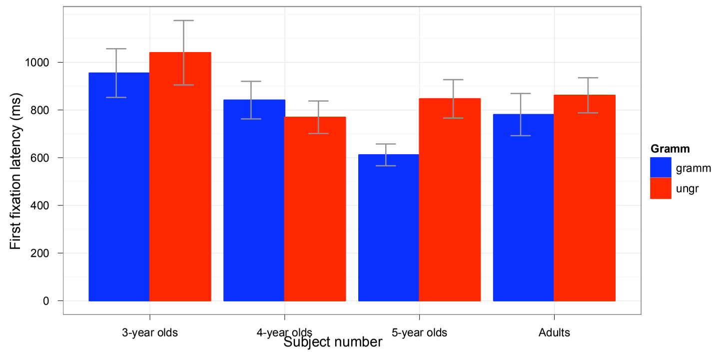


Figure 8.8: Mean latency of the first fixation onto the target picture, collapsed across subject number, for each grammaticality condition and group. Blue bars show grammatical trials and red bars show ungrammatical trials. (Errorbars: $+/-$ 1 SE)

8.3.8 Summary of Results

Experiment 1 aimed to determine when German-speaking children would be able to detect subject-verb agreement violations employing a preferential looking task. Participants were presented with SVO-sentences that contained a subject noun that was clearly marked for number (visually and verbally) and a verb that either agreed or disagreed with the number information on the subject. Children's eye gaze towards a target picture representing the object noun following the verb was measured. This was done to examine whether children were faster and more accurate in looking at a target picture when the object noun was part of a grammatical sentence with correct subject-verb agreement, compared to an ungrammatical sentence, containing an agreement violation. Additionally, the looks towards the subject picture, which reflected the number information of the sentence subject, were measured.

The adults' were tested as a control group. When only the subject noun with an appropriate numeral was presented (*Look, two dogs!*), roughly three-quarters of adults' looks on screen was directed towards the subject picture. It was concluded that adults paid attention to the number information provided by the subject picture. After the test sentence was presented, in Phase 2, adults looked significantly more to the target than the distractor picture, indicating that they had processed and comprehended the object noun. The amount of looks towards the target picture was not influenced by the test sentences' grammaticality. Thus, adult's were not found to be faster or more accurate at looking towards the target picture when the object noun was presented in a grammatical compared to an ungrammatical sentence. But, the amount of looks towards the subject picture was influenced by the grammaticality of the test sentences. Adults looked more and longer towards the subject picture in the ungrammatical compared to the grammatical trials. This pattern suggests that adults launched more 'control looks' to the subject picture when presented with sentences that contained a subject-verb agreement violation. This suggests that adults detected the agreement violation between subject number and verb inflection.

The three year old children looked to the subject picture during Phase 1 for about half of the measured time stamps. In Phase 2 they preferred the target picture over the distractor picture, just as the adults did. Therefore, they seemed to have paid attention to the verbal and visual stimuli as well. The proportion of looks to target did not reveal any influence of grammaticality. Only the fixation durations during the singular trials indicated that three year olds looked more towards the target in the grammatical compared to the ungrammatical trials. This can be interpreted as an indicator that German three year olds detected that a singular subject does not agree with an *-en*-inflected verb. But the latency to fixate on the target picture was not influenced by grammaticality, as the proportions of looks. Sensitivity to subject-verb agreement in three year olds therefore seems to be not very stable. Subject number was only found to play a role when the proportions of looks to the subject picture were calculated (only over a small time window). Thus, it seems to be the case that the three year old children attended to the verbal and visual stimuli, but the evidence that they detect subject-verb agreement violations is not very robust.

The eye gaze of the four year old children is difficult to interpret. They displayed an effect of grammaticality in Phase 1, which was not expected considering the information in the verbal stimuli. None of the other groups displayed such an effect. When proportion of looks to the target picture and looking times were used as dependent variable, no effect of grammaticality or subject number emerged. When the proportion of looks to the subject picture were analyzed, children seemed to look more towards the subject picture whenever they heard the verb inflection *-t*. It is unclear why this difference might arise. The result pattern obtained for the four year olds is not conclusive.

The five year old children's looking behavior finally was influenced by the experimental conditions. First, they looked to the subject picture during Phase 1 for about half of the measured time stamps, without any unexpected grammaticality effect. Additionally, they preferred the target picture over the distractor picture during Phase 2, thus they processed and comprehended the verbal material. The proportions of looks to target did not reveal an effect for grammaticality when the PLTs were calculated over the whole Phase 2 time window. But when the PLT-values were only calculated over a smaller, earlier time window (5300 to 6500 ms post onset), five year olds were found to look more to the target picture in the grammatical compared to the ungrammatical trials. The same pattern was found in the fixation duration analysis, although this failed to reach significance (A trend was found in the plural trials, but no effect appeared in the singular trials. This is possibly due to the fact that a lot of participants had to be discarded from the fixation-based analysis). Finally, when the latency of the first fixation to target picture was used as dependent variable, five year olds were faster to look at the target picture in grammatical compared to ungrammatical trials. It can be carefully concluded that five year old German children were faster and more accurate at looking towards a target picture, when the object noun was presented in a grammatical compared to an ungrammatical sentence.

8.4 Discussion

This experiment was intended to detect when German-speaking children would show sensitivity to subject-verb agrement violations in a preferential looking task. Three year olds showed a first sign of being sensivite to the agreement relation between subject and verb, but only in the singular subject condition and only when looking times were used as dependent variable. Five year old's looking behavior on the other hand was measurably affected the grammaticality of the test sentences. They looked more to the target picture in the grammatical trials containing correct subject-verb agreement compared to the ungrammatical trials in which agreement violations occurred. It was concluded that they are more accurate at finding a target picture showing the referent of an object noun when subject and verb in the SVO-sentence agreed in number. The same pattern was found by Van Heugten and Johnson (2011) who investigated Dutch-speaking children's sensitivity to the presence of grammatical or nonce determiners in spoken sentences. The two year old children looked more to the target picture when a grammatical determiner (irrespective of gender) preceeded a noun than when a nonce determiner was presented. This is interpreted as indicating that Dutch-speaking children comprehend utterances better when they contain real as opposed to nonce determiners. Zangl and Fernald (2007) as well found a higher proportion of looks to a target picture in grammatical compared to ungrammatical trials, but additionally found that children were faster at detecting the target picture if it was preceded by a grammatical determiner compared to an ungrammatical determiner. Such earlier detection of the target picture in the grammatical trials was apparent in the five year old children in the present experiment as well. Thus, when the object noun followed a verb that did not agree with the subject noun, the five year olds were slower at looking towards the target. This reaction time result is interpreted such that five year olds were *disrupted* in their sentence processing when an agreement violation was encountered. Therefore it is concluded that five year old German children are sensitive to the agreement relation between subject and verb and that they know about the surface manifestation of this morphosyntactic relation.

A number of topics are left to be discussed. First, how can the present findings be reconciled with results from other HPP- and grammaticality judgement studies that tested sensitivity towards the same linguistic structure? Second, why is no sensitivity to agreement found in the adults' looks to the target picture, although their looking behavior towards the subject picture indicates that they have noticed the agreement violation? Finally, how can the findings from the present experiment be reconciled with further comprehension and production data on verb inflection.

Before turning to these questions, let us consider what is actually necessary to detect agreement violations during sentence processing. Many psycholinguistic studies have examined agreement processing in adults, and usually found that adult speakers and listeners are highly sensitive to agreement relations and very adept at detecting agreement violations, at least under normal processing conditions (e. g. Eberhard et al., 2005; Pearlmutter et al., 1999; Solomon and Pearlmutter, 2005).[101] But critically, some researchers have argued that adults do not always compute subject-verb agreement during sentence comprehension (e. g. Frazier, 1987). This shall be discussed further below with regard to the looking behavior found in the adults. But even for researchers that assume regular agreement computation in sentence comprehension, the nature of the specific processes that lead to the checking of agrement relations are not perfectly clear. Researchers are at least united on the following: "Central to the agreement checking must be a process of comparison that evaluates the compatibility between a pair of representations." (Nevins et al., 2007, pg. 82). But whether agreement relations are evaluated in a *bottom-up* or *top-down* fashion is still a matter of debate. In the former case, the search for an agreement controller (the subject) would

[101] Further research shows that agreement processing can easily be disrupted by greater processing demands induced via stress or a noisy speech signal (e. g. Dick et al., 2004; Blackwell and Bates, 1995).

not begin until the verb had been encounterd, while in the latter case, an incoming word (the verb) would have to be compared to a predictively created morphological template.[102] A further matter of discussion is whether multiple agreement features (person, number, possibly gender) are checked in a unitary process or whether partially independent subprocesses exist (Nevins et al., 2007). Finally, it is not clear whether agreement checking is an operation that purely relies on morphosyntactic knowledge or whether the 'notional number' of the subject noun has an impact on comprehension or production. Nicol et al. (1997) and Eberhard et al. (2005) claimed that the computation of subject-verb agreement only involves syntactic aspects of the NP and not conceptual ones. Vigliocco and Franck (2001) on the other hand concluded that conceptual information does influence the syntactic operation of agreement. It has additionally been claimed that crosslinguistic variation might play a role here, such that languages with a richer inflectional system rely more heavily on notional number while languages with a rudimentary inflectional system do not (e. g. Foote and Bock, 2011). This latter question of whether conceptual or semantic number information influences the establishment on an agreement relation, bears relation to the present study. The participants saw either one or two referents of the subject, depending on the number condition of the sentence. Additionally, the subject sentence was first introduced with an explicit numeral, so provide further cues to the number value of the sentence subject. It is unclear whether adults and children they were able to use this information to establish the grammatical agreement relation between subject and verb.

Summarizing, to detect the agreement relation between subject and verb, participants in the present study had to access the morphosyntactic number feature of the sentence subject (possibly with help of the semantic information provided visually and verbally), the morphosyntactic number feature of the verb signalled by the inflectional affix and they needed to compare

[102] The 'directions' of parsing are explained in relation to English, a strict SVO-language.

whether these feature representations were compatible or not. It is thus usually assumed that adults process and compare some kind of abstract representation in the form of features, instead of relying on distributional properties of the speech stream, as suggested for young infants.

To return to the questions stated above, let's first discuss possible reasons of why we did not find sensitivity to agreement violations in three and four year olds, although a similar ability has been found in English-, Dutch- and French-speaking children (Soderstrom et al., 2007; Polisenska, 2010; Nazzi et al., 2011). In Chapter 4.3 it was argued that young children showed sensitivity to verb inflection because of their ability to detect distributional patterns and dependencies in the input. Because children are not able to infer number information from nouns, auxiliaries, verbs or demonstratives before the age of 22 months (e. g. Wood et al., 2009; Barner et al., 2007; Li et al., 2009), it was argued that their early preference for sentences containing correct subject-verb agreement does not rely on underlying knowledge of number features but of knowledge about the distributional patterns of determiners, nouns and verb inflections. Young infants' ability to detect statistical patterns in the linguistic input and to track even remote dependencies has been shown in artifial language learning studies (e. g. Gómez and Maye, 2005; Gómez, 2007). Studies testing infants' knowledge of their native language have found that they are sensitive to distributional patterns and remote dependencies between function words at around 17 months to 19 months (e. g. van Heugten and Shi, 2010; Höhle et al., 2006; Santelmann and Jusczyk, 1998). Therefore, it is claimed infants early capacity to detect subject-verb agreement patterns in the input is solely based on statistical knowledge (e. g. Nazzi et al., 2011; Soderstrom, 2002).

In the present experiment, children were presented with pictures that showed the referents of the subject nouns and the object nouns. The children's looking behavior indicates that they paid attention to the visual stimuli in relation to the test sentences, because they looked at the subject picture upon naming it and they preferred the target over the distractor picture once the object noun had been presented (leaving aside the data of the four year olds for now). This makes it much more likely that children were engaged in a comprehension task which might induce a different kind of processing of the linguistic input compared to a pure preferential listening task. A very similar explanation was put forward by Sundara et al. (2011) who claimed that the 22-months old children tested in their study were engaged in a *referential* task and that "adding a referential context to auditory stimuli renders the sentence processing task much more challenging for young children" (Sundara et al., 2011, pg. 57). In accordance with this, it is claimed that no sensitivity to subject-verb agreement in three and four year old German children was found because of the referential aspect of our preferential looking task, even though such sensitivity has been found for children before the age of 2 years in non-referential HPP experiments.[103]

The claim that meaning adds heavily to the computational requirements necessary to show adequate sensitivity to or comprehension of a linguistic structure is also made by Naigles (2002) and Soderstrom (2008). After giving an extensive overview of experimental findings regarding young children's early language processing capacities, Naigles argues that asymmetries between early abstract knowledge and later (seemingly) non-abstract, item-based knowledge in language learners can best be explained by considering the different *content* of the stimuli used in the studies. Simply put: "learning form is easy but learning meaning, and especially linking meanings and forms, is hard" (Naigles, 2002, pg. 157). We will return to this very important argumentation in the General Discussion of the thesis (Chapter 12), as it might explain result patterns found in detection, comprehension and production tasks of of verb inflection in general.

It should be noted that one reason for adding a referential context in the present experiment was to enhance the likeliness that children would detect the correct number value of the sen-

[103] But note that such early sensitivity before the age of 2 has not been found for German-learning children yet. So an explanation that takes into account specific aspects of the German number system cannot be ruled out on the basis of the current evidence.

tence subject, possibly making the task of detecting subject-verb agreement easier. This referential context might have helped the older children tested, as explained further below. But it is questionable whether young children were able to make use of the number information provided. Very recent work by Barner and colleagues indicates that the number word *zwei* 'two' is not a plural marker in early child language, since children hardly ever used numerals when labelling sets (Barner, Lui, and Zapf, 2012).[104] Other work has shown that English-speaking children can make use of noun plural marking at the age of three (Kouider et al., 2006). Whether young children were able to use the visual information (e. g. one dog vs. two dogs) is unclear as well. Studies on children's processing of garden-path sentences indicates that children up to the age of five have difficulties taking referential context into account during language comprehension (e. g. Trueswell et al., 1999; Snedeker and Trueswell, 2004). On the other hand it has repeatedly been shown that children use the referential context to derive word meaning, thus it is not the case that children are ignorant to the visual context that accompanies spoken language (e. g. Fisher, 2002; Naigles et al., 2005). But whether visually presented referential context aids or hinders children in displaying their language processing capacities seems to be highly determined by the linguistic structure under investigation, the task and the age of the children.

Compared to children's ability to detect subject-verb agreement violations in grammaticality judgement tasks as reported in the literature, the present preferential looking task revealed earlier sensitivity to subject-verb agreement. While five year olds were faster and more accurate at finding a target picture in a sentence containing correct agreement compared to sentences with agreement violations, even school-aged English-speaking children are reported to have considerable difficulties in detecting agreement violations in metalinguistic tasks (e. g. McDonald, 2008b; Wulfeck et al., 2004).[105] This difference can rather

[104] But see Hurewitz et al. (2006) for evidence that three year olds can make use of numerical expressions in sentence comprehension.

[105] Since no data is currently available on how well German-speaking school-aged children detect subject-verb agreement in metalinguistic tasks, again a

easily be explained with varying task demands and additional metalinguistic abilities that need to be in place to perform an explicit metalinguistic judgement. As Bialystok (1986) claims, children need to control the linguistic processing to select the specific linguistic information that is under investigation and they need to analyze what they have selected. In the preferential looking experiment presented here, children 'only' need to linguistically process the incoming speech. An influence on the looking behavior is expected to reflect automatic and unconcious linguistic processing. Thus, the two tasks, preferential looking and metalinguistic judgement, enforce very different demands upon children. Recent studies provide further evidence that more adult-like comprehension can be found in young children when eye tracking methods are employed compared to other methods, like act-out or picture selection tasks (e. g. Chan et al., 2010; Sekerina et al., 2004; Höhle et al., 2009; Bergmann et al., 2011; Beyer and Hudson-Kam, 2009).

An alternative explanation, not only taking into account different task demands but also the nature of the stimuli presented, might lie in the fact that participants were provided with *explicit number information* regarding the sentence subject, visually and verbally. This redundant information was supposed to help the participants to correctly detect the number of the sentence subject, which might have helped them to establish an agreement relation. It might have helped the participants in the present experiment to establish the conceptual number and the abstract number feature of the sentence subject, which might have had a supporting effect for checking the agreement features on the verb. Regarding the visual information to subject number, Zapf and Smith (2008) found that two year olds' rate of noun plural production is influenced by the number and similarity of the items in a visually presented set. Children were more likely to produce noun plural inflection when four instead of two referents were presented and when the instances were identical rather than merely similar. It was concluded that the underlying meaning of the plural affects retrievel of morphological

language-specific explanation cannot be ruled out.

markings (at least for nouns). Recent work by Kouider and colleagues provides evidence that 24-month old children can grasp the meaning of visually presented sets of one versus more than one object in sentence comprehension tasks, although children can relate this difference in number only to sentences in which noun, auxiliary and quantifier are number-marked, not when simply a singular or plural noun is presented (Kouider et al., 2006). It is therefore assumed that at least the five year old children were able to represent the number feature of the sentence subject, although it is unclear how such knowledge influences the establishment of agreement relations.

Psycholinguistic research with adults has provided contradicting evidende on how conceptual number information influences the syntactic processing of the subject-verb agreement relation. Nicol et al. (1997) and Eberhard et al. (2005) claimed that conceptual number does not seem to influence grammatical subject-verb agreement processing in production and comprehension, at least in English. But Foote and Bock (2011) found contrary that notional number of the sentence subject affected verb inflection production in Spanish-speaking participants. Therefore the influence of semantics on the syntactic operation of agreement checking is an open question. Another experiment would need to shed some further light on this, possibly constrasting trials in which subject number information was additionally presented with those in which no futher visual and verbal subject number information is given to participants.

The adults in the present experiment were tested to serve as a control group, but interestingly they revealed a very different looking pattern compared to even the oldest children. Most importantly, adults were not disrupted in their ability to find the target picture when a sentence containing an agreement violation was presented. Nevertheless, they noticed the mismatch between subject number and verb inflection, as evident in their looks to the subject picture. In the ungrammatical trials adults directed significantly more looks to the subject picture than in the grammatical ones. There is ample experimental evidence that adults do process subject-verb agreement during sentence comprehension. Freedman and Forster (1985) for example found that adults were slower in a 'sentence matching task' when the sentence contained an agreement violation between subject number and verb inflection. In this kind of task, adults are presented with two word sequences, one above the other, and their task is to indicate as rapidly as possible whether the two sequences are alike or not. Pearlmutter et al. (1999) found an influence of subject verb agreement on adults' reading times, both in a self-paced reading experiment and in an eye-tracking study. ERP-data additionally revealed evidenc that adult language users compute subject-verb agreement relations in sentence processing (e. g. Osterhout and Mobley, 1995; Münte, Matzke, and Johannes, 1997). For additional evidence see Nicol et al. (1997).[106]

Thus, it is concluded that adults paid attention to the verbal stimuli and that they detected the agreement violation. But the task as well as the verbal and visual material were probably very easy for adults, therefore the test sentences' grammaticality did not affect their ability to find a target picture. Only their looks to the subject picture revealed that they had noticed a difference between the sentences of the two grammaticality conditions.[107] It can be hypothesized that adults' online processing system is flexible enough to listen through a mismatching verb inflection in a simple SVO-sentence when asked to find a target picture. But this is not to say that the mismatching verb inflection went unnoticed, since they 'checked' on the number of the sentence subject by fixating the subject picture when the verb inflection provided them with an unexpected number feature. Zangl and Fernald (2007) argued very similar to explain that they had found an effect of grammaticality in younger children (18 months), but not in the older ones (36 months), when presented with sentences containing a grammatical or an ungrammatical (pseudo) determiner in a preferential looking experiment. The older children were only affected at findings a

[106] But see Frazier (1987) for the claim that adults do not always compute subject-verb agreement during sentence comprehension.

[107] The effect in the looks to the subject picture additionally occurred later, as can be seen in the time course analysis graph and in the fact that no early effect was found in the short time window analysis (5300-6500 ms post onset).

target picture by the presence of an ungrammatical determiner when they were tested on novel nouns that they had just been introduced to. Thus, when the task was altered such that it required more advanced processing and was less predictive, even older, 36-months old, children were affected such that they were slower and less accurate in finding a target picture when the object noun was preceded by an ungrammatical determiner. It can be hypothesized that adults would be affected in their ability to find a target picture if the syntactic structure of the test sentences was more complex or if they would have to find a target picture referring to a just newly learned word.

The last question do be discussed is how the present data regarding German children's sensitivity to agreement violations can be reconciled with findings on when children produce and comprehend verb inflections. Van Heugten and Johnson (2011) found sensitivity to determiner well-formedness at 17 and 20 months of age. In this age range Dutch children usually do not produce determiners on a reliable basis yet, and if they do, they often produce acoustically reduced 'proto-determiners', mostly realized as a schwa. The authors interpret their findings as indicating that Dutch children process determiners in sentence comprehension with more phonological detail than the production pattern reveals (Van Heugten and Johnson, 2011). In the case of Dutch determiners, children's comprehension abilities seem to proceed their production. The same conclusion was drawn by Kedar et al. (2006), who found that children aged 18 and 24 months were sensitive to differences between grammatical and ungrammatical function words, even though none of the infants was said to produce function words in a correct and consistens manner.

This pattern does not seem to hold true for German verb inflections. Considering that children start producing verb inflections around 2 years of age, the fact that we did not find sensitivity before the age of five years seems at first sight puzzling. Of course, it is not possible to directly compare the processing of verb inflection to the processing of determiners as investigated by Van Heugten and Johnson (2011), Kedar et al. (2006) and Zangl and Fernald (2007), since the former are bound morphemes whose correct usage depends on the abstract feature of a remote constituent (the sentence subject) and the latter is a free morpheme which use depends on the constituent within the same phrase. In addition, determiner and noun appear in one syntactic phrase while subject and verb cross a syntactic phrase boundary and can often be remote with various constituents in between (e. g. *The* man *that I had seen in the library last week* was...).

But if children's early productions of verb inflections are viewed as rather item-based and if one further assumes that adult-like knowledge of verb inflections can only be attributed when children are able to correctly produce verb inflections on all kinds of verbs (including pseudo verbs) in all kinds of tasks (i. e. in elicited production tasks), the gap between sensitivity and production is not that big anymore. As Rice et al. (1995) found, correct verb inflection in an elicited production task in English-speaking children can not be found at the age of 3, but rather at the age of 5 years. This is exactly the age when we find German-speaking children to be able to detect subject-verb agreement violations in verbal sentences that were presented in a referential context. In the remainder of this thesis, it will be investigated if slightly younger German-speaking children (aged 3 to 4) are able to use verb inflections in comprehension.

9 Experiment 2: Comprehension of inflection without pointing

In this second experiment[108] *I examined whether German three year olds (and adult controls) were able use verb inflections in sentence comprehension to infer the number of the sentence subject. In a preferential looking task, children were presented with two pictures that only differed in the number of actors performing an action (one girl vs. two girls). The children heard sentences that either contained a verb inflected for* 3^{rd} *person singular or* 3^{rd} *person plural. Since the subject in both types of sentences was an ambiguous pronoun, children had to rely on the verb inflection information to correctly understand whether a sentence referred to a picture showing one girl or a picture showing two girls. Thus, it was investigated whether children understood the connection between the form of the verb inflection and the semantic properties of the subject.*

9.1 Rationale

This experiment aimed to examine whether German-speaking children were able to use verb inflectional affixes as a cue to subject number, and thus if they showed 'comprehension' of verb inflection. Using verb inflections in sentence interpretation has been found to be difficult for English- and Spanish-speaking children, at least when tested with a picture selection task (Johnson et al., 2005; Pérez-Leroux, 2005), as discussed in Chapter 5. We investigated whether German three year old children could use the inflectional affixes marking verbs as 3^{rd} person *singular*

[108] The children's data for this experiment are published in Brandt-Kobele and Höhle (2010) and referred to in Brandt-Kobele and Höhle (2011).

or 3^{rd} *person* plural (present tense) to semantically disambiguate an ambiguous sentence subject, in this case a pronoun. Thus we wanted to know whether German children *understood* the connection between the *form* of the verb inflection and *semantic properties* of the subject. Children's receptive abilities regarding this functional morpheme might shed light on their representation of verb inflections.

9.2 Method

Child language research has shown that linguistic knowledge in children can be assessed easier when using methods with less task demands like preferential looking (e. g. Hirsh-Pasek and Golinkoff, 1996) or eye tracking Sekerina et al. (e. g. 2004); Trueswell (e. g. 2008). We took advantage of such a testing method which requires no obvious reaction of the child and thus recorded the children's eye gaze towards pictures which differed in the number of actors in relation to test sentences in which the verb was either inflected for singular or plural. No further reaction was asked of the children or the adults, who were tested with the same material as a control group.

9.2.1 Participants

Twenty-eight children between 3;0 and 4;1 (mean age: 3;6, 10 girls) participated in this experiment. All children were monolingual native speakers of German from the Berlin/Potsdam area, all without known language deficits and not born prematurely. The children's parents were reimbursed for their travel costs to the lab, the children received a little toy for taking part in the study.

The parents were asked to fill out a short questionnaire, in which they were asked if their children had ever presented any language deficits in the course of development. The questionnaire was further used to obtain information about each child's productive use of verb inflection, as no parallel production task was administered in this study. For this purpose, six unambiguous examples of pronoun-verb combinations were provided and the parents had to mark the forms their child had already produced. These examples consisted of regular inflected verbs, inflected for 1^{st}, 2^{nd} and 3^{rd} person singular and plural, all marked for present tense, e. g. 'ich gehe' (*I go*), 'du gehst' (*you go*), 'er geht' (*he goes*), etc.

This parental information was used to provide a rough estimation of our participants' production of verb inflection and thus extend beyond the spontaneous speech data reported in the literature (e. g., Clahsen, 1986). According to parental reports, three quarters (77 %) of the children produced 3rd person singular inflections at the time of testing, while just over half (51 %) of the children produced 3rd person plural inflections. These findings show that (according to their parents) not all children participating in the study were already producing correctly inflected verbs in their spontaneous speech at the point of testing.

Additionally, 17 adult participants were tested in this experiment, serving as a control group. The mean age of this group was 24 years (range: 21 years to 38 years), and 15 participants were women. All adults were students of the University of Potsdam, and received class credit for taking part in the experiment. Further, all adult participants had learned German as their native language.

9.2.2 Material

To make sure that the verbal affixes were the only available cue to subject number, we created simple SVO-sentences containing pronominal subjects. In German, the personal pronouns for 3^{rd} person singular female (*sie*) and 3^{rd} person plural (*sie*) are homophones, making the test sentences temporarily ambiguous until the inflection marker of the verb has been parsed. The verbs were either inflected for 3^{rd} person singular (*-t*) or 3^{rd} person plural (*-n*), providing the number variation in the experimental design, see Table 9.1.[109]

[109] In German, the pronoun *sie* combined with the inflectional ending *-n* is additionally ambiguous, since it does not only refer to plural subjects, but is also used as a politeness form when adressing an unaquainted or respected

Number condition	Example		
Singular	Sie	fütter-t	einen Hund
	Pronoun-3SG (fem.)	feed-3SG	a dog
	She	*is feeding*	*a dog.*
Plural	Sie	fütter-n	einen Hund
	Pronoun-3PL	feed-3PL	a dog
	They	*are feeding*	*a dog.*

Table 9.1: Experimental condititions (Experiment 2).

Four different verbs were used (‘öffnen’, *to open*, ‘basteln’, *to craft*, ‘füttern’, *to feed*, ‘angeln’, *to fish*). The criteria for verb selection were that the actions would be depictable, the verbs and actions were known to young German children[110] and the verbs contained two syllables when inflected for both number conditions. The latter criterion was included because many German verbs contain two syllables when inflected for 3^{rd} person plural, but only one when inflected for 3^{rd} person singular. We intended to keep the inflectional affix information as the only cue to subject number and avoid any further or possibly confounding information by the number of syllables.

Each verb was combined with two different object nouns each, e. g. ‘Hund’, *dog* and ‘Pferd’, *horse* were used with the verb *feed*, so the girls were either feeding a dog or a horse, to create eight experimental sentences. The object NPs always contained an indefinite article (‘ein’, ‘eine’, *a*). German indefinite articles start with a vowel, while definite articles start with a voiced alveolar plosive (*d-*). To avoid any coarticulation effects of the verb inflectional affix and the following determiner onset, only

adult. We consider it as very unlikely that young children are confused by this ambiguity, as the polite form is very rarely used in child directed speech and because German children start producing it very late (some not even in their elementary school years). Since the polite form is semantically 2^{nd} person singular and thus not felicitous when describing pictures, no interference is expected in this experiment.

[110] German data from the CHILDES-database was searched for the productive use of the test verbs. Three of the verbs were found to be produced before age 3, the fourth one was only found in the input of children, but this again before age 3. The verb used in the practice trials is produced at age 3;4.

indefinite articles were used. An additional verb ('streicheln', *to pet*), again combined with two different objects, was used for the practice trials. All verb-object combinations were presented in both number conditions, yielding four practice trials and sixteen test trials. All verbal stimuli were recorded by a female native speaker of German in a child-directed manner. The mean length of sentences was 1580 ms, ranging from 1392 ms to 1832 ms (mean singular trials: 1575 ms; mean plural trials: 1585 ms). All practice and test sentences are provided in the Appendix (see A).

For each sentence and each number condition, a simple coloured drawing of the described situation was created. The two pictures used for a pair of sentences only differed with respect to the number of actors accomplishing the action denoted by the verb. In the 1-actor picture, only one person was performing the action mentioned in the sentence (e. g. one girl was feeding a dog). In the 2-actor picture, two persons were performing the same action together (e. g. two girls feeding a dog together). See Figure 9.1. To conform to the female pronoun in the singular condition, all depicted characters were girls.

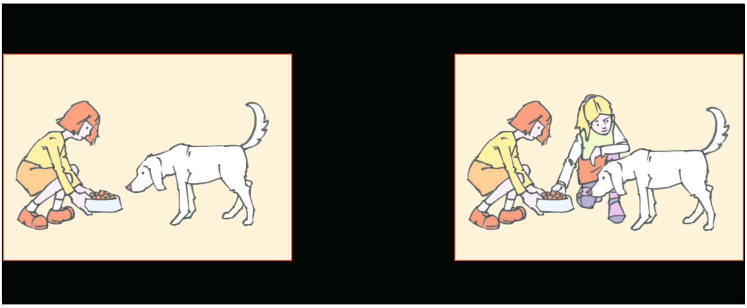


Figure 9.1: Visual material used in Experiment 2. 1-actor picture (left) and 2-actor picture (right).

9.2.3 Design

Pictures were presented pairwise in the experiment, always accompanied by a matching test sentence that contained either a singular or a plural inflected verb. Each participant experienced a given picture pair only once during the experiment, either with a sentence containing a singular inflected verb or with a sentence containing a plural inflected verb. Note, that in the singular number condition, the 1-actor picture served as the target, while the 2-actor picture served as the distractor. In the plural number condition, this relation was reversed, with the 2-actor picture serving as target and the 1-actor picture as distractor.

Whether a given picture pair was presented with the singular or the plural sentence was counterbalanced across children. The location of the target picture (right or left half of the monitor) and the side of presentation of the 1-actor picture were also counterbalanced across the children. The test trials were presented in a pseudo-randomized order, with no more than three items of either number condition, either side of 1-actor picture (left vs. right) and side of target picture (left vs. right) in a row. Each experimental session contained four practice trials without feedback, and eight experimental trials, four containing singular inflected verbs and four containing plural inflected verbs.

In each trial, one pair of pictures was presented side-by-side on the eye-tracking monitor for 3s, accompanied by an attention getting phrase ('Schau mal!' *Look here!*) to direct the child's attention to the screen. These first 3 seconds of picture presentation served as a *baseline phase* and were included to control for initial picture preferences or biases. After that, the screen turned black for 2s, during which time the test sentence was presented auditorily. The sentences were aligned to the visual presentation such that the presentation of the acoustic stimulus ended exactly when the pictures reappeared for the test phase. Thus, the onset of the sentences varied slightly in relation to the disappearance of the pictures at the end of the baseline phase due to the variation in sentence length. After the sentence presentation, the same pair of pictures reappeared again for 3s. This second presentation of the pictures served as a *test phase*.

This was again followed by a black screen for 1 s. Then, the trial ended automatically. Thus, a trial had a duration of 10s. The inter-trial interval lasted about 2s, during which time the screen was blank. See Figure 9.2 for a schematic description of a trial course.

Trial phase	Duration	Auditory stimuli	Left visual stimulus	Right visual stimulus
Baseline phase	3s	'Schau mal!' *Wow, look there!*		
Sentence presentation	2s	'Sie füttert einen Hund!' *She is feeding a dog*		
Test phase	3s			
	1s			

Table 9.2: Time course of a trial (Experiment 2).

After every fourth experimental trial, a short clip (e. g. Elmo jumping up and down) was presented to redirect the child's attention to the screen. The participants' eye movements were recorded throughout the whole experiment, which lasted about 3 min.

9.2.4 Procedure

To measure eye gaze, a tabletop Tobii 1750 eye-tracking system (Tobii Technology AB, Sweden) was used, which tracks eye position with a resolution of 50Hz, thus every 20 ms. Stimulus presentation and eye-gaze data collection was conducted using ClearView®(Version 2.5.1, Tobii Technology AB, Sweden) in a dual-computer set-up.

During the testing session, all participants were sitting in the reclined chair, their heads approximately 60 cm away from the monitor, watching the pictures and listening to the sentences.

They were simply instructed to sit still and listen to the sentences as well as to watch the screen. The parent, in the case of child participants, was sitting behind the child in the corner of the test room and the experimenter sat in another part of the room, not visible to the child, controlling the experiment on the second computer.

As soon as the participant was sitting comfortably and was attentive, a 5-point calibration procedure was performed. Before starting the experiment, the participant was presented with a short pre-recorded story. This story was accompanied by an introductory picture presented on the eye-tracking monitor, which showed three girls standing side by side. The story alerted the participant that he would now see three girls performing some actions, which would either be done by one girl or by two of them together. The story is provided in Appendix A. After this short intro, the four practice trials were presented to acquaint the participant with the experimental procedure. If necessary, children were reminded to watch the screen or not to talk in-between. After the practice trials, the eight testing trials were presented. Since there was no obvious task, no feedback was given to the children before or during the testing trials.

9.2.5 Data Analysis

Eye movements were analyzed automatically by the ClearView®-Software, using the standard settings. The output data that was used for further analysis was one text file per participant providing information about the specific time course of the experiment (e. g. name of each trial, temporal onset and time course of each trial) and the accurate position of the eye gaze (as X-Y-coordinates) at each time point, i. e. every 20 ms. Additionally, the number and duration of *fixations* according to ClearView default settings were part of the automatic output. A fixation is defined as an eye gaze with a minimal duration of 100 ms on a certain area of the monitor that has a maximal radius of 30 pixels.

To analyze the looking behavior in relation to the verbal and visual stimuli presented, we defined two spatial areas of interest ('AoI'). Each area of interest was 400 x 286 pixels in size, corresponding to the size of each of the pictures presented on the monitor. In addition, two temporal areas of interest ('Phase') were defined, one corresponding to the baseline and the other to the test phase; both phases had a duration of 3s.

If children comprehended the inflected verb correctly and if they used the information supplied by the verb inflection to interpret the number of the sentence subject, we expected more and longer looks towards the 1-actor picture in the test phase of the singular trials and more and longer looks towards the 2-actor picture in the test phase of the plural trials. Our expectation was based on the finding that children usually prefer the visual stimuli that match the the auditory information in the standard intermodal preferential looking procedure (Hollich, Hirsh-Pasek, and Golinkoff, 1998; Houston-Price et al., 2007).

9.3 Results

9.3.1 Dependent measures

The looking behavior was analyzed in three different ways. For the first two analyses (time course visualization and proportion of looks analysis), the looking behavior at all measured data points was taken into account, while in the last one, only actual fixations with a minimum duration of 100 ms on a maximal radius of 30 pixels entered analysis.

First, looks to the 1-actor and the 2-actor picture were averaged for every given time stamp (every 20 ms) over trials and participants in each number condition and group, yielding a **percentage of looks** measure, which is depicted in Figure 9.2 for every time stamp over the course of a trial. This served for a first rough inspection of participants' looking behavior towards the 1-actor picture and the 2-actor picture in both number conditions over the whole course of a trial.

Second, the **proportion of looks to the target picture (PLT)** was calculated for each phase, condition and group (see Chapter 9.3.3). PLT was calculated by dividing the number looks to the target picture by the sum of all looks that were directed to target and distractor picture (Tar/(Tar+Dis). The mean PLT values for each group, phase and number condition can be seen in Table 9.3. These values were first compared to the chance level 0.5, which represents equal amount of looks to both pictures, and second entered into two separate two-way analysis of variance (ANOVA). In these 2*2-ANOVAs, one for adults and one for children, Phase (baseline, test) and Verb number (singular, plural) were compared as within-subject factors. Only F1-analyses were performed to account for the low number of experimental trials per participant.

Third, **fixations durations** were used to calculate summed 'looking times' (see 9.3.4). For this, the fixation durations to the 1-actor and the 2-actor picture were added within and over trials separately for the baseline and the test phase as a function of the number condition, separately for every participant. These summed fixations were then averaged over participants. Mean fixation durations were again entered into statistical analyses, as they were subjected to four separate 2x2 ANOVAs, with Number and Phase as within-subject factors. Separate ANOVAs were calculated for the looking times to the 1-actor picture and the 2-actor picture (Picture can not be regarded as an independent factor, such as verb number, because the latter can be controlled by the experimenter, while the amount of looks that are directed to either picture are not). Separate ANOVAs were performed for the group of children and adults, because adults only served as control participants (and no adults were tested in Experiment 3). Already Golinkoff et al. (1987) used the total visual fixation time towards matching and non-matching pictures as a dependent measure when analyzing preferential looking data.[111]

[111] But see Plunkett and Schafer (1998) for a discussion on the theoretical and empirical value of the total visual fixation time as a preferential looking measure.

9.3.2 Time course

For a first visual inspection of children's and adults' looking behavior over the course of a trial, looks to 1-actor and 2-actor pictures were plotted separately as a function of the number condition (see Figure 9.2). Presented is the mean percentage of looks to the 1-actor and 2-actor pictures during the course of the trial, at any given time-stamp, averaged over the four experimental trials in the singular number condition (left panel), and in the plural number condition (right panel). The looking behavior of the adults (N=17) is depicted in the top row, while the looking behavior of three year olds (N=28) is depicted on the bottom row.[112]

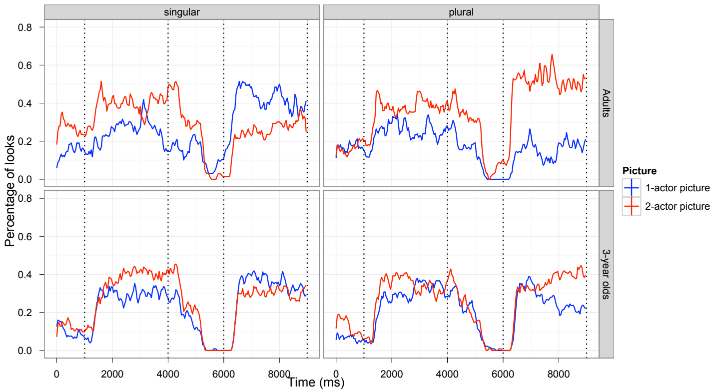


Figure 9.2: Time course of percentage of looks to 1-actor and 2-actor picture, separated by Number condition and Group. (The dotted lines represent the onset and offset of the baseline and test phase.)

The time course plot depicting the looking behavior over the whole course of a trial (Figure 9.2) shows that children looked at the pictures during the baseline and during the test phase.

[112] Note that 40 % of looks at one picture at a given timestamp refers to 40 % of all looks (averaged over trials and participants), not just 40 % of those looks that were directed at the pictures. For this reason, the sum of the percentage of looks directed at either picture a or b in Fig. 2 is less than 100 %.

While the test sentence is presented (4 to 6 s after trial onset), looks in the two spatial AoIs drop close to zero, which can be considered a consequence of the blank screen during this phase. In the baseline phase there seems to be a higher percentage of looks towards the 2-actor picture, irrespective of number condition or group.

In the test phase, the proportions of looks to the 1-actor and 2-actor picture seem to vary as a function of the number condition, especially in the group of adult participants. For those participants, the percentage of looks to the 1-actor picture (blue line) outranks the percentage of looks to the 2-actor picture (red line) in the singular condition over the whole course of the test phase. Vice versa, the percentage of looks to the 2-actor picture is much higher than the percentage of looks to the 1-actor picture in the test phase of the plural number condition.

For the children, the time course graphs are less conclusive. While the percentage of looks to the 2-actor picture seems to outrank the percentage of looks to the 1-actor picture in the test phase of the plural number condition, the advantage of the percentage of looks to the 1-actor picture in the singular number conditions is only barely visible. To test whether adults' and children's looks to the 1-actor and 2-actor picture actually differed reliably as a function of number condition, further eye-tracking data analyses as well as statistical analyses were performed.

9.3.3 Proportion-based analysis

The looks to the target picture were set in relation to the overall looks to the two presented pictures, providing a proportional measure of the amount of looks to target (PLT). Such a PLT-value was calculated for the baseline and test phase of each trial and then averaged for each number condition and group. To level out the effect the initial preference for the 2-actor picture has in the different number conditions, data was additionally collapsed over number condition. All PLTs are listed in Table 9.3.

Group	Phase	verb number	PLT (*Std.dev.*)	t-test vs. chance
Children (N=30)	baseline	singular	0.441 (*0.136*)	$t(27)$=-2.295*
		plural	0.548 (*0.121*)	$t(27)$=2.115*
		collapsed	**0.495 (*0.139*)**	$t(27)$=0.298
	test	singular	0.518 (*0.168*)	$t(27)$=0.557
		plural	0.552 (*0.125*)	$t(27)$=2.18*
		collapsed	**0.535 (*0.148*)**	$t(27)$=1.79(*)
Adults (N=30)	baseline	singular	0.393 (*0.094*)	$t(16)$=-4.703***
		plural	0.589 (*0.099*)	$t(16)$=3.72**
		collapsed	**0.491 (*0.137*)**	$t(16)$=-0.539
	test	singular	0.6 (*0.208*)	$t(16)$=1.989(*)
		plural	0.736 (*0.184*)	$t(16)$=5.301***
		collapsed	**0.668 (*0.205*)**	$t(16)$=3.976**

Table 9.3: Proportion of looks to target (PLT) for each group and phase (signif. codes: *** <0.001, ** <0.01, * <0.05, (*) <0.1).

Adults As adults were simply tested as a control group, their data is analyzed separately and presented first. The PLT-values for each test phase and number condition were statistically compared to chance level 0.5, which represents an equal amount of looks to both pictures, via paired t-tests. The t-values plus level of significance are listed in Table 9.3. When the adult data was collapsed over number conditions, the effect of verbal material on the looks to the target picture became apparent. The PLT did not differ from chance level in the baseline phase, but was significantly above chance level in the test phase.

When the data was analyzed separately for the number condition, the picture got a little more complex. In the baseline phase, the PLT values in both number conditions differed from chance level. As PLT is lower than 0.5 in the singular condition, but higher than 0.5 in the plural condition, these values again represent the baseline preference for the 2-actor picture, which is the to-be target in the plural, but not in the singular trials. A highly significant difference between PLT and chance is found in the test phase of the plural trials, but only a marginal significant effect in the singular trials. It seems to be the case that the baseline preference for the 2-actor picture lasts into the test phase, lowering the PLT in the singular condition but increasing it in the plural condition. Nevertheless, adults look to the target picture in both number conditions more than would be expected by chance.

The PLT-values were additionally entered into a 2x2 ANOVA to analyze the effects of Phase and Number condition. It was found that the PLT values in the test phase were significantly higher than in the baseline phase, indicated by the main effect for Phase ($F(1,16)=11.65$, $p=0.003$). Further, the main effect for Number condition was significant ($F(1,16)=30.311$, $p<0.001$), because the proportion of looks in the plural trials were higher than those in the singular trials. No interaction between the factors was found. Adult participants were thus able to use verb inflection information to find a matching picture. They looked more to the target picture in the test phase compared to the baseline phase in both number conditions, and looks to the target picture exceeded chance level.

Children For the three year old children, the PLTs were analyzed in the same way. First, PLTs were compared to chance level 0.5. When the data was collapsed over number condition, children's results were similar to the ones obtained for the adults, although less pronounced. As for the adults, PLT did not differ from chance level in the baseline phase, but PLT was above chance in the test phase, although this effect was only marginally significant.

When the PLTs from the separate number conditions were compared to chance level, the pattern was again very similar to the one found in the adult control group. In the baseline phase, PLT in the singular number condition was significantly below chance level and PLT in the plural number condition was significantly above chance level. This pattern was just a simple reflection of the preference for the 2-actor picture. In the test phase, PLT only differed from chance in the plural but not in the singular trials. It again seemed to be the case that the singular verb inflection was not 'strong enough' to override the 2-actor picture preference. The PLT in the plural trials seemed to profit from the visual preferences of the children, because in this condition the PLT exceeded chance level significantly. Nevertheless, when data was collapsed over verb number condition, the proportion of looks to the target picture was slightly higher than would be expected by chance, indicating that the children's looks were at least to some extent influenced by the verbal input.

Additionally, PLTs were entered into a 2x2 ANOVA to analyze the effect of Phase and Number condition. This ANOVA revealed a significant main effect for Number condition ($F(1,27)=5.622$, $p=0.025$) and a very marginal trend for a Phase X Number interaction ($F(1,27)=3.034$, $p=0.093$). To disentangle the marginal interaction, the effect of Phase was investigated separately for both number conditions.

In the singular trials, the difference between the baseline and the test phase PLT was marginally significant ($F(1,27)=3.317$, $p=0.079$), while the PLT values in the baseline and the test phase of the plural condition did not differ ($F(1,27)<1.0$). Thus, one can carefully conclude that the presentation of the singular inflected verb affected the 2-actor preference, which was constantly found in the baseline looking behavior of the two year olds. No decrease in looks to the 2-actor picture was found when the children were presented with a plural inflected verb.

9.3.4 Fixation-duration analysis

In an alternative way to analyze eye tracking data, only actual fixations are taken into account. This filters out all the eye gazes that are too short to be considered as part of information processing. Such an approach is taken for example by Bergmann et al. (2011) and Poltrock (2011).

The durations of the fixations directed to the 1-actor and the 2-actor picture were added for each participant within and across trials for each condition. These values will in the following be called 'looking times'. These summed looking times were then averaged across participants, separately for the two number conditions and the two testing phases. The mean looking time for the singular number condition and plural number condition during the baseline and the test phase are plotted in Figure 9.3.

The graphs show that, during the baseline phase, participants in both groups looked longer at the 2-actor picture than at the 1-actor picture, irrespective of number condition. In the test phase, looking times seemed to vary as a function of number condition, because the participants looked longer towards the 1-actor picture in the singular number condition and longer to the 2-actor picture in the plural number condition. This effect is much more pronounced in the group of adults, but is still visible in the group of three year olds as well. For inferential statistics, only F1-analyses were performed on account of the low number of experimental trials in the experiment.

Adults Separate ANOVAs were conducted for two pictures. Looking times were calculated in the usual way (described in detail on page 202). The looking times are depicted in Figure 9.3.

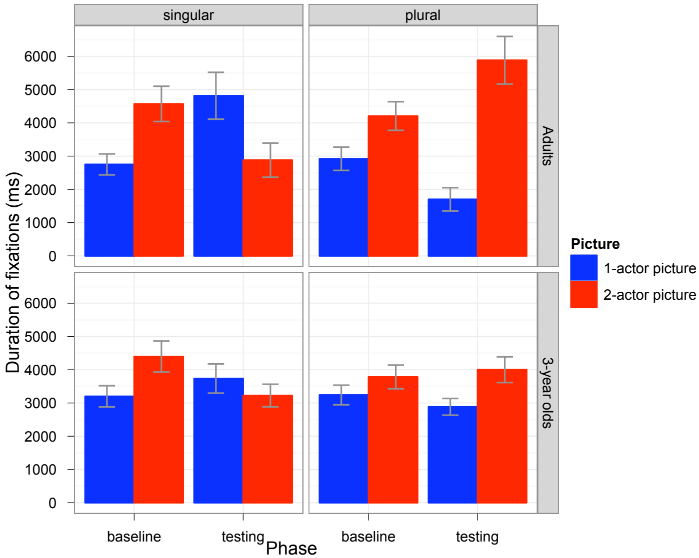


Figure 9.3: Mean summed duration of fixations for baseline and test phase, separately for singular and plural number condition and adults and three year olds. (Errorbars: $+/-$ 1 SE).

1-actor picture The mean durations of fixations that landed on the 1-actor picture were entered into a 2x2 ANOVA comparing Phase and Number condition. This revealed a main effect for Number condition ($F(1,16)=11.483$, $p=0.004$) and a significant Phase X Number interaction ($F(1,16)=12.654$, $p=0.003$). To gain further insight into this interaction, fixation durations in both number conditions were analyzed separately. For the singular trials, a significant effect of Phase ($F(1,16)=10.179$, $p=0.006$), confirmed that the looking times to the 1-actor picture increased significantly from baseline to test phase. For the plural trials, on the other hand, the significant effect of Phase ($F(1,16)=8{,}851$, $p=0.009$) signaled that the looking times to the 1-actor picture decreased from baseline to test phase. Overall, the information of the verb inflection directed adults' looking behavior even when only looks towards the 1-actor picture are considered. When adults heard a singular inflected verb, the looking times on the 1-actor picture increased, and when they heard a plural inflected verb, the looking times to the 1-actor picture decreased.

2-actor picture The mean durations of fixations that were directed to the 2-actor picture were entered into a separate 2x2 ANOVA, again comparing Phase and Number condition. In this analysis, a main effect of Number condition ($F(1,16)=13{,}250$, $p=0.002$) and a significant Phase X Number interaction emerged ($F(1,16)=9.601$, $p=0.007$). To solve the interaction, fixation durations for both phases were compared separately in the singular and plural trials. In the singular trials, a significant effect of Phase ($F(1,16)=8.172$, $p=0.001$) confirmed that the looking times to the 2-actor picture significantly decreased from baseline to test phase when participants heard a singular inflected verb. In the plural trials, this pattern was reversed. A significant effect of Phase ($F(1,16)=6.905$, $p=0.018$) confirmed that the looks to the 2-actor picture increased from baseline to testing, but only when participants had heard a plural inflected verb in the test sentence. Overall, the looking times to the 2-actor picture again showed that adults' eye gaze was directed by the verb inflection in the test sentence. When the participants were presented with a singular inflected verb, looking times to the 2-actor picture decreased, but when they were presented with a plural inflected verb, looking durations significantly increased.

Three year olds Separate ANOVAs were conducted for two pictures. Looking times were calculated in the usual way (described on page 202). The looking times are depicted in Figure 9.3.

1-actor picture The mean fixation durations that the children directed to the 1-actor picture were entered into a separate 2x2 ANOVA comparing Phase and Number condition. This analysis yielded a marginally significant main effect for Number condition ($F(1,27)=4.973$, $p=0.054$), without any further effect or interaction. Still, to investigate the looking times in the separate number conditions further, post-hoc one-way ANOVAs were performed. No effect of Phase was found in the singular or in the plural trials. Therefore, children's looking times towards the 1-actor picture do not reveal any evidence that verb inflection influences their eye gaze.

2-actor picture To compare children's looking times towards the 2-actor picture, a separate 2x2 ANOVA comparing Phase and Number condition was performed. This revealed a significant main effect for Phase ($F(1,17)=5.297$, $p=0.029$) and a significant interaction between the two factors ($F(1,27)=8.623$, $p=0.007$) under investigation. The interaction warranted further exploration, which revealed that the children had a significant decrease in their looking times from the baseline phase to the test phase in the singular number condition ($F(1,27)=11.235$, $p=0.002$), while no significant decrease in their looking times towards the 2-actor-picture from the baseline to the test phase occurred in the plural number condition ($F<1.0$). Thus, when we tested whether the presentation of the test sentences had a differential effect on children's preference for the 2-actor picture as evidenced in the baseline phase, we did find that verb inflection affected children's eye gaze. The preference for the 2-actor picture vanished after the presentation of a singular inflected verb, but stayed constant after the presentation of a plural inflected verb.

9.3.5 Summary of Results

In Experiment 2, we investigated whether participants were able to use verb inflection during sentence comprehension to infer the number of an ambiguous sentence subject. This was tested using a preferential looking task. We measured whether participants looked more, longer and/or faster to a picture that was consistent with the number information provided by the verb inflection. Correct comprehension was indicated by more and longer looks to the 1-actor picture after hearing a singular inflected verb, and to the 2-actor picture after hearing a plural inflected verb.

Summarizing our results, the following picture emerges. When the children and adults were presented with a singular sentence, their general preference for looking at the 2-actor picture vanished. This was not the case when hearing a plural sentence. In this case, adults showed an increase in looking time to the 2-actor picture while children's looking times stayed constant. This pattern was found when calculating proportion of looks and looking times for analysis. It suggests that, without a verbal stimulus that is systematically related to one of the pictures, the children tend to look more and longer at the 2-actor picture. This may reflect the fact that the 2-actor picture shows an additional person, is thus informationally more complex than the 1-actor picture and for this reason attracts attention for a longer period of time. The differences in the looking times at the two pictures that were found during the baseline phase emphasizes the necessity to include such a control phase in experiments using this method, as only the changes from baseline to test phase can reveal the effects of the sentence presented.

The finding that the children's and adults' looking times at the 2-actor picture decreased after the presentation of a singular inflected verb in the test compared to the baseline phase suggests that the preference for the 2-actor picture has been overridden by the presentation of a sentence that better describes the 1-actor picture. It is therefore concluded that the sentence presentation drives the children's and adults' attention to the corresponding target picture. This pattern additionally discards the possibility that children consider the 2-actor picture as a felicitous reference for a singular inflected verb. Thus, we conclude, that the children and adults in this study did process the singular verb inflection (*-t*) and that they were able to use this information to correctly infer the number of the sentence subject. When a plural sentence was presented, a similar pattern of increased looking times at the target picture was only observed in the adults, but not in the group of children. In this case, the target picture corresponded to the picture that the children already had preferred in the baseline phase. Nevertheless, the observation that looking time from the baseline to the test phase stayed constant indicates that the presentation of the plural test sentence has kept children's attention on the 2-actor picture. This in turn suggests that children have processed the number information given by the verb inflection in the plural condition (*-n*) as well. Adults clearly have processed and comprehended the plural inflection, considering the significant increase in looks towards the 2-actor picture in the plural condition.

9.4 Discussion

In this experiment, it was examined whether German-speaking children aged three to four years were able to make use of verb inflection in sentence comprehension to infer the number of an ambiguous sentence subject. Eye-gaze towards pictures that either depicted one or two actors preforming an action were measured as children (and adult control participants) heard sentences in which the verb was either inflected for 3^{rd} person singular or 3^{rd} person plural. Adults' looked more to the 2-actor picture before the presentation of a verbal stimulus, but looked more to the target picture in both number conditions after they had heard the test sentences. Children's looking patterns were very similar, although differences in relation to the verbal stimulus were less pronounced. As the adults, children looked more to the 2-actor picture before the presentation of a verbal stimulus, and this preference vanished when they were presented a singular inflected verb, but stayed constant when they were presented with a plural inflected verb.

The finding hat the looking pattern in the singular trials provided clearer evidence of children's ability to use of verb inflections to infer the sentence subject than the plural trials needs further consideration. This contrasts with findings from Lukyanenko and Fisher (2010) and Grüter and Fernald (2011), who better results for plural inflected verbs . The critical difference might lie in the verbal material, since we used inflected main verbs but the other studies examined children's comprehension of auxiliaries. Three possible explanations can be considered for our result pattern.

A first reason relates to grammatical assumptions underlying the number contrast. To explain the phenomenon of 'agreement attraction errors' in adult sentence production and comprehension (which occur more often with plural subjects than singular subjects), researchers have put forward the assumption that plural is marked while singular is unmarked (e. g. Eberhard et al., 2005; Wagers, Lau, and Phillips, 2009). Agreement attraction errors, in which an agreement-bearing element (verb) fails to match the agreement features on its grammatical controller (subject) but instead realizes agreement with a nearby distractor noun, are well documented in spontaneous speech recordings and can even be found in well-edited textbooks or newspapers (Wagers et al., 2009). The usual pattern looks as follows. Adults make more agreement mistakes in production and take longer in comprehension tasks when they process sentences containing a singular subject NP combined with another plural NP (e. g. *The key to the cabinets is/*are...*) than when they have to process sentences containing a plural subject NP combined with another plural or singular NP (e. g. *The keys to the cabinets *is/are...*). One of the explanations for this kind of error pattern assumes feature-perculation and the idea that singular nouns are unmarked, thus lack a number feature, while plural nouns are marked and possess a number feature (e. g. Eberhard, 1997). However, recent works in semantics has questioned the assumption of plural as the marked form (e. g. Sauerland, Anderssen, and Yatsushiro, 2005). Nevertheless, Grüter and Fernald (2011) call on the attraction effect to explain why they have found 24-months olds to be able to use plural marking preceding a noun as a predictive cue in language processing (but not singular marking).

The idea that plural nouns are marked, in the sense that they possess a feature that needs to be checked in sentence comprehension, while singular nouns do not (e. g. Eberhard, 1997; Pearlmutter et al., 1999), could be applied to the result pattern of the current experiment. Under the assumption that pronouns should have an underlying number feature as full subject NPs do, one could assume that the (additional) step of checking a marked feature might be demanding for young children, thereby making plural verb inflection more difficult to use in sentence comprehension, simply because this feature has to be checked. In this idea, the absence of a feature and therefore the lack of need to check it would provide an advantage for singular sentences. Thus we would expect to find a similar pattern of behavior in processing of other marked vs. unmarked feature pairs, with the marked feature condition not evincing much of a change from baseline behavior, and the unmarked feature condition showing a change from baseline behavior. Note also that, according to this analysis, the fact that children prefer the 2-actor picture in general is not related to the fact that plural is the marked value, and so, had the 1-actor picture been the preferred one, we would have expected to find no change in looks in the plural (marked) condition, and increased looking to the 1-actor picture in the singular (unmarked) condition.

The second possible explanation relates to the acquisition of the verb inflection paradigm, claiming that the 3^{rd} person plural inflection might be acquired later than the 3^{rd} person singular inflection. This might be due to frequency of morphological forms in the input or syncretism between 3^{rd} person plural inflection and the German infinite verb ending or some other unknown factor. One could claim that the German children have yet only acquired a rule regarding verbs with a *-t* affix or a more abstract rule targeting all 3^{rd} person singular verbs and are not able to infer what a verb carrying an *-en* affix should mean. This would mirror acquisition data as found for verb inflection production. Production studies suggest that singular inflections appear earlier and with higher rates of correct use than plural inflection (e. g. Poeppel and Wexler, 1993; Clahsen, 1986), a pattern that has been found in the parental report for the participants of the present study as well. This explanation would then entail that the children's behavior did not change in the plural condition because they didn't yet understand plural marking (as well).

The third explanation relates to the visual material used in the experiment. The visual baseline preference for the 2-actor picture might have obscured further effects when this picture was the target. A preference for pictures that show multiple entities has also been found by Soderstrom (2002). Interestingly, the proportion of looks analysis revealed better than chance performance in the plural trials, not in the singular ones. Thus, the children's ability to keep their eye gaze on the 2-actor picture when being presented with a plural inflected verb is interpreted as evidence that they process and comprehend the plural inflectional affix. The different looking patterns in the test phase of the two number conditions are critical to validate the claim that children 'kept' their gaze on the 2-actor picture after hearing a plural inflected verb. This is only revealing in combination with the finding that children did not keep their gaze on the 2-actor picture after hearing a singular inflected verb. Contrary, the drop in looks to the 2-actor picture in the singular trials should not be explained by a lack of interest that simply occurred because the pictures had been visible for a while.

Therefore, based on the eye-tracking results we conclude that German three to four year old children have acquired the morphosyntactic knowledge about number agreement in the domain of the German 3^{rd} person verb inflections, and that they can use this knowledge when determining the number of an ambiguous sentence subject. This finding contrasts with the earlier studies testing English- and Spanish-learning children who were not found to be able to make use of the number information provided by the verbal inflection before the age of five to six years (Johnson et al., 2005; Pérez-Leroux, 2005). One possible explanation for the earlier comprehension evidenced by German children could be the morphosyntactic differences across these languages. While German has a rather rich system of verb inflection, in English main verbs are only marked for 3^{rd} person singular and past tense. The greater variety of inflectional endings in German may allow German children to acquire the verb inflections representing number differences earlier than their English-speaking peers (e. g. Guasti, 2002; Phillips, 1995). But the relevance of morphological richness is questioned by the findings from the Spanish-learning children who do not demonstrate earlier comprehension of verb inflections than the English-learning children (Pérez-Leroux, 2005).

A second possible explanation for the discrepancy between our findings and those of the earlier studies could be the different methodological paradigms employed. In contrast to the previous studies which used a picture selection task, children in this experiment did not have to perform any particular action but were only instructed to watch the pictures and listen to the sentences. This difference in task demands, merely looking in the eye tracking study versus deciding and pointing in a picture selection task, may account for the different findings. To test this possibility, a second experiment was conducted, in which the eye-tracking technique was combined with an explicit picture selection task. This experiment is presented in Chapter 10. An impact of task has been reported for the production of verb inflection, since children's spontaneous speech data usually reflect higher levels of competence than elicited production tasks. And an impact of task demands has also been found for the comprehension of other linguistic structures than verb inflections, e. g. pronouns (e. g. Bergmann et al., 2011) and focus particles (e. g. Höhle et al., 2009).

According to the present findings, the ability to use verb inflections in sentence comprehension is in place between three- and four-years of age, at least in German-speaking children. This suggests that children have access to a rather abstract, almost adult-like rule (see 2.5) for processing verb inflection information. If the children tested only had a rule that would link a certain kind of verb form (a verb with a *-t* affix) to certain subject forms (NPs plus a plural marker or pronoun), they would need an additional rule that would link the pronoun form to a meaning regarding number. Since the German pronoun *sie* is homophone for singular and plural, such an additional rule that is based on the surface form of the pronoun is hard to conceive. Therefore, it is concluded that German children aged three and a half have acquired an abstract adult-like rule that at least allows them to link a *-t* inflected verb to a singular subject and and possibly an *-en* inflected verb to a plural subject.

The findings further call into the question the existence of an asymmetric acquisition path regarding production and comprehension of verb inflection, that had been postulated (see Chapter 6.3). Since it is concluded that the structural competence necessary to use the information provided by verb inflections can be assessed in three to four year old German children when appropriate measures are employed, receptive abilities seem to be in place earlier than previously assumed.

Concerning production of German verb inflections, we can rely on the finding that according to our parental report not even all the tested children were producing the verb inflections in question. This matches the findings from Rice and Wexler (2001), de Villiers and de Villiers (1973a) and Rice et al. (1995) who did not find more than 90 % of children producing verb inflections before age four. Additionally, the parental report data and the cross-sectional studies point to the problem of generalizing findings from a very small groups of children or single case studies (Poeppel and Wexler, 1993). Taken together, the production data indicate that most children only use verb inflection on an adult level around the age of four years or even later. Theoretical implications of this will be discussed in the General Discussion (Chapter 12). But this matches the age that we found German children capable of using verb inflection in comprehension. Therefore, we conclude that our data question the hypothesis of an asymmetrical development of verb inflection. Our findings emphasize that empirical evidence that seems to support the hypothesis of an asymmetrical development must be evaluated critically with respect to the methods that generated it.

Still, the result patterns that we obtained for the children are not as clear as the ones obtained for the adult controls. It seems to be the case that making use of verb inflections in sentence comprehension is not an easy endeavor. One can hypothesize that the use of more familiar or frequent verbs with simpler phonological structure as test items might improve children's performance, as children's production of verb inflections is enhanced in familiar verbs (e. g. Theakston et al., 2003) or verbs that have a simpler phonological structure (e. g. Song et al., 2009). Another possibility to improve children's performance would be to use intransitive sentences with verbs in final position, since verb inflections are found to be easier to produce and detect in sentence-final compared to sentence-medial position (e. g. Sundara et al., 2011). These ideas should be examined in further research. On the other hand, one can as well hypothesize that any additional load that makes the task or the verbal stimuli more difficult, like using a different methodology like Johnson et al. (2005), or using less known subject nouns like Leonard et al. (2000) can make children's comprehension performance even worse, to the point where no comprehension can be found. This latter possibility, more precisely the impact of task, was examined and is presented in the following two experiments (Chapter 10 and 11).

10 Experiment 3: Comprehension of inflection with pointing

The results of Experiment 2 show that German children can make use of verb inflections during sentence comprehension to infer semantic meaning of the sentence subject. This contrasts with earlier findings from English- and Spanish-speaking children who were not able to comprehend verb inflection as a cue to subject number before the age of five or six years. Critically, earlier studies had employed a picture selection task while children in Experiment 2 were tested using the preferential looking paradigm. To investigate whether the earlier comprehension ability we found should rather be attributed to language-specific properties or methodological differences, we conducted a another comprehension experiment in which children's eye gaze was tracked while they had to perform a picture selection task.[113]

10.1 Rationale

The results of Experiment 2 (Chapter 9) indicate that German children at age three are able to comprehend the 3^{rd} person singular and 3^{rd} person plural verb inflections and use this information to infer the number of an ambiguous sentence subject. This contrasts with the findings reported in Chapter 5 for English- and Spanish-learning children who were not able to make use of the number information provided by the verb inflection before the age of five to six years (Johnson et al., 2005; Pérez-Leroux, 2005; Leonard et al., 2000).

[113] The children's data for this experiment are published in Brandt-Kobele and Höhle (2010) and referred to in Brandt-Kobele and Höhle (2011).

One possible explanation for the discrepancy between our findings and those of the earlier studies could be the different methodological paradigms employed. In contrast to the previous studies which used picture selection tasks, children in Experiment 2 did not have to perform any particular task but were only instructed to watch the pictures and listen to the sentences. It was hypothesized that this difference in task demands, i. e. merely looking in the eye tracking study vs. deciding and pointing in a picture selection task, may account for the different findings.

10.2 Method

To test the influence of task demands on children's comprehension of verb inflection, a third experiment was conducted. In this experiment, the eye-tracking technique was combined with an explicit picture-pointing task, thus children had to make a picture selection decision while their eye gaze was tracked.

10.2.1 Participants

Twenty-eight children participated in this pointing-while-looking experiment. Their mean age was 3;8 years (min: 3;2 years, max: 4;4 years) and 15 of the participants were girls. All children were monolingual speakers of German, did not suffer from any known language disorders and were not born prematurely. Parents were asked to fill out the same questionnaire to control for any history of language deficits and gather information about children's productive use of verb inflections, just as in pure eye-tracking experiment (see Chapter 9). According to the parental report, 72 % of the children produced 3^{rd} person singular inflections at the time of testing, and 48 % of the children produced 3^{rd} person plural inflections. This shows that not all children tested produced verb inflections at the time of testing. Additionally the group does not differ from the one tested in Experiment 2 (Chapter 9) with respect to verb inflection production (3^{rd} person singular inflection: $chi^2(1)=0.29$, $p>0.1$; 3^{rd} person plural inflection: $chi^2(1)=0.38$, *p*>0.1). Only child participants were tested in this experiment, as adults' ability to look at the correct picture had already been shown in Experiment 2. It was assumed that adults would be as well able to point to a matching picture when presented with the test sentences.[114]

10.2.2 Material

The same verbal and visual material was used as in Experiment 2, the pure eye-tracking experiment, presented in Chapter 9.

10.2.3 Design

The same design was used as in Experiment 2, the pure eye-tracking experiment, presented in Chapter 9. Every participant was presented with four practice and eight pseudo-randomized testing trials.

10.2.4 Procedure

In this 'pointing-while-looking' experiment, children had to perform a picture selection task while their eye gaze was tracked. Thus, the procedure was slightly different compared to Experiment 2. Most importantly, before testing began, children were instructed to point to the picture which they thought would match the auditorily presented sentence. To monitor the children's pointing reactions, the experimenter was sitting next to the child and, if necessary, encouraged her to select a picture.

Before the eye-tracking experiment started, children were presented with two pictures printed on a sheet of paper, one picture showing one boy and the other one showing two boys. Children were prompted to point to each picture, by being asked 'Wo ist/sind *ein Junge/zwei Jungen*? Siehst du das Bild, auf dem *ein Junge/zwei Jungen* zu sehen ist/sind?' ('Where is/are *one boy/two boys*? Do you see the picture with *one boy/two boys*?'. This pre-testing phase was included to ensure that all children

[114] See Experiment 4, Chapter 11, for further experimental support of this assumption.

understand the difference between one and more than one actor, and that they were able to point to a picture when asked. All children performed well at this pre-test. Some had to be prompted more than once but finally came up with a correct pointing reaction.

A further variation affected the duration of the testing phase. To account for the fact that young children usually need some time to select a picture, the testing phase was extended from the 3 s allotted in the pure pointing experiment up to a maximum of 15 s. Still, as soon as a child had selected one of the pictures, the experimenter started the next trial via a button-press. Children did not receive feedback regarding the correctness of their pointing reaction. The experimenter simply said 'Yeah, very good' or 'Okay', when the children pointed to the screen, no matter which picture they were pointing at. In all further aspects the procedure and the technical equipment was the same as Experiment 2.

10.2.5 Data Analysis

First of all, the pointing reactions were analyzed. They were averaged per number condition and then compared to chance level (0.5) for statistical analysis. In the eye gaze analysis, only the looking behavior in the first three seconds of the testing phase of each experimental trial was considered so as to keep data analysis comparable to Experiment 2. Temporal Phases as well as spatial AoIs were defined as in Experiment 2.

10.3 Results

10.3.1 Dependent measures

Pointing reactions were noted on a protocol sheet during the testing session as pointings to the left or to the right picture. Later, these reactions were coded as pointings to the 1-actor or the 2-actor picture, separately for the two number conditions. Pointings to the 1-actor picture in the singular number condition as well as pointings to the 2-actor picture in the plural number condition counted as correct reactions. The mean percentages of correct reactions per number condition were statistically compared to chance level performance (50 %) via paired t-tests.

Again, looking behavior was analyzed in three different ways. First, the pure **percentage of looks** to the pictures relative to number condition and testing phase was used to create a visual description of the looking behavior during the course of a trial. Second, the **proportion of looks to the target picture (PLT)** was calculated by putting the looks to the target in relation to the looks to both target and distractor picture during baseline and testing phase. Third, the **duration of fixations** to the target and distractor picture were summed per condition and phase, delivering the looking time measure.

Additionally, the pointing reactions allowed for the separation of trials according to the correctness of the children's comprehension. Thus, looking behavior in correctly and incorrectly pointed trials is presented separately. This procedure is pursued by Trueswell et al. (1999) and Beyer and Hudson-Kam (2009), and is employed here to get a better grasp of what children are actually doing or looking at while they are exercising a pointing-while-looking task. For sake of simplicity, only a time course analysis based on the percentage of looks and a fixation duration analysis are performed for the correctly and incorrectly pointed trials.

10.3.2 Pointing reactions

The mean percentage of the pointing reactions to the target picture as a function of the number condition was calculated. In the four trials containing a singular inflected verb, the mean percentage of correct pointing reactions (to the 1-actor picture) was 59.8 %, (SD: 26.6 %) averaged over all 28 children. In the four plural trials, the mean percentage of correct pointing reactions (to the 2-actor picture) was 50.9 % (SD: 25 %), averaged over all children.[115]

[115] SDs are rather large due to the low number of test trials and because children varied between 0 % and 100 % correct pointing reactions.

In the singular condition, the pointing reactions were slightly above chance level, as the t-value shows a trend for significance ($t(27)=1.95$, $p=0.062$). In the plural number condition, the pointing seemed exactly to represent chance performance ($t(27)=0.189$, $p>0.1$). When all pointing reactions were collapsed over number condition, on 55.4 % of the trials children pointed to the correct picture (SD=14.9). Compared to chance level performance, again only a trend for significant difference emerged ($t(27)=1.89$, $p=0.069$).

Thus, when asked to point to the picture which best matched the auditorily presented sentence, children were mostly guessing and pointing randomly. This impression is further supported by the finding, that only two out of 28 children pointed significantly more correctly than would be expected by chance.[116] One out of 28 children pointed only to the 2-actor picture and another one only to the 1-actor picture, both irrespective of number condition.[117] For the other 25 children, pointing reactions were distributed randomly, supporting the assumption of a guessing-strategy employed by most of the children.

10.3.3 Time course

For a first visual inspection of the recorded eye gaze data, the looks to the 1-actor picture and the 2-actor picture over the course of a trial were averaged as a function of number con-

[116] To compare every single participants pointing scores to chance level, eight values were randomly generated per participant (with a mean of 0.5 and a standard deviation of 0.3) and were compared to the eight pointing scores (either 1 or 0) every participant had reached. Only one participant with 100 % (8/8) correctly pointed trials and one with 87.5 (7/8) correctly pointed trials performed better than chance when compared to the simulated data.

[117] Other studies have discarded children from analysis who only point to one side or one kind of target picture (e. g. Fernandez, Marcus, Nubila, and Voulomanos, 2005), but in this study, such behavior is interpreted as evidence that children do not understand the linguistic structure under investigation. Especially since all children pointed correctly in the pre-testing phase, see 10.2.4. But, excluding the two children who had pointed to only one of the picture types did not change group results such that the group mean was different from chance.

dition (see Figure 10.1). Depicted is the mean percentage of looks to both pictures at any two given time stamps (the looking behavior at two time stamps was averaged, because the whole course of a trial was divided into 40 ms bins), averaged over the four experimental trials of each number condition.

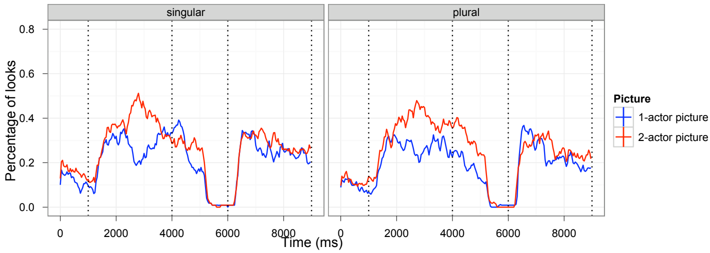


Figure 10.1: Time course of percentage of looks to 1-actor and 2-actor picture, separated by Number condition and Group (The dotted lines represent the onset and offset of the baseline and testing phase.).

These time course plots show that the children in this experiment looked more to the 2-actor picture in the baseline phase in both number conditions (as was found in the previous experiment). While the test sentence was presented, percentage of looks towards the critical spatial areas of interest dropped close to zero, which can be considered a consequence of the blank screen presented during those two seconds. In the testing phase, children did not seem to have a preference for any of the pictures, a pattern found in both number conditions. The eye-tracking data had to be further aggregated and statistical analyses had to be performed to test whether this impression would hold true.

10.3.4 Proportion-based analysis

As in Experiment 2, the proportion of looking time to the target picture (PLT) was calculated, setting the looks to the target picture in relation to the overall amount of looks to both pictures. The mean values plus standard deviation are listed in Table 10.1.

Phase	verb number	PLT (*Std.dev.*)	t-test vs. chance-performance
baseline	singular	0.44 (*0.109*)	$t(27)$=-2.916**
	plural	0.568 (*0.142*)	$t(27)$=2.523*
	collapsed	**0.504 (*0.141*)**	$t(27)$=0.237
test	singular	0.467 (*0.168*)	$t(27)$=-1.037
	plural	0.52 (*0.129*)	$t(27)$=0.834
	collapsed	**0.494 (*0.151*)**	$t(27)$=-0.308

Table 10.1: Proportion of looks to target (PLT) for each phase (signif. codes: *** <0.001, ** <0.01, * <0.05).

The mean PLT for each number condition and phase was compared to chance level performance 0.5 via t-test. The t-values and significance information is listed in Table 10.1 as well. The baseline pattern again confirmes the initial preference for the 2-actor picture. In the testing phase, the proportion of looks to target did not differ from chance level, indicating that the participants were equally likely to look towards the target as the distractor picture and therefore did not prefer for any of the pictures, irrespective of number condition.

Again, to level out the effect the initial preference for the 2-actor picture has in the two number conditions, the data was collapsed over subject number. When doing so, the baseline preference for the 2-actor picture vanished, but looking behavior in the test phase was at chance level as well and presumably not influenced by the verbal material children were presented with. The PLT measures indicated the same as the pure percentage of looks measure: while children in this experiment clearly looked more to the 2-actor picture in the baseline phase, no preference was found in the testing phase.

Further, PLT values were entered into a 2x2 ANOVA comparing Phase and Number condition. The main effect for Number ($F(1,27)=19.444$, $p<0.001$) simply confirmed higher PLTs in the plural compared to the singular trials. This reflects the preference for the 2-actor picture, which did not vary as a function of Phase, as indicated by the absence of a significant interaction. Therefore, the preference for the 2-actor picture did not get overridden by the verb inflection information.

10.3.5 Fixation duration analysis

Additionally, the looking times towards the 1- and 2-actor pictures were calculated and compared for the two number conditions and test phases.[118] The mean looking times for the singular and plural number condition during the baseline and the testing phase are plotted in Figure 10.2.

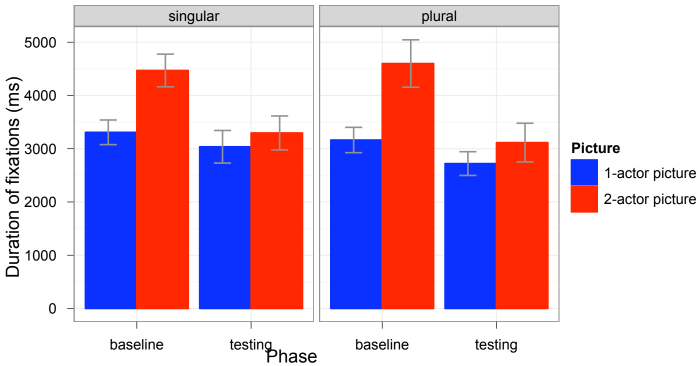


Figure 10.2: Mean duration of fixations to 1-actor and 2-actor picture, averaged over Phase and Number condition (Errorbars: +/− 1 SE).

1-actor picture The 2x2 ANOVA comparing Phase and Number for the looking times on the 1-actor picture did not reveal any significant effects or interactions. No further comparisons were performed. Thus, children's looking times to the 1-actor picture stayed constant, irrespective of test phase and number information on the verb.

[118] Two participants did not have any fixations in the test phase of the experiment and were thus not included in this analysis. This left the data of 26 participants in the analysis. One other participant did not fixate on the 1-actor picture during the testing phase of the singular trials, so a duration of 0 ms was inserted to include this participant in the statistical analysis. (R is not able to calculate ANOVAs with missing data.)

2-actor picture The 2x2 ANOVA for the looking times towards the 2-actor picture revealed that participants looked significantly more to the 2-actor picture during the baseline phase than during the test phase ($F(1,25)=15.107$, $p<0.001$). Even though no interaction was found, looking times in the two number conditions were analyzed separately. It was found that the duration of looks to the 2-actor picture significantly decreased from baseline to testing, but was found in the plural trials ($F(1,25)=10.068$, $p=0.004$) and in the singular trials ($F(1,25)=8.966$, $p=0.006$). Thus, no influence of number condition on the looking times could be found in the children. The analysis of the fixation durations confirmed the pattern found using the proportion of all looks, namely that the children prefer the 2-actor picture in the baseline phase, but that their looking behavior in the testing phase was not at all guided by comprehension of the test sentences, but was rather random.

10.3.6 Looking behavior divided by pointing reactions

Another way to analyze the looking data in this experiment is to divide the data sets into those trials in which the children pointed to the correct picture and those trials in which children pointed to the incorrect picture. Such an approach is taken by Trueswell et al. (1999), Hurewitz et al. (2000) and Beyer and Hudson-Kam (2009) to get a better impression of of what children were doing during the experiment and whether eye movements differed depending on the type of reaction. As 60 % of the trials containing a singular inflected verb and 53 % of the trials containing a plural inflected verb have been answered correctly, the amounts of trials in both analyses are not too different and should thus be comparable.

Here, only a subset of the analyses pursued above are presented. To get a first impression of the data, time course plots of the percentage of looks for the correctly and incorrectly pointed trials are depicted. Additionally, the mean looking time to the 1-actor and the 2-actor picture is plotted per phase and number condition, and entered into statistical analysis.

As has been presented in Chapter 10.3.2, the overall percentage of correct pointing reactions does not differ from chance level, therefore it could be argued that the children are just guessing when asked to point to the picture that best matches the auditorily presented sentence. To be able to discuss this possibility of pure guessing while pointing, and to discuss possible reasons why the looking-results of Experiment 2 (see Chapter 9) could not be replicated in the pointing-while-looking setting, it is interesting to investigate the looking behavior in relation to the reactions given by the children.

Correctly pointed trials Figure 10.3 shows the eye gaze that was recorded in those trials in which the children pointed correctly to the target picture. As can be seen, the initial preference for the 2-actor picture during the baseline phase is not very pronounced in the singular trials, while it surely is in the plural trials. Especially at the end of the baseline phase (around 4000 ms post onset) of the singular trials, we find that the children looked more to the 1-actor picture, although the singular inflected verb had not been mentioned yet.

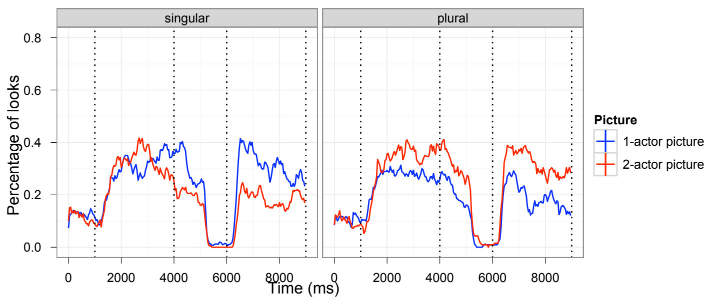


Figure 10.3: Time course of percentage of looks during **correctly pointed trials** to 1-actor and 2-actor picture, separated by number condition. (The dotted lines represent the onset and offset of the baseline and testing phase.)

The looking behavior in the test phase seems to be very similar to the eye gaze pattern found for the adults in Experiment 2 (see Figure 9.2 for comparison). In the test phase of the singular trials, they look more to the 1-actor picture, while in the plural trials, they look more to the 2-actor picture. This pattern can be interpreted as showing (a) that children comprehended the verb inflection, were able to use it to infer the number of the sentence subject and directed their eye gaze accordingly or (b) that children merely looked towards the picture which they were about to point towards anyways, without any influence of the verbal input. These possibilities will be discussed in Chapter 10.5.

To verify the impression of 'correct comprehension', the data of the correctly pointed trials was further analyzed. The looking time was determined and entered into statistical analysis. The mean looking times towards the 1-actor and 2-actor picture per condition and phase are depicted in Figure 10.4.[119]

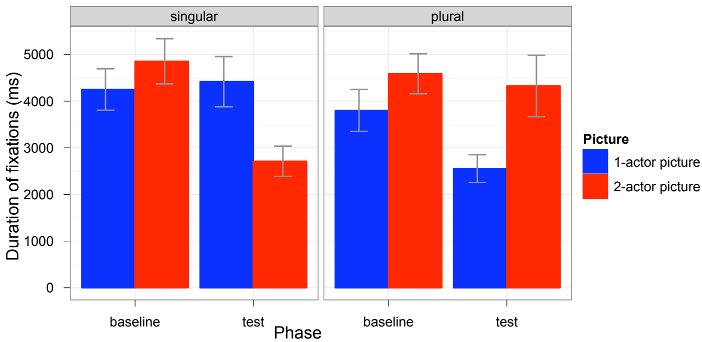


Figure 10.4: Mean duration of fixations on 1-actor and 2-actor picture of the **correctly pointed trials**, averaged over Phase and Number condition. (Errorbars: $+/-$ 1 SE)

[119]Note that only participants with data in all conditions and both phases could be included in this fixation based analysis. Because of missing data, 6 participants had to be discarded from analysis, leaving the data of 22 subjects.

1-actor picture To examine the looks to the 1-actor picture in the correctly pointed trials, a 2x2 ANOVA comparing Phase and Number condition as within-subject factors was performed. This yielded a significant main effect of Number ($F(1,21)=6.783$, $p=0.017$), indicating longer looks to the 1-actor picture in the singular trials. But a significant Phase X Number interaction was found ($F(1,21)=8.483$, $p=0.008$), which warranted further exploration. It turned out that no effect of Phase occurred in the singular trials ($F<1.0$), but that looking times significantly decreased from baseline to test phase in the plural trials ($F(1,21)=13.299$, $p=0.002$). Therefore, the duration of children's looks to the 1-actor picture stayed constant when children were presented with a singular inflected verb, but significantly decreased when they were presented with a plural inflected verb.

2-actor picture To examine the looks to the 2-actor picture, another 2x2 ANOVA comparing Phase and Number condition was performed. This revealed a significant main effect for Phase ($F(1,21)=12.958$, $p=0.002$), and again an interaction between the two factors ($F(1,21)=10.184$, $p=0.004$). Post-hoc one-way ANOVAs confirmed that looking durations to the 2-actor picture significantly decreased in the singular trials ($F(1,21)=26.803$, $p<0.001$), but stayed constant in the plural trials ($F<1.0$). Thus, the looking behavior towards the 2-actor picture was exactly opposite to the looking behavior towards the 1-actor picture in the correctly pointed trials. Looking times stayed constant when a plural verb was presented but significantly decreased when a singular verb was presented. It is interesting to note that differences in looking behavior in the correctly pointed trials are only marked by decreases. Adults in Experiment 2 displayed decreases and increases in looking times. This will be discussed in Chapter 10.5.

Incorrectly pointed trials Figure 10.5 depicts the time course of looking behavior averaged only over those trials in which the children pointed to the incorrect, i. e. non-machting, picture. In the singular trials, the children showed a clear preference for the 2-actor picture whenever the pictures were visible on the screen, thus in the baseline and in the test phase. In the plural trials, however, there seemed to be only a slightly higher percentage of looks to the 2-actor picture in the baseline phase, while in the test phase the children looked more to the 1-actor picture, which served as distractor in these trials. The looking behavior in the test phase of these incorrectly pointed trials seemed to be just the reverse of the looking behavior found in the correctly pointed trials.

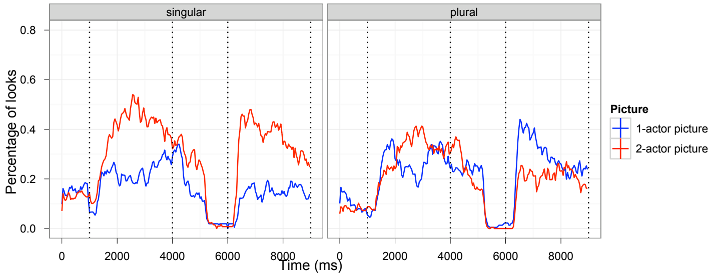


Figure 10.5: Time course of percentage of looks to 1-actor and 2-actor picture, separated by number condition. Only **incorrectly pointed trials** are included in this analysis. (The dotted lines represent the onset and offset of the baseline and testing phase.)

To validate the impressions from the time course plot, the looking times, i. e. summed fixation durations, were calculated again for each phase and condition and entered into a 2x2x2 ANOVA. The mean looking times towards the 1-actor and the 2-actor picture can be seen in Figure 10.6.[120]

[120]Note that only participants with data in all conditions and both phases could be included in this fixation based analysis. Because of missing data, 9 participants had to be discarded from analysis, leaving the data of 19 subjects.

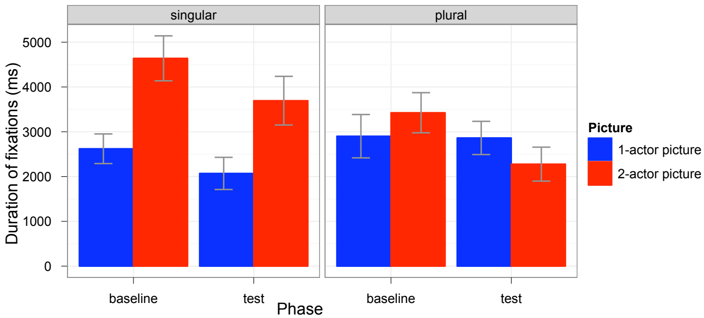


Figure 10.6: Mean duration of fixations to 1-actor and 2-actor picture of the **incorrectly pointed trials**, averaged over Phase and Number condition. (Errorbars: $+/-$ 1 SE)

1-actor picture The 2x2 ANOVA comparing Phase and Number condition in the looks to the 1-actor picture of the incorrectly pointed trials did not reveal any effects or interaction. No further statistical analyses were performed on the data. The duration of looks to the 1-actor picture was therefore constant in both number conditions in both test phases.

2-actor picture The 2x2 ANOVA comparing Phase and Number condition in the looks to the 2-actor picture instead revealed a main effect for Phase ($F(1,18)=8.816$, $p=0.008$) and a main effect for Number condition ($F(1,18)=5.675$, $p=0.028$). Even though the interaction failed to reach significance, further analysis showed that looking times to the 2-actor picture significantly decreased in the plural trials ($F(1,18)=21.128$, $p<0.001$), but not in the singular trials ($F(1,18)=2.252$, $p>0.1$). Clearly, the decrease in looking times to the 2-actor picture in the test phase is due to the fact that children pointed towards the non-matching, i. e. the 1-actor picture, in these trials. Such an effect of Phase was not found in the singular trials because in the baseline phase, children preferred the 2-actor picture anyways, and in the test phase, they pointed incorrectly to it. The implications of these findings will be discussed in Chapter 10.5.

10.4 Summary of Results

In Experiment 3, children were presented with the same material as in Experiment 2, but they had to point to the best matching picture while their eye gaze was tracked. The main aim of this experiment was to investigate the impact of task on children's comprehension of verb inflection. The analysis of children's pointing reactions showed that they did not differ from chance level performance in both number conditions.[121] This suggests that children were mostly guessing when asked to point to the matching picture.

[121] They did, however, show a trend in the singular condition.

This assumption is strengthened by the individual scores. Only one out of 28 children performed 100 % correct in both number conditions and only two performed better than chance.[122]

The analysis of children's looking behavior revealed a strong preference for 2-actor picture over the 1-actor picture during the baseline phase in both number conditions. This was found in the analysis of the proportions of looks to target and in the analysis of the looking times. Like in Experiment 2, we consider this a consequence of the greater amount of visual information encoded in the 2-actor picture. No influence of the verb inflection was found in the test phase of the experiment, independent of the picture examined and the analysis pursued (based on proportions or fixation durations). Thus no effect of the test sentence on children's looking behavior, as evident in Experiment 2, could be found here.

Since the only difference between the two experiments was the presence or absence of an additional picture selection task, we propose that this difference in tasks accounts in some way for the different findings across the two experiments. Possible underlying reasons for this task effect will be addressed in the following and in the General Discussion (Chapter 12).

Still, collection of off-line pointing reactions allowed to further analyze children's looking behavior according to the correctness of these reactions. This revealed further insights in the way children seemed to approach the picture selection task. The looking behavior in the correctly pointed trials indicated correct comprehension of verb inflection, as children's looks to the 2-actor picture decreased in the singular trials from baseline to test phase and conversely, the looks to the 1-actor picture decreased in the plural trials from baseline to test phase. The pattern is still different from the one found for the adult control participants in Experiment 1. While the adults displayed increases in their looking times from baseline to test phase in relation to the verb inflection information, children's 'compre-

[122] Remind that children only encountered 8 trials, four per number condition. The child who performed 100 % correct pointed to the target picture in 8/8 trials, while the other two children pointed to the target in 7/8 trials.

hension' was only apparent in decreases. Even more revealing are the graphs depicting the time course of looking behavior. At the end of the baseline phase, children looked more to the 1-actor picture in the singular trials and to the 2-actor picture in the plural trials.[123] Thus, one might suspect that children's early preference of the 1-actor picture during those singular trials is rather guided by chance than by comprehension. More precisely, the looking pattern in the correctly pointed trials suggests that children might already have made up their mind of which picture is more interesting and they would point to. Since the test sentence had not been presented yet, a preference to the 1-actor picture must have been driven by some other unknown factor, but not by the verbal stimulus. In the baseline phase of the plural trials, the usual preference for the 2-actor picture was apparent. Whether this pattern is just an instance of the usually found 2-actor preference or as well a possible early decision for the 2-actor picture is nondistinctive.

To conclude, the picture selection task does not provide evidence that children can make use of the number information provided by the verb inflection. These findings are in line with findings from earlier studies with English- and Spanish-learning children that were using the same kind of picture selection task (Johnson et al., 2005; Pérez-Leroux, 2005). This strengthens the assumption that it is not structural differences across the languages considered that account for the earlier competence found in German children (Experiment 2), but rather the different methods used. In the following, I will discuss how task dependent comprehension can best be explained in terms of current language acquisition and language processing models, and what this within-modality asymmetry means for the postulated asymmetries between comprehension and production with regard to verb inflection.

[123] The incorrectly pointed trials point toward a similar direction. A strong preference for the 2-actor picture is found in the baseline phase of the singular trials, which did not get overridden by the verbal material. No preference for any picture was found in the baseline phase of the plural trials, which hints that the general preference for the 2-actor picture was already overridden, and no further differences were found in the test phase.

10.5 Discussion

The present experiment aimed to examine German-speaking children's use of verb inflection as a cue to subject number. To be more comparable to experiments testing English- and Spanish-speaking children and to investigate the impact of task, children were asked to perform a picture selection task while their eye gaze was tracked. In all other respects, this experiment was a replication of Experiment 2. While the findings of Experiment 2 provide evidence that German three to four year old children are able to infer the number of the sentential subject solely by relying on the inflectional information, such evidence could not be found in Experiment 3. Neither the the looking data nor the pointing data indicated that children matched sentences containing a singular inflected verb to the 1-actor picture and sentences containing a plural inflected verb to the 2-actor picture. As different groups of children were tested in Experiment 2 and 3, it cannot be ruled out that this difference is due to group differences between the children. But this possibility is considered unlikely, as the groups differed neither in age nor with respect to the productive use of verb inflections (p>0.5, t-test). Because of the different result patterns found across the two experiments, we assume that some kind of processing factors must account for children's failure to demonstrate the morphosyntactic knowledge displayed in Experiment 3.

The results obtained in Experiment 3 are in line with the lack of comprehension found in English- and Spanish-speaking children who have as well been tested using a picture selection task (Johnson et al., 2005; Pérez-Leroux, 2005). Johnson and colleagues propose that agreement morpheme checking is a purely syntactic operation and that agreement features carried by verb inflection affixes do not survive to LF, as suggested by Chomsky (1995) and therefore cannot be used for semantic analysis of the sentence (Johnson et al., 2005). This assumption surprisingly does not take into account that a verb enters into an agreement relation with the subject of a sentence, such that the number features of the verb are 'coordinated' with the subject's features. These subject features critically survive to LF, so that interpretation of the verb morphology should be possible nonetheless. Otherwise it would be unclear how adults and older children could infer the semantic number of the subject based of verb inflection information. Alternatively, Johnson and colleagues discuss the possibility that children use a rule that relies on co-occurrence patterns of subject NPs, pronouns and certain verb forms, and does not involve the number contrast (see 2.6). They immediately reject this idea, under the assumption that "[...], 'NP without /s/' is hardly the kind of rule that languages, or children, represent" (Johnson et al., 2005, pg. 327). I nevertheless claim that the existence of such a rule that relies on co-occurence patterns is very likely, especially in younger children. It can even be assumed that the two rules that are postulated in 2.5 and 2.6 (and repeated in 6.5 and 6.6) co-exist in the child's grammatical system at a certain point in development, but that children are only able to 'work with' the more abstract and later learned rule when further processing demands are kept as low as possible. Otherwise children might retreat to the more basic rule that they are more familiar with or that is easier for them to use, since they have 'known' it for much longer (assuming that children use a non adult-like co-occurence rule for early productions and/or early detection of subject verb agreement).

Before further discussing possible reasons for the result pattern obtained and put the comprehension data in relation to earlier reported production data, let's consider how the findings of the correctly and incorrectly pointed trials can be interpreted. It is assumed that the picture selection task in Experiment 3 interfered with children's ability to interpret the test sentences correctly. This interference is caused the explicit decisions the children were forced to make in the picture selection task. The looking patterns found in the correctly pointed trials can be interpreted as providing indirect support for such an 'interference effect'. It seems to be the case that children pointed to that picture which they have looked at at the end of the baseline phase of a given trial. This suggests that children, who of course knew they had to come up with a pointing reaction at the end of every trial, decided for one picture, even before the verbal stimulus was presented. It might be the case that children felt pressure to come up with an answer, and this explicit decision which had to be made interfered with their ability to make use of verb inflection information in sentence processing (to access the adult-like rule, so to speak). Thus the pattern might reflect a 'strategy' used by the children to simply stick to the picture to which they have looked at the end of the baseline phase, and to finally point to exactly that one. This would support the notion of 'pure guessing when pointing'.

Trueswell et al. (1999) and Hurewitz et al. (2000) as well found a tendency in children to stick to a visual referent, regardless of linguistic information presented. The authors report that children's eye fixations showed a pattern of 'egocentricity' when they had to interpret sentences with temporarily ambiguous structures (e. g. *Put the frog on the napkin in the box*) in the presence of two possible referents (one frog on a napkin and a second frog not on a napkin). The basic findings was that children were not able to take the referential information into account during sentence processing and interpret the PP *on the napkin* as a restrictive modifier, even when two frogs were present. It was found that children mostly returned to the one frog they had looked at when the word 'frog' was presented, and used this one frog (whether on a napkin or not) for the putting action (Trueswell et al., 1999).[124] See as well Trueswell and Gleitman (2007) on this aspect. Beyer and Hudson-Kam (2009) have also found in the eye gaze of the incorrectly pointed trials that children seem to stick to the picture they had looked at at the beginning of a trials.

An alternative explanation for the looking pattern found in the correctly pointed trials would be that the picture, which children looked at, actually influenced their sentence comprehension. In this scenario, children might have expected to hear a sentence that would describe what they were just looking at (this seems to represent very 'natural' situation for children,

[124] Interestingly, such a pattern, to stick to the one frog they had looked at before, was not found when the modifier reading was enforced by the use of a complementizer (*..the frog that's on the napkin...*). Thus children were able to revise their initial reference decision when they easily comprehended the linguistic input.

because that's what happens when adult read picture books to them). Thus, they could have 'parsed the sentence correctly', simply because the sentences in these cases matched what they had expected. In those cases when the sentence did not match what children had expected, they were unable to revise their initial 'parsing decisions'. There is experimental evidence that children have difficulties revising their initial sentence parse (e. g. Kidd, Stewart, and Serratrice, 2010; Trueswell et al., 1999; Trueswell and Gleitman, 2007). Such an idea might be far fetched, and the role of the visual referent scene in sentence processing, especially in children, is still a matter of ongoing discussion. To date, experimental results indicate that children are not able to use the referential scene information to guide their parsing (e. g. Trueswell et al., 1999; Snedeker and Trueswell, 2004; Hurewitz et al., 2000). Studies investigating noun and verb learning on the other hand repeatedly report that children use referential information to detect the meaning for newly presented word forms (e. g. Waxman and Lidz, 2006; Golinkoff and Hirsh-Pasek, 2008; Arunachalam and Waxman, 2011; Fisher, 2002). Interestingly, the ability to use referential information to detect word mearning seems to be strongly constrained by the complexity of the linguistic and non-linguistic information (e. g. Scott and Fisher, 2012).

Overall, our findings raise some important questions. First, what could the possible underlying reasons for the *within-modality asymmetry* be, i. e. what underlies children's success in demonstrating their comprehension ability on verb inflections when tested via preferential looking versus their failure to do so when tested with a picture selection task? Second, what implications does this within-modality asymmetry have for *asymmetries across modalities*, namely production and comprehension? I will provide some answers to these questions in the General Discussion (Chapter 12). To validate our findings that German-speaking children actually fail to understand verb inflection in a (more demanding) picture selection task, we conducted a follow-up experiment. Because we reasoned that children might have difficulties matching the pictures to a sentence that was presented while the screen was blank, we altered the trial procedure. Additionally, we wanted to know whether children assigned a singular or plural interpretation to the pronoun *sie*, which was not detectable when the test sentence was presented in the absence of visual material.

11 Experiment 4: Online comprehension of inflection

The results from Experiment 2 and 3 indicate that children's ability to use verb inflections in sentence comprehension is strongly affected by the way children's comprehension abilities are examined. While children were able to link a picture showing one actor to verbs inflected for 3^{rd} person singular (and a picture showing two actors to verbs inflected for 3^{rd} person plural) when tested in a pure eye tracking paradigm, no such ability was found when children had to perform an explicit picture selection task. To examine whether a simpler trial structure would improve children's comprehension abilities when asked to point to a matching picture and to further investigate whether children assigned a number interpretation to the ambiguous pronoun sie, this last experiment was conducted. Adult control participants were tested as well to provide a frame of reference for 'adult-like' eye gaze patterns.

11.1 Rationale

To validate the influence of task demands on children's comprehension of verb inflection, a fourth experiment was conducted. In this experiment, the eye-tracking technique was again combined with an explicit picture selection task. Children had to decide and point to the matching picture again while their eye gaze was tracked. The critical difference to Experiment 3 was that, in the present experiment, the pictures were visible over the whole course of a trial and the test sentences were presented *while* the pictures were visible. This allowed us (a) to verify the findings of Experiment 3 with another group of three year old children, (b) to examine whether the disappearance and reappearance of the pictures had an effect on children's pointing reactions as found in Experiment 3, and c. to examine participants interpretation of the subject pronoun. Further, an adult control group was tested in a picture selection task on their comprehension of verb inflection.

11.2 Method

11.2.1 Participants

In this version of the comprehension experiment, 30 German speaking children aged 3 years were tested. The man age was 3;6, ranging from 3;0 to 3;11 years. Sixteen of the participants were girls. All children came from monolingual german homes in the Berlin/Potsdam area, had no history of speech or language disorder and were not born prematurely (this was ensured using the same parental checklist as used for Experiment 2 (see Chapter 9.2.1). The parents gave informed content and were reimbursed for their travel costs to the lab, children received a small toy for taking part in the study. Additionally, 20 adults were tested in this experiment. All adults were students or employees of the University of Potsdam, with a mean age of 24 years, ranging from 20 to 42 years. Fourteen of the adult participants were women. The students received class credit and the employees a small financial reimbursement for taking part in the study.

11.2.2 Material

The same SVO-sentences as in Experiments 2 and 3 were used. The verbal stimuli that had been recorded for Experiment 2 were used here as well, but had to be adjusted to match the structure of the on-line trial procedure (see Chapter 11.2.3). More precisely, adjustments were made to keep the onset of the pronoun and the offset of the verb inflection at the same time point in all trials. The pronoun onset was at exactly 2000 ms post trial onset and the offset of the verb inflection was at 2900 ms post trial onset. Therefore, in some trials, the pause between the pronoun *sie* and the verb had to be lengthened or shortened to account for these temporal restrictions. The mean length of time between pronoun onset and verb offset (i. e. verb inflection) before manipulation was 765 ms (min: 659 ms, max: 934 ms), did not differ significantly between number conditions (singular inflected verbs: 761 ms; plural inflected verbs: 769 ms). After manipulation, the mean length of time between pronoun onset and verb inflection offset was 900 ms (min: 855 ms; max: 908 ms). Additionally, the same visual material as in Experiments 2 and 3 was used.

11.2.3 Design

The trials were pseudo-randomization and counter-balanced across children as in Experiment 2. Every participant was presented with four practice and eight pseudo-randomized testing trials.

Importantly, the course of each single trial was different from the one in Experiment 2 and Experiment 3. In those two experiments the pictures appeared twice on the screen and the test sentences were presented while the screen was black. In the present experiment however, the pictures of a pair were visible over the whole course of a trial, that is until the participants had pointed at one of the pictures and the trial was terminated by the experimenter. Thus, the test sentence was presented *while* the pictures were visible. See Figure 11.1 for a schematic description of of a trial course. The alignment of visual and verbal material was made to ensure that the looks in all trials would be time-locked to the pronoun onset and the critical disambiguating information, i. e. the verb inflection. Additionally, the duration of the baseline and test phase were altered such that they lasted for 2s instead of 3s.

These changes allowed us to 1) examine the looking behavior at the verbal presentation of the pronoun *sie* to examine whether children or adults would interpret the pronoun *sie* as rather referring to one (female) person or more than one person, and 2) to examine whether the structure of a trial has an influence on the comprehension of verb inflection as found in Experiment 2. We wanted to tested whether a constant presen-

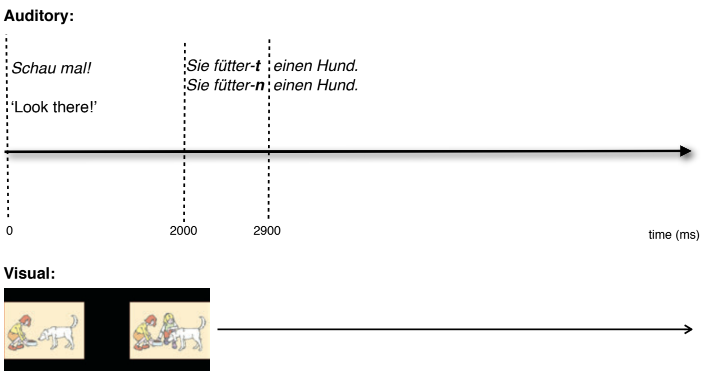


Figure 11.1: Schematic time course of a trial in Experiment 4.

tation of the visual stimuli would enhance the children's comprehension performance compared to the previous pointing experiment (where comprehension of verb inflection was found neither in the pointing reactions nor in the looking behavior). Lastly, we wanted to test adults on their comprehension of verb inflection when asked to point to a matching vs. non-matching picture.

11.2.4 Procedure

The same technical equipment as in Experiment 2 and 3 was used (Tobii 1750 tabletop eye-tracking system, 50 Hz). The testing procedure when coming to the lab was the same as in Experiment 3, since children and adults were instructed to point to the matching picture. Pre-testing (point to one or two boys presented on two different pictures) was only performed with children.

11.2.5 Data Analysis

The eye gaze data was analyzed in same way as in Experiment 2 and 3, except that the temporal regions of interest (Phases) in this experiment were defined differently due to the different trial procedure. The 2000 ms time span before the test sentence was presented was defined as *baseline phase*, in which it was possible to check for a visual preference for one of the pictures, that was not related to any verbal input. The 2000 ms after the offset of the verb inflection was used as *test phase*. This test phase would actually need to start at 2900 ms post trial onset, because this is when the verb inflectional information had been presented, but to account for time needed to analyze the incoming auditory information and to initiate and launch an eye movement, 300 ms were added before the actual test phase started.

Pointing reactions were noted on protocol sheets as pointings to the left or right picture and later coded for being on target or distractor picture. The mean percentages of correct reactions per condition and group were calculated and statistically compared to chance-level performance (0.5).

11.3 Results

Looking behavior was, as in Experiment 2 and 3, analyzed in the following ways. First, the percentage of looks was used to create a **time course plot** showing the looking behavior over the course of a trial. This only served for a rough visual inspection of the data. Then, the **proportion of looks to target** (PLT) was calculated for each group, trial phase and number condition. In this procedure, looks to target are set in relation to the looks to target and distractor. Further, **looking times** to the 1-actor and 2-actor picture were calculated. The PLT-values and the looking times were entered into separate ANOVAs for statistical analysis. Since the baseline and the test phase in this experiment lasted only 2 sec instead of 3 sec, the total looking times might well be lower than the ones found in the previous experiments.

Additionally, a **pronoun analysis** targeting the looking behavior following the pronoun *sie* was performed (2300 to 2800 ms post onset). Because no number information was presented at that point in the trial, all trials were collapsed and the percentage of looks to the 1-actor picture and the 2-actor picture was calculated and compared separately for each group. This was done to determine whether participants interpreted *sie* as rather referring to a singular subject or a plural subject. One open question in the previous experiments testing the comprehension of verb inflection (Experiment 2 and 3) was whether the participants had assigned a particular interpretation to the pronoun *sie*. To the best of my knowledge, there is no study addressing the question whether the German pronoun *sie* is interpreted by default as referring to a singular or to a plural entity in adults or children.

Finally, as in Experiment 3, trials were analyzed separately according to the **pointing reactions** of the children. This analysis was again not feasible for the adult-data because of the very few incorrect pointing reactions obtained from the adults (n=3).

11.3.1 Pointing reactions

The mean percentages of correct pointing reactions to the target picture as a function of number condition was calculated separately for the two age groups. These mean percentages are listed in Table 11.1, alongside the information on the statistical comparison to chance level 0.5. For each number condition, four trials were analyzed.

Group	singular condition	plural condition
three year olds	50 %	50 %
	$t(29)$=-0.216	$t(29)$=-0.202
Adults	96 %	100 %
	$t(19)$=8.283***	$t(19)$=9.178***

Table 11.1: Mean percentage of correct pointing reactions (signif. codes: *** <0.001, ** <0.01, * <0.05).

As can be seen in Table 11.1, three year olds point to the target picture exactly half of the times, performing perfectly at chance-level. Therefore, might reasonably conclude that the three year old children in this experiment are simply guessing when asked to point to the matching picture. This behavior is found in the singular and in the plural trials.

The adult participants on the other hand are close to ceiling in the singular condition and at ceiling, i. e. at 100 % correct, in the plural condition. Not surprisingly, they pointed at significantly better than chance-level, indicating correct comprehension of the verb inflection and no problems with a picture-selection task. The pointing reactions gathered in the current experiment confirmed the findings from Experiment 3 that three to four year old German children were not able to show comprehension of verb inflection morphemes differing in number when tested with a simple picture selection task.

11.3.2 Time course

For the first visual inspection of the participants' eye gaze pattern, looks to the 1-actor and to the 2-actor picture were plotted separately as a function of number condition and group over the whole course of a trial (see Figure 11.2). As before, the eye gaze pattern at each recorded time-stamp is averaged over trials and participants (and only eye gaze data that were recorded as being 'on screen' were considered for analysis).

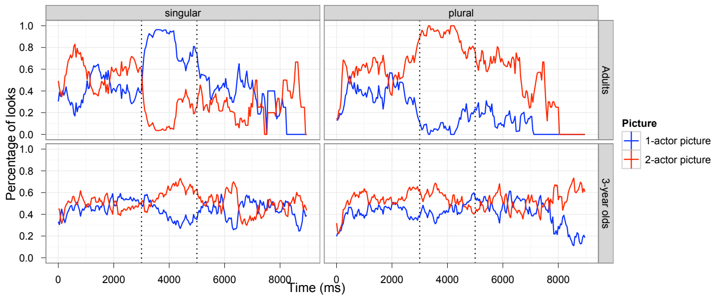


Figure 11.2: Time course of percentage of looks to 1-actor and 2-actor picture, separated by Number condition and Group. (The dotted lines represent the onset and offset of the test phase.)

The graphs in Figure 11.2 show that adult participants in the singular trials looked more towards the 1-actor picture after roughly 3000 ms post trial onset. In the plural trials, adult participants looked more to the 2-actor picture. This difference emerged earlier, although the preference for the 2-actor picture got enhanced after roughly 3000 ms post onset. Further, a slight preference for the 2-actor picture is visible at the very beginning of the trial and in the pronoun phase. The three year olds show a less clear looking pattern. Actually, no striking preference for one of the pictures is visible in any number condition. To test whether the amount of looks to the 1-actor and 2-actor picture differed as a function of number condition, the eye gaze data were further aggregated and entered into statistical analysis.

11.3.3 Proportion-based analysis

The mean PLT-values that were calculated for for each group, phase and number condition are listed in Table 11.2. The comparison to chance via t-test and the level of significance are reported as well.

Group	Phase	verb number	PLT (*Std.dev.*)	t-test vs. chance
3-year (N=30)	baseline	singular	0.477 (*0.179*)	*t*(29)=-0.752
		plural	0.542 (*0.116*)	*t*(29)=2.689*
		collapsed	**0.51 (*0.153*)**	*t*(29)=0.951
	test	singular	0.426 (*0.132*)	*t*(29)=-3.274**
		plural	0.535 (*0.191*)	*t*(29)=-0.319
		collapsed	**0.48 (*0.172*)**	*t*(29)=-2.623*
Adults (N=30)	baseline	singular	0.464 (*0.092*)	*t*(19)=-0.978
		plural	0.561 (*0.102*)	*t*(19)=0.642
		collapsed	**0.512 (*0.108*)**	*t*(19)=0.056
	test	singular	0.703 (*0.178*)	*t*(19)=3.725**
		plural	0.742 (*0.179*)	*t*(19)=5.664***
		collapsed	**0.722 (*0.177*)**	*t*(19)=6.517***

Table 11.2: Proportion of looks to target (PLT) for each group and phase (signif. codes: *** <0.001, ** <0.01, * <0.05).

Adults In the baseline phase, adults' PLT-values did not differ from chance level in the singular condition or in the plural condition. Thus, no significant visual preference was found in the baseline phase in this experiment. In the test phase however, PLTs exceeded chance level significantly. This was again found for the singular trials, the plural trials and when trials were collapsed over number condition. Therefore one can conclude that adults' eye gaze in the test phase was influenced and directed by the information conveyed by the verb inflection. When the verb was inflected for singular, adults preferred to look to the 1-actor picture and when the verb was inflected for plural, adults preferred to look to the 2-actor picture.

The PLT-values calculated for the adults were further entered into a 2x2 ANOVA comparing Phase and Number condition. This confirmed that adults had higher PLT values in the test phase than in the baseline phase ($F(1,19)=23.288$, $p<0.001$). Further, the main effect for Number condition ($F(1,19)=8.377$, $p=0.009$) indicated overall more looks to the 2-actor than the 1-actor picture, a pattern that has been repeatedly found in the previous experiments. No interaction between the factors was found. Overall, adults look more to the target picture in the test phase than in the baseline phase. This is found for both number conditions.

Children The children displayed a different looking pattern than the adults. In the baseline phase of the plural condition, they looked more to the 2-actor picture than would be expected by chance. This again reflects a strong preference for the 2-actor picture, as found in the previous experiments. In the test phase, though, PLT differed from chance in the singular condition, this difference was because the PLT was still *lower* than expected by chance. Thus, in the test phase of the singular trials, the three year olds still preferred to look towards the 2-actor picture even though the verb inflection should have pointed them towards the other, the 1-actor picture. The PLT for the plural trials did not differ from chance in the test phase. Neither did the PLT value calculated from all trials, collapsed over number condition.

The 2x2 ANOVA, that was performed to further investigate the proportion of looks to target, only revealed a main effect for Number condition ($F(1,29)=8.586$, $p=0.007$). This confirmed to overall higher PLT for the plural trials compared to the singular trials. This is most likely due to the fact that the 2-actor picture served as target in the plural trials and was preferred simply for visual reasons.

11.3.4 Fixation duration analysis

In Experiment 2, looking times revealed the most compelling evidence for correct comprehension in the group of three year olds. To test whether participants looked *longer* to the matching than the non-matching picture in the test phase of the current experiment, the summed fixation durations per trial were used as dependent measure. These looking times are depicted in Figure 11.3.[125]

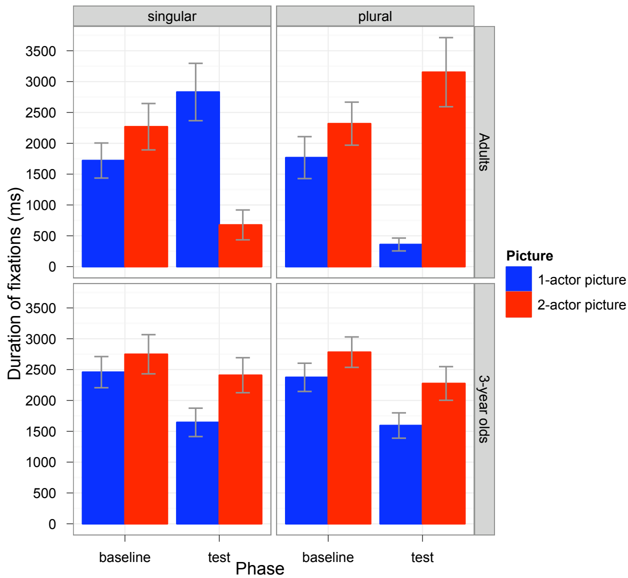


Figure 11.3: Mean duration of fixations on 1-actor and 2-actor picture, separated by Number condition and Group (Errorbars: +/- 1SE).

[125]Note that only those participants that delivered fixation data in all conditions could be included in the analysis. Six adults and four children had to be excluded from analysis, leaving overall 40 participants (i. e. the data of 14 adults and 26 children).

Adults

1-actor picture To examine whether the looking times towards the 1-actor picture changed in relation to trial phase and verb number, a 2x2 ANOVA was performed for the adult data. The main effect of Number condition was highly significant ($F(1,13)=24.648$, $p<0.001$) as was the Phase X Number interaction ($F(1,13)=20.133$, $p<0.001$). This interaction was further investigated via post-hoc ANOVAs, one for each number condition. In the singular trials, a significant increase in looking times was found from baseline to test phase ($F(1,13)=12.98$, $p=0.003$) and in the plural trials, a significant decrease in the duration of looks to the 1-actor picture was found ($F(1,13)=14.704$, $p=0.002$). Therefore, rather similar looking times to the 1-actor picture were found in both number conditions in the baseline phase, but these were found to decrease when a plural inflected verb was presented and they were found to increase when a singular inflected verb was presented. Looking times to the 1-actor picture therefore confirm that adults were able to use verb inflection to interpret the number of the sentence subject.

2-actor picture To examine the looking times towards the 2-actor picture, another 2x2 ANOVA comparing Phase and Number condition was performed. This time, a highly significant main effect for Number condition ($F(1,13)=20.05$, $p<0.001$) and a marginal effect for Phase were found ($F(1,13)=3.417$, $p=0.087$). But the interaction between the factors proved to be significant as well ($F(1,13)=15.529$, $p=0.001$), warranting further investigation. The post-hoc ANOVA comparing Phase in the singular trials revealed a significant decrease in the duration of looks towards the 2-actor picture from baseline to test phase ($F(1,13)=16.248$, $p=0.001$). In the plural trials, on the other hand, a significant increase in looks to the 2-actor picture from baseline to test emerged ($F(1,13)=5.923$, $p=0.003$). Thus, verb inflection information affected adults' looks to the 2-actor picture as well, providing grounds for an increase in the plural trials from baseline to test phase, and for decrease in the singular trials.

Children

1-actor picture To examine children's looking times towards the 1-actor picture and to test whether these changed as a factor of number condition, another 2x2 ANOVA comparing Phase and Number condition was performed. This revealed a main effect for Phase ($F(1,25)=12{,}457$, $p=0.001$), without any further interaction. Therefore, the looking times towards the 1-actor picture simply decreased from baseline to test, to the same extend in both number conditions. This pattern was confirmed by two post-hoc one-way ANOVAs comparing the factor Phase in both number condition. A significant decrease was found in the singular trials ($F(1,25)=6.758$, $p=0.015$) and in the plural trials ($F(1,25)=8.126$, $p=0.009$). Thus, children's looking behavior towards the 1-actor picture was not found to be influenced by verb inflection information.

2-actor picture Another 2x2 ANOVA comparing Phase and Number was performed to examine children's looking times towards the 2-actor picture. This revealed simply a marginal effect for Phase ($F(1,25)=3.771$, $p=0.064$) without any interaction. Nevertheless, again post-ho ANOVAs comparing the factor Phase in both number conditions were pursued. These revealed no difference in looking times from baseline to test for the singular trials ($F<1.0$). For the plural trials, a decrease in looking times towards the 2-actor picture emerged, which almost proved to be significant ($F(1,25)=4.014$, $p=0.056$). This pattern is actually the opposite of what would be expected if children's eye gaze was directed by verb inflection information.

The findings mirror the children's results obtained in the PLT analysis, revealing no difference in looks to the 1-actor and 2-actor picture as a function of number condition. Thus, there was no evidence found thus far that children comprehend and use the verb inflectional information as a cue to subject number when tested in a pointing-while-looking experiment with a online trial procedure.

11.3.5 Pronoun analysis

Additionally, the looking behavior directly after the pronoun towards the 1-actor and the 2-actor pictures was examined. Because the test sentences were the same up to the disambiguating verb inflection, all test sentences were analyzed together (collapsed over number condition). To this end, the percentage of looks towards both pictures in the 500 ms following the pronoun presentation was analyzed (with an additional 300 ms time to analyze the linguistic input and to plan eye movements, thus from 2300 to 2800 ms post trial onset). Critically, only a 500 ms window could be used for the pronoun analysis, because at 2900 ms post onset, the verb inflectional information is presented. Additionally, the first look following the presentation of the pronoun was singled out and checked whether it landed on the 1-actor or the 2-actor picture.

Percentage of looks following *sie* The mean percentage of looks towards the 1-actor and 2-actor picture in the 500 ms phase following the pronoun were calculated and are listed in Table 11.3. The percentage of looks toward the 1-actor and 2-actor picture were compared via paired t-tests, one for each age group.

Group	1-actor picture	2-actor picture	t-test (pictures)
Adults	18 % (*16*)	28 % (*21*)	$t(19)$=-3.186**
3-year olds	29 % (*21*)	29 % (*21*)	$t(29)$=-0.146

Table 11.3: Mean percentage of looks (*Std.dev.*) following *sie* (signif. codes: *** <0.001, ** <0.01, * <0.05).

The adults preferred to look towards the 2-actor picture after the presentation of the pronoun. This could either mean that they interpreted the pronoun as referring to a plural subject, or that they did not interpret the pronoun itself at all and that the eye gaze reflected a visual preference for the 2-actor picture. This shall be discussed in Chapter 11.5.

The children on the other hand did not show a preference for one of the pictures in the 500 ms following pronoun presentation. Considering that all groups of children in Experiment 2 and 3 experiments displayed a clear baseline preference for the 2-actor picture, the finding of no preference following the pronoun *sie* seems remarkable. It could indicate that the pronoun presentation was overriding the initial preference for the 2-actor picture, which would only be possible if the pronoun would be interpreted as referring to a singular subject.[126] This idea is discussed in Chapter 11.5.

First look following *sie* The percentage of first fixations to the 1-actor and the 2-actor picture after encountering the pronoun *sie* were calculated for each group. The means and standard deviations are listed in Table 11.4.[127]

Group	1-actor picture	2-actor picture	t-test (pictures)
Adults	42 % (*29*)	57 % (*29*)	$t(17)$=-1.027
three year olds	50 % (*24*)	49 % (*24*)	$t(26)$=0.151

Table 11.4: Mean percentage of first fixations towards the 1-actor and the 2-actor picture following *sie* (*Std.dev.*)
(signif. codes: *** <0.001, ** <0.01, * <0.05).

It was found that a higher percentage of adults' first fixation after the pronoun landed on the 2-actor picture, but the difference failed to reach significance (F<1.0). For thet children, almost exactly half of the first fixations landed on the 1-actor and half landed on the 2-actor picture. This pattern revealed thus no preference and is therefore in line with the percentage of looks following the pronoun *sie*.

[126] It has to be noted though that in the present experiment a 2-actor picture preference was only found in the PLT analysis in the plural trials.

[127] Note that only participants that actually fixated the 2-pictures in the 500 ms time phase could be included in the analysis. Due to empty data cells, two adults and 3 children had to be excluded, leaving 45 participants in the analysis.

11.3.6 Looking behavior divided by pointing reactions

As in Experiment 3 (see Chapter 10) we were able to divide the trials according to the correctness of children's pointing reactions. This was done to get a better impression of children's online sentence interpretation and how they ended up with the pointing reaction as they did.

Correctly pointed trials Figure 11.4 shows the averaged percentage of looks in those trials in which children finally pointed to the picture that matched the test sentence.[128] It can be seen that the 2-actor picture was slightly preferred before the test phase started at around 3000 ms. In the singular trials, a peak in looks to the 1-actor picture, the target, appeared around 7000 ms post onset. In the plural trials, a strong preference for the 2-actor picture can be seen, although this only began when the test phase ended, at roughly 5000 ms into the trial.

Although the time course graphs do not indicate a preference for the target picture in the test phase, as it was defined for this experiment, the looking times to both pictures were calculated and entered into statistical analysis to confirm this impression. Again, separate ANOVAs were conducted for two pictures. Looking times were calculated in the usual way, described in detail above (see page202). The looking times for the correctly pointed trials are depicted in Figure 11.5.

1-actor picture The 2x2 ANOVA comparing Phase and Number condition in the looks that were recorded during the correctly pointed trials revealed only a main effect of Phase ($F(1,17)=10.754$, $p=0.004$). Since no interaction was found, the main effect indicates that the looks to the 1-actor picture decreased from baseline to test phase in both number conditions. No further analyses were performed. Thus, looking times to the 1-actor picture did not indicate any influence of verb inflection,

[128]Note that only 18 children were included in the analysis of the correctly pointed trials, due to empty data cells in the fixation analysis. If children did not fixate both pictures in the singular and plural trials in the baseline and test phase, they had to be discarded from analysis.

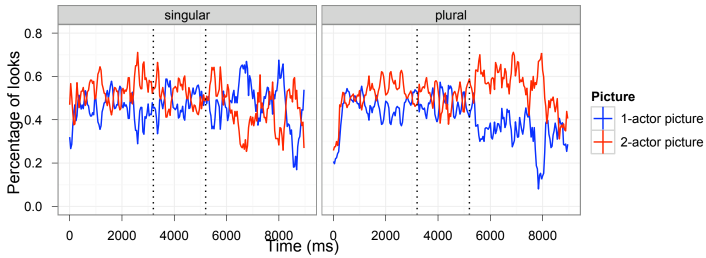


Figure 11.4: Time course of percentage of looks to 1-actor and 2-actor picture in the **correctly pointed trials** only, separated by Number condition. (The dotted lines represent the onset and offset of the test phase.)

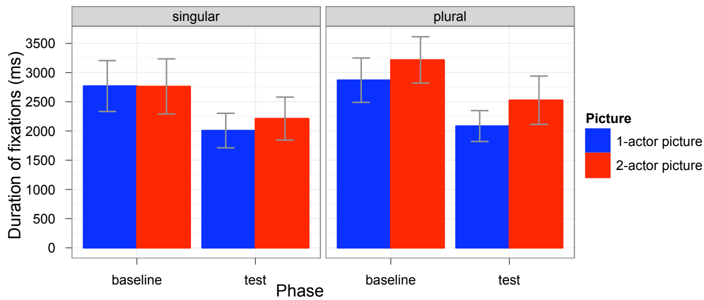


Figure 11.5: Mean fixation durations to 1-actor and 2-actor picture for the **correctly pointed trials** only, separated by number condition.(Errorbars: +/- 1SE)

even when only the trials were analyzed in which the children pointed correctly. Looks to the 1-actor picture even decreased from the baseline to the testing phase, no matter whether a singular inflected verb or a plural inflected verb was presented.

2-actor picture The compare the looking times to the 2-actor picture, another 2x2 ANOVA with the factors Phase and Number condition was performed. This did not reveal any main effects or interactions. Thus, as for the looks to the 1-actor picture, no influence of verb inflection was found, even though only the correctly pointed trials were considered for analysis. Children's amount of looks to the 2-actor picture stayed constant over the course of a trial, and did not decrease like the looks to the 1-actor picture, possibly due to visual complexitiy of the 2-actor picture.

The time course graph shows that children's preference for the target picture emerged only later during the trial. But since this preference began to emerge more than 2000 ms after the critical verb inflection information had been presented, it is hard to argue that any preference found is actually related to the verbal stimulus. This will be further discussed in Chapter 11.5.

Incorrectly pointed trials Figure 11.6 depicts children's percentage of looks to the 1-actor and the 2-actor picture in those trials in which they pointed to the incorrect picture.[129] At the very beginning of the trial, a slight preference for the 2-actor picture can be seen in both number conditions. This preference stayed constant throughout the course of the singular trials, with a further increase after the test phase had ended, roughly around 6000 ms post onset. During the incorrectly pointed plural trials, participants first prefer the 2-actor picture as well but then, during the test phase, switch to prefer the 1-actor picture. This is of course the one to which they finally pointed to.

Still, for reasons of comparison, the looking times were as well calculated for these incorrectly pointed trials. The mean looking times to the 1-actor and 2-actor picture per condition can be seen in Figure 11.7. They were entered into two separate ANOVAs for the two pictures presented.

1-actor picture The 2x2 ANOVA comparing Phase and Number condition for the incorrectly pointed trials revealed a marginal main effect for Phase ($F(1,16)=3.849$, $p=0.067$) and a marginal effect for Number ($F(1,16)=3.822$, $p=0.068$). Because no interaction emerged, no post-hoc analyses were performed.

[129] Note that due to empty data cells in the fixation duration analysis performed for these data, only 17 children could be included in this analysis.

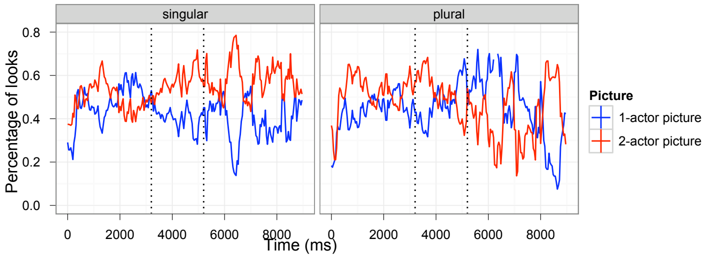


Figure 11.6: Time course of percentage of looks to 1-actor and 2-actor picture in the **incorrectly pointed trials** only, separated by Number condition. (The dotted lines represent the onset and offset of the test phase.)

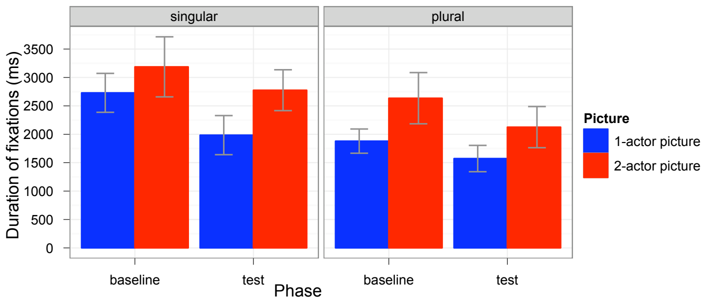


Figure 11.7: Mean fixation durations to 1-actor and 2-actor picture for the **incorrectly pointed trials** only, separated by number condition.(Errorbars: +/- 1SE)

Overall, children looked more to the 1-actor picture in the singular condition than in the plural trials, but in both number conditions, looks to the 1-actor picture decreased from baseline to testing.

2-actor picture In the 2x2 ANOVA that was performed on the looking times to the 2-actor picture, no effect or interaction was found. Therefore, the looks to the 2-actor picture does not seem to be guided by verb inflection.

These looking patterns of the correctly and correctly pointed trials do not indicate that children first find the correct picture but then just point to the wrong one or that children stick to the picture they have first decided upon. The pattern can best be explained in terms of simple guessing strategies that do not seem to have any connection to the linguistic input. It must be noted that the division according to pointing reactions does not provide any further insight into this online-experiment, except that it shows that children come up with a decision rather late, about 5 to 7 sec. into the trial.

11.4 Summary of Results

The aims of this last experiment were to verify the findings from Experiment 2, that children did not show comprehension of verb inflection when asked to select a matching picture via pointing, and to examine whether a more on-line version of a trial could maybe enhance children's comprehension performance. Additionally, adults' use of verb inflection in sentence comprehension was investigated using a picture selection task.

First of all, adults performed at ceiling in the picture selection task. This pattern was as well reflected in their looking behavior. Adults' proportions of looks to the target picture exceeded chance level in the test phase after the critical verb inflection had been presented, while it was at chance level before the sentence had been presented. The same pattern was found in the looking time analysis. The looking times to the 1-actor picture increased when a singular inflected verb had been presented, but decreased when a plural inflected verb had been presented. Looking times to the 2-actor picture showed just the opposite pattern: significant decrease in the singular trials, but significant increase in the plural trials. Further, adults displayed a preference (albeit non-significant) to look towards the 2-actor picture after the pronoun *sie* had been presented, indicating a plural interpretation of the pronoun (although it might well be the case that a visual 2-actor preference was still at work).

For the children, a more on-line trial procedure without disappearing and reappearing pictures did not seem to have an impact on comprehension performance. The results of the three year old children were remarkably similar to the ones obtained in Experiment 3. First, the children in the current experiment did not seem to be able to comprehend the verb inflection as a cue to subject number when asked to *point* to the picture that best matched the test sentence. Second, children's eye gaze did not reveal an influence of test sentence on the proportion of looks to target or the looking times towards the pictures presented on the eye-tracking screen. When the trials were analyzed separately according to correctness of pointing reaction, no effects in the critical test phases could be observed. But trial course data for the correctly and incorrectly pointed trials showed that it took children longer than 5 sec. post trial onset to decide on a pointing reaction (the critical verb inflection was presented 2900 ms post trial onset). This can be seen as an indication that late preference for one of the pictures was not guided by verbal input.[130]

Thus, three year old German children do not seem to be able to show correct comprehension of verb inflection when tested in a task that asked them for an explicit reaction, no matter if the pictures stayed on the screen or first disappeared and then reappeared. Still, the current experiment provided some further insights: First, we found that adults have no problems at all to point to a matching picture when presented with a sentence in which the only cue to subject number is the verb inflection. Second, the adults eye gaze does not seem to be influenced by the additional pointing task. Third, the finding that German children aged three to four years are not able to show comprehension of verb inflection when tested with a picture selection task is a robust finding.

[130] Swingley and Aslin (2000) only analyze looking behavior that occurred within 2 sec after the onset of the target word, because they claim that "previous research has suggested that the few eye-movements occurring after this time are usually spontaneous re-fixations unrelated to the spoken stimulus." (pg. 55). This criterion is applied to the investigation of young children's word comprehension in sentences like *Where is the X?*.

11.5 Discussion

This discussion only deals with specific findings of this last control experiment. How the lack of comprehension found in a picture selection task can be explained in relation to children's ability to use verb inflection in sentence comprehension when tested with a pure eye tracking task, was discussed in great length in Chapter 10.5. The reader is referred to the mentioned chapter.

The findings from Experiment 3 were replicated in this present experiment. Since the results in Experiment 2 and 3 mutually support each other, the finding that German three and a half year olds are not able to infer the subject number by relying on the verb inflection is considered to be robust. As discussed, the inability to access knowledge about the subject-verb agreement relation and use the rules during sentence comprehension might be due to weak and unreliable representations in the grammatical system of three to four year olds. Adults on the other hand did not have any problems directing their eye gaze to matching picture (a replication of the findings from Experiment 2) and to point to the target picture. Therefore, adults knowledge of verb inflection is stable and can be easily used during sentence processing.

Adults' looking behavior following the pronoun *sie* suggests that they assign a plural interpretation to this homophonic pronoun. Alternatively, one could assume that adults' looking behaviour in the time frame following the pronoun simply reflects a preference for the 2-actor picture, but there are two points of evidence that make this latter assumption less likely. The first is the general premise of eye-tracking studies in psycholinguistic research, namely that language comprehension and eye gaze are closely time-locked. Thus, when adults are presented with a spoken language and a visual display that depicts what the language refers to, they spontaneously look at the visual referents of the words that they have heard (e. g. Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy, 1995; Tanenhaus, Magnuson, Dahan, and Chambers, 2000; Altmann and Kamide, 2007; Eberhard, Spivey-Knowlton, Sedivy, and Tanenhaus, 1995). Arnold, Eisenband, Brown-Schmidt, and Trueswell (2000) have found in an eye tracking study that adults were able to use gender information provided by a pronoun (*he* versus *she*) in sentence comprehension approximately 200ms after the pronoun offset to find a target referent that was visually displayed. This makes it likely that the adults' looking behavior measured in the present study as well reflects semantic processing. The preference for the 2-actor picture would then indicate that adults assigned a plural reading to the pronoun. The second hint comes from the time course graphs (see Figure 11.2). When the picture first appeared on the screen, adults looked more to the 2-actor picture in both number conditions. At around 2000 ms post onset, when the verbal stimulus (the pronoun) began, the percentage of looks to both pictures was equal, again in both number conditions. Then, right after the presentation of the pronoun, again a preference for the 2-actor picture emerged (in both number conditions). This looking pattern indicates a link between the verbal input adults received and their eye gaze.

It is so far unclear why adults assigned a plural interpretation to the pronoun *sie*. Differing frequencies for the singular pronoun *sie* and the plural homophone come first to mind. But lists of word frequency in German merge the singular and the plural form of the pronoun and only count the phonological form. Wängler (1963), for example, lists *sie* as the 10^{th} most frequent word in German, but this is collapsed over all word types.

It is unclear whether children have assigned an interpretation regarding number to the pronoun. If so, it rather seems to be the singular interpretation. Children did not show a difference in percentage of looks or number of first looks towards the two pictures, a pattern which contradicts the usually found preference for the 2-actor picture. Thus, one could carefully suspect that children assign a singular reading which overrides the 2-actor picture preference. But this idea is not very well supported by the data. If children would have assigned a singular reading to the pronoun immediately, their looking behavior and their pointing in the singular trials should be better than chance. Because children are found to be unable to revise initial parsing decisions, at least concerning garden-path sentences, even if they lead to completely wrong sentence interpretations (e. g. Kidd et al., 2010; Trueswell et al., 1999; Choi and Trueswell, 2010), a singular interpretation of the pronoun should have made them point or look to the 1-actor picture much more often, irrespective of verb inflection. This is not found to be the case. But Arnold, Brown-Schmidt, and Trueswell (2007) found that English-speaking four year old children can use gender information carried by pronouns online for findings a referent in an eye tracking experiment. Thus, children around the age tested seem to compute grammatical features of pronouns on-line, at least regarding gender information. There is no available evidence indicating that they already compute number information, especially for number ambiguous forms.[131]

Additionally, this experiment revealed that the procedure of a trial did not influence children's performance, since results in Experiment 3 and 4 did not differ, but the trial procedure did. In Experiment 3, the pictures on the screen disappeared and reappeared, more closely resembling the trial procedure used by Kedar et al. (2006), although Kedar and colleagues repeated the test sentence once the pictures had re-appeared on the screen. Still, the first presentation of the test sentence occured when the screen was black. Kedar and colleagues examined how fast and accurately children were at finding a target picture when the object noun referring to the target was preceded by a grammatical or a nonce determiner. The most revealing measure in their study was 'number of first looks to target' per condition.[132] Difference in proportion of looking times to target was no revealing measure.

The trial procedure that we used in the last experiment more closely resembled the one used by Zangl and Fernald (2007), Van Heugten and Johnson (2011), or Mani and Plunkett (2010b),

[131] Tyler (1983) concluded from a mis-pronounciation detection task in Dutch (in which she tested pronoun processing) that children could not use cues like gender or number for online pronoun comprehension until age seven. But the mispronounciation detection task might be rather unnatural and challenging for children, thereby giving rise to late pronoun 'comprehension'.

[132] This measure did not reveal any differences between the verbal conditions in Experiment 2, 3, or 4, therefore it is not reported.

simply because the visual stimuli were visible over the whole course of a trial, roughly 2 sec before the verbal stimulus was presented and 5 sec after it was presented. If anything, we expected the latter kind of trial procedure to be more likely to elicit a looking pattern than indicated comprehension, because children would not be confused by the disappearence and reappearence of the pictures within a trial. As stated above, this was not the case. Therefore, the influence of task demands seems to be the most critical factor in accounting for the within-modality asymmetry we obtained with regard to comprehension of verb inflections. How such within-modality asymmetries can be explained with regard to current theories of language acquisition and how they can be informative when discussing across-modalitiy asymmetries will be subject of the General Discussion.

Part III

General Discussion

12 General Discussion

12.1 Summary

The aim of this thesis was to investigate the receptive side of verb inflection processing in German pre-school aged children. To investigate *comprehension*, we tested whether children were able to match a sentence that contained a singular inflected verb (and a number ambiguous subject pronoun) to a picture showing one actor performing an action and a sentence that contained a plural inflected verb (and the same number ambiguous subject pronoun) to a picture showing two actors performing an action together. Additionally we examined whether the methodology used affected children's comprehension performance by testing children (1) with a pure eye tracking experiment and (2) with an experiment in which children had to perform a picture selection task while their eye gaze was tracked. Further, children's *sensitivity* to subject-verb agreement violations was tested in a pure preferential looking task. The comprehension tasks relate rather to semantic aspects of agreement processing while the latter sensitivity task is thought to rather reflect syntactic processing or processing of distributional regularities of forms.

The preferential looking task examining sensitivity to form did not provide any evidence that children before the age of five years detect subject-verb agreement violations (Experiment 1).[133] A referential context which was provided to clarify the number of the sentence subject, so that children 'simply' had to access the number value of the verb and match it to the number

[133] Three year olds only showed very limited influence of grammaticality on their looking behavior, which is not considered persuasive evidence that they detected agreement violations. The younger children might be able to detect subject-verb agreement violations but the task employed did not show this.

value of the subject, did not seem to have any helpful effect for children aged three and four. It is argued that, rather on the contrary, this referential aspect of the experiment accounts for the fact that we only found much later sensitivity to agreement violations than previous HPP-studies had found (e.g Soderstrom, 2002; Polisenska, 2010; Nazzi et al., 2011). The referential context might have prompted children to perform semantic processing instead of relying on distributional regularities, which seems to be much more difficult for children (Naigles, 2002; Soderstrom, 2008).

Regarding the comprehension of verb inflection, we found that German-speaking children aged three and a half years were able to make use of verb inflection to infer the number of an ambiguous sentence subject but only in an ideal processing situation (Experiment 2). At the same time, children did not seem to understand the connection between verb form and subject meaning when additional task demands in the form of picture selection were imposed on them (Experiment 3). The detrimental effect of additional task demands did not vanish when the trial course was altered so that the visual material remained visible during the presentation of the test sentences (Experiment 4). The preferential looking data is interpreted as evidence that German-speaking children aged three to four years possess the morphosyntactic knowledge necessary to understand the connection between verb form and subject meaning, but the picture selection data and the children's eye gaze under the influence of additional task demands indicate that children are not able to access the knowledge in every context.

As stated earlier, the findings of the present experiments raise some important questions, which shall be repeated here. First, what underlies children's success in demonstrating their comprehension ability when tested with a pure preferential looking task versus their failure to do so when tested with a picture selection task? That is, what gives rise to the *within-modality asymmetry* in the receptive modality (a similar question arises for the productive modality)? Second, what can such within-modality asymmetries reveal about proposed *across-modality asymmetries* with regard to verb inflection acquisition?

12.2 Within-modality asymmetries

Concerning the first question of why children were able to show comprehension of verb inflection in a preferential looking task (Experiment 2) but not in a picture selection task (Experiment 3 and 4), at two possible explanations are conceivable. Both call on the notion of 'higher task demands' in the picture-selection task. It should be noted that it is not perfectly clear why these might arise. It is rather unlikely that the pointing gesture itself is too demanding, since already infants point at people and objects, usually before producing referential speech (e. g. Tomasello, Carpenter, and Liszkowski, 2007; Behne, Liszkowski, Carpenter, and Tomasello, 2011). But a picture selection task demands further abilities from children, namely storing linguistic and visual information in parallel, comparing the information and finally making a decision. One could hypothesize that any necessary capacities to accommodate for these additional demands might not be fully developed in children at age three.[134] Preferential looking, on the other hand, considers a dependent variable that reflects a rather automatic reaction of the organism which is under little, if at all, conscious control of the participant (e. g. Huettig and Altmann, 2005; Tanenhaus et al., 1995; Trueswell, 2008). Additionally, it is closely time-locked to the processing of the incoming speech (e. g. Trueswell and Gleitman, 2007; Thorpe and Fernald, 2006).

[134] In an experimental study from Fernandez et al. (2005), children aged 27 to 35 months were tested on their comprehension of transitive and intransitive sentences using a picture selection task. The authors report data from 48 children, but mention that an additional 48 children had to be discarded from analysis because of various reasons (among these were 'fussing out' (N=21) and 'picking only one side' (N=11). This enormous drop out rate can be seen as a hint that picture selection is not always an easy task, especially not when rather challenging linguistic structures are under investigation. But see for example Shady and Gerken (1999) for a very successful testing procedure and a low drop-out rate when employing the picture selection method.

As Gerken and Shady (1996) point out, we do not yet know whether preferential looking and picture selection tap the same processes, with looking being easier because of not requiring a choice.

Recent findings by Love, Walenski, and Swinney (2009) indicate that different processing routines might be engaged in on-line and off-line tasks. The authors investigated the effects of slowed speech on the processing of sentences containing pronouns using a cross-modal priming task (to tap on-line processing) and a sentence/picture matching task (to tap off-line processing). Children between the age of five and thirteen were tested. They found an interesting dissociation between the two measures. Slowed speech had an improving effect on children's performance in the off-line measure. But the children's automatic syntactic parsing (on-line processing) was rather disrupted by a slowed rate of speech (just as found for adults). Interestingly, children were found to correctly activate the antecedent of a pronoun in on-line measures but showed chance-level performance in off-line measures, a similar pattern to the one found in the present study. Love and colleagues interpret their findings as evidence that different processing routines are engaged in the two types of tasks. On-line measures are thought to reflect automatic 'unconcious' processing while off-line measures are thought to depend on meta-linguistic 'reflective' processes.

Other researchers who investigated children's comprehension of linguistic structures like pronouns and focus particles have found the same pattern, i. e. earlier or better performance when measuring children's eye gaze instead of relying on explicit pointing decisions (e. g. Bergmann et al., 2011; Höhle et al., 2009; Chan et al., 2010; Beyer and Hudson-Kam, 2009). But this pattern does not always hold true, as some researchers obtained very similar results from preferential looking data and pointing reactions (e. g. Legendre et al., 2010). But overall, eye tracking might be more adequate to examine children's processing of syntactic structures (e. g. Golinkoff et al., 1987; Kouider et al., 2006; Heine, Thaler, Tamm, Hawelka, Schneider, Torbeyns, Smedt, Verschaffel, Stern, and Jacobs, 2010).

Thus, the first explanation is strongly related to the one put forward by Love et al. (2009). It refers to different ways of processing speech during sentence comprehension that might be reflected by differences in on-line measurements (eye tracking) vs. off-line responses (picture selection). It seems probable that the latter kind of response is more heavily influenced by heuristic processing strategies than the former one, leading to false interpretations or guessing in the latter. An effect of heuristics in sentence interpretation, giving rise to "quick-and-dirty processing principles" (Hurewitz et al., 2000, pg. 622), can even be found in adults when off-line tasks are used (Ferreira, 2003).[135] If such heuristics can lead to 'false' interpretations in adults, it seems reasonable to expect similar effects in children's sentence interpretation. Such an explanation is also discussed by Hurewitz et al. (2000). According to this, I suggest that the eye-tracking data in Experiment 2 might reflect an automatic analysis of the structural properties of a sentence. Such an analysis might have been pursued in the children tested in Experiment 3 as well, but this was shadowed by heuristics and/or possibly the use of a non adult-like rule that did not provide information about semantic features of the verb inflection. The idea of different stages in the interpretation process that might be assessed in on-line and off-line tasks is also discussed by Trueswell and Gleitman (2007) and Sekerina et al. (2004). This explanation implies that children in general have the linguistic competence to interpret the test sentences adequately, but that heuristic processing strategies kick in because some aspect of the task and/or the stimuli posed additional demands on children. A similar point is made by Clark and Hecht (1983), already extending to performance differences which might extend across modalities.

[135] A mismatching heuristic would for example be an agent-first strategy when interpreting a sentence like *The dog was bitten by the man*, resulting in the assumption that the dog would be the agent of the biting-action (which is certainly more likely according to ones word knowledge).

> Prior to the acquisition of a linguistic rule, children are usually said to rely on heuristics or strategies for comprehending or producing language. But such strategies may often be isomorphic with the adult's "rules", except that they apply only in one kind of task, in imitation, say, but not in production, or in comprehension but not in elicited judgements of anomaly or grammaticality, instead of in all domains of language use. (Clark and Hecht, 1983, pg. 335)

The second explanation is more concerned with children's ability to access morphological information during sentence processing. Experiment 2 rules out the possibility that the morphosyntactic knowledge regarding verb inflection has not been acquired by children around the age of three and a half. But it might be the case children are not able to access the information under certain circumstances and therefore fail to use it during sentence comprehension. Two theoretical frameworks explain processing difficulties this way. The Competition Model (Bates and MacWhinney, 1989; Bates et al., 1984, e. g.) and the Constraint-based lexicalist theory (CBL) (e. g. Trueswell and Tanenhaus, 1994; Trueswell and Gleitman, 2004, 2007) both emphasize the role of cues in sentence processing and language development.[136]

In the framework of the Competition Model, it is assumed that (English) verb morphology is only a weak cue to subject number (e. g. Bates et al., 1991; Dick et al., 2004).[137] Because there are usually other, more reliable cues to subject number in

[136]Trueswell and Gleitman themselves note that their acquisition account of the CBL is reminiscent of the Competition Model, as both theories assume constraint-satisfaction mechanisms for language acquisition and use. They further emphasize the important role of information reliability to explain developmental patterns. But the critical difference between the models lies in the role they assign to detailed linguistic representations. CBL assumes that detailed linguistic representations on multiple, partially independent dimensions (lexical, syntactic, semantic) play a critical role in language processing (e. g. Trueswell and Gleitman, 2007).

[137]Note that empirical work within the Competition Model typically assesses the validity of subject-verb agreement as a cue to thematic role assignment,

English and German sentences (such as the determiner, subject noun morphology, quantifiers etc.), verb inflection is a redundant cue for the process of 'finding the number value' of the sentence subject. All of these cues vary in validity, reliability and availability[138], thus the mappings are acquired at different rates[139] and they vary in the cost of using them in sentence processing once acquired. The results of Experiments 2 and 3 could be interpreted as indicating that German children aged three to four years have acquired the verb inflection paradigm, but they are unable to use this verb inflection cue in sentence processing *under certain circumstances*. When the task requires further attention, as the picture selection task does compared to the pure eye tracking task, even a cue that a child knows about might be too costly to use. Such an explanation is consistent with with other findings. Von Berger et al. (1996) claimed that children have relatively low processing resources when it comes to sentence comprehension, so even highly valid cues can be too costly for a child to use. Leech et al. (2007) found that children aged five to seventeen years were affected in their sentence comprehension ability when attentional demands increased. Such findings can be viewed as evidence that comprehension of a certain linguistic structure is not all or nothing, but might be working well under optimal processing conditions and deteriorate when additional demands are put on the children.

The assumption that verb inflection is a weak cue to subject number, is supported by language processing studies with adults. Blackwell and Bates (1995) found that agreement detec-

not to number assignment on the subject noun. But I claim that this should be extendable to determine the number of the sentence subject, since there are other cues to subject number which are more frequently available and easier to perceive. Additionally, the phonetic realizations of verb agreement markers are the same, irrespective whether this information is used to determine the agent of the sentence or the number of the agent.

[138] Eberhard et al. (2005) propose an account according to which determiners and nouns make separate contributions to the overall number marking of the noun phrase, with the contribution of nouns being weighted much more heavily than that of determiners. This is phrased for adult agreement production.

[139] Noun plural morphology is acquired before verb morphology (e. g. Leonard et al., 2000; Hsieh et al., 1999; Zapf and Smith, 2007).

tion was vulnerable in healthy adults when tested in a dual-task paradigm in which listeners had to memorize digits while judging the grammaticality of sentences. Similar, German adults' ability to detect agreement violations was found to deteriorate under stress (Dick, Bates, Ferstl, and Friederici, 1999; McDonald, 2008a), "indicating that agreement information is especially vulnerable under adverse processing conditions, including perceptual degradation and dual task conditions" (Wulfeck et al., 2004, pg. 225). Agreement is also known to be difficult for Broca aphasics (e. g. Wulfeck, Bates, and Capasso, 1991) and for adults learning a second language. English-Spanish and Chinese-English bilinguals have difficulties detecting subject-verb number agreement despite high levels of proficiency in their second language (Foote, 2011; Jian, 2004). It is argued that 'integrated knowledge', which is supposed to be necessary to process and produce language efficiently like native speakers, is not present in L2 (Jian, 2004).

The cross-linguistic data presented in this thesis, especially with regard to Spanish and Italian, does not fit the predictions made by the competition model. Considering that Italian-speaking adults rely most heavily on agreement cues, and that verb inflection is perceptually more salient in Spanish and Italian than the English *-s* is, Spanish- and Italian-speaking children should be able to exploit agreement cues in sentence comprehension earlier than their English-speaking peers. This was found not to be the case (Pérez-Leroux, 2005; Miller and Schmitt, 2009; Dispaldro and Benelli, 2012).

The constraint-based lexicalist model (CBL) as well calls on the notion of cues, i. e. the ability to use various linguistic and non-linguistic sources of information, to explain why children differ in their sentence processing from adults. Importantly, CBL proponents do not assume a qualitative difference in the parsing strategies of adults and children, but children are thought to process language in a statistics-driven incremental manner, just as adults, and to be sensitive to various cues while trying to recover the meaning of a sentence (e. g. Trueswell and Gleitman, 2007). Critically, children attend to the most reliable (and most frequently encountered) cues to sentence interpretation and are only able to utilize less reliable cues after they have accumulated sufficient evidence to support the use of a particular cue (Trueswell and Gleitman, 2004).

> ...the more frequent and reliable in the input is an observable property of the system being learned, the sooner a learner will exploit this property in making parsing decisions. Thus, a youngster may appear to be deaf to a particular cue [...] not because his or her comprehension architecture is immune to such information but because the relevant knowledge base for using such a cue type either doesn't exist (yet), isn't fully automatized, or hasn't been integrated with other cue types. (Trueswell and Gleitman, 2004, pg. 338).

The CBL model additionally assumes 'probabilistic processing continuity' as one of its organizing principles. Thus, early sensitivity to distributional properties of the speech signal and the ability to extract grammatical properties of utterances are viewed as the first critical steps into parsing. Young children's ability to distinguish structures with correct subject-verb agreement from those with agreement violations (based on the distribution of forms), as found by Soderstrom (2002) and Nazzi et al. (2011) therefore does not come as a surprise. Additionally, the theory aims to explain "how children construct the automatic mechanisms for rapid and efficient language understanding, in the age range from 4 to 6 years." (Trueswell and Gleitman, 2007, pg. 27). This age range matches what has been found for the productive mastery and comprehension of verb inflections. Interestingly, the CBL model actually expects *comprehension-specific deficits*, as it predicts that children might be able to produce a particular linguistic structure correctly while showing an inability to understand the very same structure in a comprehension setting (Trueswell and Gleitman, 2004; Hurewitz et al., 2000).

Thus, within-modality asymmetries can be explained under the assumption that on-line and off-line tasks measure different processing routines. Alternatively, within-modality asymmetries can be explained under the assumption that children might be unable to access a specific cue (in the present case verb inflection) during sentence processing under certain circumstances.

It should be noted though that the findings of earlier and/or more elaborate knowledge in children using eye tracking is not restricted to language. Other researchers have found earlier understanding of false belief (e. g. Clements and Perner, 1994; Southgate, Senju, and Csibra, 2007) or earlier knowledge of numerical magnitude (e. g. Heine et al., 2010), when children's eye gaze was used as dependent measure instead of gestures or verbal answers. In these other domains of cognitive development, asymmetric findings across tasks are often explained with the level of knowledge that is assessed using the different measures, i. e. eye gaze measures are thought to tap *implicit* knowledge while behavioral measures are thought to rely on *explicit* knowledge (e. g. Clements and Perner, 1994; Poltrock, 2011).

Overall, the relation between tasks that only rely on looking behavior and those that demand overt responses, plus the question of which consequences the second type of task has on children's looking behavior, is far from clear and must should be examined in future research. Additionally, "[the] dissociation in performance also underscores the need to examine child language processing from the perspective of multiple methods" (Love et al., 2009, pg. 301).

This need to examine child language using multiple methods is as well evident for the productive modality. As repeatedly pointed out, children show higher performance rates regarding the production of verb inflections in spontaneous speech data than in elicited production tasks (e. g. Rice et al., 1995; Poeppel and Wexler, 1993; Ott, 2011). This dissociation might be due to frequency and familiarity factors, since children are likely to produce familiar and high frequent verbs in their own speech while they have to deal with whatever they get in elicited production tasks, sometimes even pseudo verbs. But the dissociation might also be due to the kind of knowledge necessary to perform well in both tasks. In spontaneous speech, children can recall complete constructions, in which case the form of the verb inflection might be linked to the form of the (subject) noun. Recent evidence shows that children can produce irregular noun plural forms (e. g. *teeth*) much better when those forms are embedded in lexically specific constructions, i. e. in a frame that actually precedes the irregular plural in actual speech (Arnon and Clark, 2011). This finding is interpreted as evidence that children not only attend to the relation between specific words and phrases, but that they also draw on this knowledge in language production. The same procedure, i. e. drawing on knowledge about specific relations between words, might be at work when children produce regular verb inflections in spontaneous speech. Children can rely on this knowledge much less in elicited production tasks, especially when pseudo verbs are presented. Additionally, most elicited production tasks require semantic processing from the child, such that they are specifically asked to encode the semantic information of e. g. 'two actors' in a sentence or at least an inflected verb. Thus, within-modality asymmetries are not only found in the receptive but also in the productive modality.

12.3 Across-modality asymmetries

The second question, what our experimental findings can tell us about the relationship between production and comprehension of verb inflections in children, has to be faced against the background of the previous discussion. The assumption that productive capacities for some linguistic expressions may precede their comprehension in language acquisition is mainly based on the following observations: (1) children display non-adult-like performance in sentence comprehension tasks like picture selection, truth-value judgements or act-out tasks while (2) they display adult-like use of the same kind of structures in spontaneous sentence production.

In line with other studies, the results of the present study show that it is not trivial to determine at which age a given linguistic structure is correctly produced and comprehended by children. Based on the eye gaze data in Experiment 2, we conclude that the structural competence necessary to use the information provided by verb inflections can be assessed in three to four year old German children. Our data indicate that the comprehension ability is in place between three- and four-years of age, at least in German-speaking children under 'optimal' processing conditions. Based on the parental report data gathered for the present studies and the results available from spontaneous speech and elicited production studies (e. g. Rice et al., 1995; de Villiers and de Villiers, 1973a), it is claimed that adult-like rules for verb inflection are available for the productive domain around the age of four years. This comparison shows that the proposal of a 'true' underlying production-comprehension asymmetry is not justified.

But why can we see the superficial pattern of early verb inflection production and late verb inflection comprehension? The answer to this question depends on the theoretical stand one takes. If one assumes that children possess adult-like grammatical knowledge rather early (when they produce verb inflections in their own speech, as proposed by Hyams (1999), Wexler (1994) or Phillips (1995)), processing constraints have to be that explain the performance patterns in the various tasks. If one on the other hand supposes that children's early productions are not based on adult-like grammatical rules, a lack of early comprehension can simply be blamed on the lack of an adult-like rule that allows for semantic processing. Both approaches therefore seem to be suitable to explain the found pattern, but they both as well reveal some drawbacks.

To begin with the earliest ability to detect subject-verb agreement, both approaches can explain children's sensitivity to the presence or absence of verb inflection in HPP studies. Critically, children around the age of one and a half or two must be credited with the ability to detect and learn distributional dependencies between word forms. This ability does not need to rely on adult-like rules. Soderstrom (2002) argues that the data can better be interpreted as reflecting sensitivitiy to statistical regularities than to adult-like knowledge of verb inflection.

Apparently conflicting findings of early receptive sensitivity and later comprehension abilities can, within a nativist framework, be reconciled by distinguishing between early perceptual knowledge (which can either be viewed as statistical or grammatical) and later semantic comprehension. As Naigles (2002) and Soderstrom (2008) claim, the acquisition of linguistic knowledge itself is best described as being accomplished in various steps. "It is highly likely that knowledge of distributional properties of functional morphemes like *-s* (PERCEPTION of grammar) greatly precedes knowledge of their interpretative implications (COMPREHENSION of the grammar)." (Soderstrom, 2008, pg. 675). Naigles (2002) as well states that language acquisition can be viewed as a "coherent developmental path where children first grapple with the sounds and forms of a language, at least somewhat independent of their meaning [...], and only later work on integrating these forms with their meanings" (pg. 185). Interestingly, this description of the language acquisition renders an item-based approach to language acquisition unnecessary (Naigles, 2002). Appealing to semantic considerations as well allows to fit the findings from Experiment 1 in the broader picture, since it was proposed that young children might not detect subject-verb agreement violations because they engage in semantic processing of the input.

The differentiation between early perceptual knowledge and later semantic comprehension can be extended to the productive modality. It can first describe why children are able to detect grammatical dependencies in HPP-experiment earlier than they are able to produce the same grammatical morphemes that make up these dependencies. But it can also be extended to within-modality asymmetries that have been found in children's language production. As explained, the requirements for semantic processing are higher in elicited production tasks than in spontaneous speech, since the former requires children to linguistically encode a specific semantic information (e. g. the number of the actors). Spontaneous speech rather allows children to pick a certain form and the experimenter has no control of whether this is the form that actually matches the meaning the child intended to convey.

Additionally, children's early productions of verb inflections are, within the nativist framework, thought to be influenced by processing costs (e. g. Phillips, 1995). Under this assumption, limited performance in elicited production tasks might be due to higher processing costs therein. Higher processing costs can also be called on to explain frequency and familiarity effects as well as higher rates of verb inflection, i. e. better performance, in shorter sentences and phonologically simpler words (e. g. Song et al., 2009; Sundara et al., 2011). Thus, varying performance with respect to verb inflection production can be explained within the nativist framework by calling on performance factors and by appealing to semantic considerations.

To explain rather late comprehension of verb inflections, proponents of the nativist framework as well have to assume severe performance limitations in young children and/or a critical influence of the semantic aspect of a task. But, using a method that poses only little demands on children's processing resources and that is likely to measure automatic processing, like preferential looking, one can find earlier comprehension, than when using a task that requires children to explicitly create a sentence interpretation. Alternatively, one can claim that children can access their linguistic knowledge easier in a preferential looking task and can therefore exploit even weak cues in sentence comprehension (see above).

The constructivist approach to language acquisition explains the patterns found in production and comprehension quite differently. The early production data, and the within-production asymmetry that can be found when comparing different methods like spontaneous speech and elicited production, is explained by children's early reliance on memorized chunks of speech. Since a child has the 'free' choice of verbs in spontaneous speech, she can choose to only produce some verbs and combine them with those inflectional endings that she has encountered with that particular verb before (e. g. Rubino and Pine, 1998; Pine et al., 1998). This is not possible in elicited production when the child is forced to produce a particular inflectional form in combination with a word given by the experimenter, possibly even a novel word (Theakston et al., 2003). But based on production data alone, it is almost impossible to decide which theoretical approach is more likely to be a good explanation of what children actually know and do.

> It has repeatedly been claimed that children's early productions rely on ready-made chunks or routine utterances, [...]. Apparently perfect utterance forms can result from reliance on routines; they do not necessarily signal that the child has fully appropriate productive control over the structure in question. Assessing the relation between what children produce and what they exactly understand may be difficult, but the point is that there is no evidence in such productions that *appropriate* production is occurring in the absence of comprehension, much less ahead of comprehension. (Clark and Hecht, 1983, pg. 334)

But the constructivist approach can as well explain an impact of task that leads to within-modality asymmetries in comprehension. If one assumes that the abstraction process is slow and gradual, with unstable and fragile representations and rules in the beginning and robust representations only later in development, an impact of task is easily conceivable (e. g. Abbot-Smith and Tomasello, 2006). While eye-tracking techniques tap weaker representations (or the lack of task enables children to make to use unstable adult-like rules), an explicit decision might need stronger representations (Chan et al., 2010) or children might fall back on an alternative rule that they have employed much more often and which has therefore a higher probability of being used. This implies that even within-comprehension asymmetries are not only due to task demands, but that the strength of representation or knowledge plays an additional role.

A similar point is made by Leonard et al. (2000) and Pawlowska et al. (2008) to explain the asymmetric pattern found across modalities. They claim that production may require less advanced knowledge than comprehension, escpecially when agreement markers are initially produced in only a limited number of lexical contexts. The authors assume that children first rely on noun plural marking to interpret subject number and need to encounter frequent co-occurences of agreement morphology with subjects whose number value is clearly specified to build up stable representations that can be accessed in comprehension tasks (Pawlowska et al., 2008). In this sense, production-comprehension asymmetries can easily be explained within an item-based framework. But one needs to assume different 'levels of knowledge' that are available to the child at various points during development.

To reconcile the two approaches, the following scenario can be sketched: (1) different rules that make up the child's linguistic competence at different stages in development and (2) different demands the various production and comprehension tasks pose on the child. Thus, children's very first productions might simply reflect memorized forms they have stored as constructions (or parts of constructions). But early on, children derive a rule from these constructions. Such a rule could have a form like 'V+/Ø/ $\rightarrow$ V+/s/ following he, she, it, NP+/Ø/' (see 2.6) and it would allow children to produce many correct verb forms and to distinguish 'grammatical' from 'ungrammatical' passages, which either conform to the rule or not. But such a rule would not help children when they have to infer semantics based on the verb affix, since the rule is devoid of meaning. If one assumes that children up to the age of three or three and a half rely on such, 'correct' production and sensitivity to form without semantic comprehension is expected. But around the age of three and a half, children[140] might have abstracted even further and build up a rule that is more adult-like, e. g. '$V_{INF} \rightarrow V_{3SG}$ iff $SubjNP_{3SG}$' (see 2.5). Such a rule would allow the child to produce most verb inflections correctly *and* to use verb inflections in sentence comprehension.[141] But such a newly es-

[140] At least German-speaking children, as Experiment 2 has shown.

[141] Several theoretical accounts explain *how* children proceed from an item-based system to a more general and abstract morphological system have been put forward. According to the *critical mass hypothesis*, children have to acquire a certain amount of lexical items to derive morphosyntactic knowledge (Marchman and Bates, 1994). It is assumed that this process of building up stable morphological knowledge takes children a couple of years. These are possibly due to limited generalization abilites in children, as suggested by Boyd and Goldberg (2011)., giving rise to the repeatedly found frequency and familiarity effects in verb inflection production (e. g. Warlaumont and Jarmulowicz, 2011; Hsieh et al., 1999; Theakston et al., 2003).

tablished rule might be difficult or costly to use during sentence comprehension and production, so that children cannot always rely on it. Additional processing demands (such as in elicited production or comprehension tasks using the picture selection or act-out method) might hinder children from using such a rule. In such cases, children might retreat to the use of earlier established non adult-like rules (which children have used more frequently and which have therefore a higher probability), giving rise to within-asymmetry patterns. In such a case, the production-comprehension asymmetry would only in the beginning be due to grammatical knowledge children have, but then be due to factors that allow or inhibit access an adult-like syntactic rule, which is still fragile or not easily accessible. Empirical predictions of such a 'model' could be examined in the future.

12.4 Final remarks and directions for future research

Critically, the case of verb inflection shows that one should not assume that comprehension is per se an 'easier' task than production, although children often seem more advanced at the former than at the latter. Comprehension and production should rather be viewed as qualitatively different processes (e. g. Clark and Hecht, 1983). It should be kept in mind that there are factors that possibly 'hindered' comprehension in the pointing-while-looking experiments presented in this thesis. First of all, all verbs used had complex codas (*basteln, angeln, fuettern, öffnen, streicheln*). This was a consequence of using only verbs that were bisyllic in both number conditions to keep the verb inflectional affix the only cue to subject number and not have the confounding factor of syllable number. Song et al. (2009) and Sundara et al. (2011) presented experimental evidence that children were better at producting and detecting verb inflections in verbs with a simple coda structure (e. g. *sees*) than a complex coda structure (e. g. *needs*). Additionally, all verbs used in the present experiments were presented in a transitive SVO sentence frame, thus the verb was presented in medial sentence position. The studies by Song et al. (2009) and Sundara et al. (2011) have also provided evidence that children are better at producting verb inflections and detecting agreement violations when the verb is in sentence-final position. It might well be the case that German-speaking children would able to use verb inflections in sentence comprehension even when task-demands are higher, if further linguistic variables were altered such that verb inflections would be easier to process (e. g. by using verbs with simple codas or intransitive verbs that appear in sentence-final position). It is thus hypothesized that linguistic processing demands and non-linguistic task demands might have added up in the present picture-selection experiments. This should be subject to further research. To further validate the claim that German-children are indeed able to comprehend verb inflection affixes as a cue to subject number when tested under ideal processing conditions, an eye-tracking experiment with pseudo-verbs should be performed. Slightly later comprehension might be found, but this could be ascribed to higher processing demands imposed by the unknown lexical material. (e. g. Zangl and Fernald, 2007). Along the same line, it would be interesting to investigate how children inflect pseudo verbs productively in a language different from English, to either support the findings by Theakston et al. (2003) or investigate cross-linguistic differences. Regarding the production-comprehension asymmetry, it would be of great interest to investigate comprehension and production within one group of children or ideally within single children. This has recently been done by Miller (2012). Critically, one should try to employ experimental tasks that pose very similar processing demands on children, but this seems to be a very challenging endeavour, since preferential looking tasks cannot be administered on a single subject basis so far (or the single subject would need to finish a huge amount of trials, which is again unrealistic when testing preschool children).

One of the main conclusions that stem from this thesis is that performance asymmetries within modalities have to be taken into account when discussing asymmetries between production and comprehension capacities. Thus, identifying areas in which production precedes comprehension needs intense empirical research that has to fulfill various methodological requirements. Methods used to study production and comprehension performance should be as similar as possible with respect to the processing demands they pose on children. This is hard to achieve as far as we do not have a detailed picture of the knowledge and the computational operations that are necessary to solve a specific task. However, an understanding of the task dependence of children's performance should not only help us to create the most reliable experimental designs to study language acquisition, but may also provide essential insights about how a child develops to be a competent producer and comprehender of her target language.

References

Abbot-Smith, K. and M. Tomasello (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review 23*, 275–290.

Aguado-Orea, J. J. (2004). *The acquisition of morpho-syntax in Spanish: Implications for current theories of development*. Ph. D. thesis, The University of Nottingham.

Akhtar, N. and M. Tomasello (1997). Young children's productivity with word order and verb morphology. *Developmental Psychology 33*(6), 952–965.

Altmann, G. T. and Y. Kamide (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language 57*, 502–518.

Arnold, J. E., S. Brown-Schmidt, and J. C. Trueswell (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes 22*(4), 527–565.

Arnold, J. E., J. G. Eisenband, S. Brown-Schmidt, and J. C. Trueswell (2000). The rapid use of gender information: evidence of the timecourse of pronoun resolution from eyetracking. *Cognition 76*, B13–B26.

Arnon, I. and E. V. Clark (2011). Why brush your teeth is better than teeth - children's word production is facilitated in familiar sentence-frames. *Language Learning and Development 7*(2), 107–129.

Arunachalam, S. and S. R. Waxman (2011). Grammatical form and semantic context in verb learning. *Language Learning and Development 7*(3), 169–184.

Balason, D. V. and C. A. Dollaghan (2002). Grammatical morpheme production in 4-year-old children. *Journal of*

Speech, Language and Hearing Research 45, 961–969.

Barner, D., T. Lui, and J. A. Zapf (2012). Is 'two' a plural marker in early child language? *Developmental Psychology 48*(1), 10–17.

Barner, D., D. Thalwitz, J. Wood, S.-J. Yang, and S. Carey (2007). On the relation between the acquisition of singular-plural morpho-syntax and the conceptual distinction between one and more than one. *Developmental Science 10*(3), 365–373.

Barriere, I., L. Goyet, T. Nazzi, S. Kresh, and G. Legendre (2011). The representatino of subject-verb agreement in French-learning toddlers: New evidence from the comprehension of an infrequent pattern of pseudoverbs. In N. Davis, N. Mesh, and H. Sung (Eds.), *Proceedings of the 35th Annual Boston University Conference on Language Development*, pp. 38–38.

Bates, E. (1993). *Comprehension and production in early language development: Comments on Savage-Rumbaugh et al.* (Serial number: 233 ed.), Volume 58 of *Monographs of the Society for Research in Child Development*.

Bates, E. (1998). Construction grammar and its implications for child language research. *Journal of Child Language 25*, 443–484.

Bates, E., A. Devescovi, and S. D'Amico (1999). Processing complex sentences: A crosslinguistic study. *Language and Cognitive Processes 14*(1), 69–123.

Bates, E. and B. MacWhinney (1989). Functionalism and the competition model. In B. MacWhinney and E. Bates (Eds.), *The cross-linguistic study of sentence processing*, pp. 3–73. New York, NY: Cambridge University Press.

Bates, E., B. MacWhinney, M. C. Caselli, A. Devescovi, F. Natale, and V. Venza (1984). A cross-linguistic study of the development of sentence interpretation strategies in. *Child Development 55*, 341–354.

Bates, E., B. Wulfeck, and B. MacWhinney (1991). Cross-linguistic research in aphasia: An overview. *Brain & Language 41*, 123–148.

Behne, T., U. Liszkowski, M. Carpenter, and M. Tomasello (2011). Twelve-month-olds' comprehension and production of pointing. *British Journal of Developmental Psychology 30*(3), 359–375.

Behrens, H. and M. Tomasello (1999). And what about Chinese? *Behavioral and Brain Sciences 22*(6), 1014.

Berger, F. and B. Höhle (2012). Restricitons on addition: children's interpretation of the focus particles auch 'also' and nur 'only' in german. *Journal of Child Language 39*(2), 383–410.

Berger-Morales, J., M. Salustri, and J. Gilkerson (2005). Root infinitives in the spontaneous speech of two bilingual children: Evidence for seperate grammatical systems. In *Proceedings of the 4th International Symposium on Bilingualism*, Arizona State University, Phoenix Arizona.

Bergmann, C., M. Paulus, and P. Fikkert (2011). Preschooler's comprehension of pronouns and reflexives: the impact of the task. *Journal of Child Language 39*(4), 777–803.

Berko, J. (1958). The child's learning of English morphology. *Word 14*, 150–177.

Bernal, S., J. L. Lidz, S. Millotte, and A. Christophe (2007). Syntax constraints the acquisition of verb meaning. *Language Learning and Development 3*, 325–341.

Bever, T. G. (1970). The cognitive bias for linguistic structures. In J. Hayes (Ed.), *Cognition and the development of language*, pp. 945–984. New York: Wiley.

Beyer, T. and C. L. Hudson-Kam (2009). Some cues are stronger than others: The (non)interpretation of third person present -s as a tense marker by 6- and 7-year-olds. *First Language 29*(2), 208–227.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development 57*(2), 498–510.

Bittner, D. (2003a). Are there correlations in the acquisition of definite pronouns and finite verbs in German? *ZAS Papers Linguistics 15*, 199–217.

Bittner, D. (2003b). Aspectual interpretation of early verb forms in German. In *ZAS Papers in Linguistics*, Volume 29, pp. 3–25.

Bittner, D. (2003c). The emergence of verb inflection in two german-speaking children. In D. Bittner, W. U. Dressler, and M. Kilani-Schoch (Eds.), *Development of Verb Inflection in First Language Acquisition. A Cross-linguistic Perspective*, pp. 53–88. Berlin: de Gruyter.

Blackwell, A. and E. Bates (1995). Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience 7*, 228–257.

Blom, E. and S. Unsworth (Eds.) (2010). *Experimental Methods in Language Acquisition Research*. Amsterdam: John Benjamins.

Blom, W. B. T. (2003). *From root infinitives to finite sentences - The acquisition of verbal inflections and auxiliaries*. Ph. D. thesis, Universiteit Utrecht.

Bloom, L., K. Lifter, and J. Hafitz (1980). Semantics of verbs and the development of verb inflection in child language. *Language 56*(2), 386–412.

Bloom, P., A. Barss, J. Nicol, and L. Conway (1994). Children's knowledge of binding and coference: Evidence from spontaneous speech. *Language 70*, 53–71.

Bock, K., K. M. Eberhard, J. C. Cutting, and A. S. Meyer (2001). Some attractions of verb agreement. *Cognitive Psychology 43*, 83–128.

Bock, K. and C. A. Miller (1991). Broken agreement. *Cognitive Psychology 23*, 45–93.

Bock, K., J. Nicol, and J. C. Cutting (1999). The ties that bind: Creating number agreement in speech. *Journal of Memory and Language 40*, 330–346.

Boyd, J. K. and A. E. Goldberg (2011). Young children fail to fully generalize a novel argument structure construction when exposed to the same input as older learners*. *Journal of Child Language 39*(3), 475–481.

Boyle, M. K. and L. Gerken (1997). The influence of lexical familiarity on children's function morpheme omissions: a nonmetrical effect? *Journal of Memory and Language 36*, 117–128.

Brandt-Kobele, O.-C. and B. Höhle (2010). What asymmetries within comprehension reveal about asymmetries between comprehension and production: The case of verbal inflection in language acquisition. *Lingua 120*(8), 1910–1925.

Brandt-Kobele, O.-C. and B. Höhle (2011). Asymmetries in child language performance within and across modalitie. In A. Grimm, A. Müller, C. Hamann, and E. Ruigendijk (Eds.), *Production-comprehension asymmetries in child language*, SOLA Studies on Language Acquisition. Mouton de Gruyter: Berlin.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Canfield, R. L., E. Smith, M. Brezsnyak, and K. L. Snow (1997). Information processing through the first year of life. In W. A. Collins (Ed.), *Monographs of the Society for Research in Child Development*, Volume 62. Society for Research in Child Development.

Caprin, C. and M. T. Guasti (2009). The acquisition of morphosyntax in italian: A cross-sectional study. *Applied Psycholinguistics 30*, 23–52.

Cazden, C. B. (1968). The acquisition of noun and verb inflections. *Child Development 39*(2), 433–448.

Chan, A., K. Meints, E. V. Lieven, and M. Tomasello (2010). Young children's comprehension of english svo word order revisited: Testing the same children in act-out and intermodal preferential looking tasks. *Cognitive Development 25*, 30–45.

Chien, Y.-C. and K. Wexler (1990). Children's knowledge of locality conditions in binding as evidence for modularity of syntax and pragmatics. *Language Acquisition 1*(3), 225–295.

Childres, J. B. and M. Tomasello (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology 37*(6), 739–748.

Choi, Y. and J. C. Trueswell (2010). Children's (in)ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology 106*, 41–61.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Clahsen, H. (1986). Verb inflections in German child language: Acquisition of agreement markings and the functions they encode. *Linguistics 24*, 79–121.

Clahsen, H. (2008). Behavioral methods for investigating morphological and syntactic processing in children. In I. Sekerina, E. M. Fernandez, and H. Clahsen (Eds.), *Developmental Psycholinguistics: Online methods in children's language processing*, pp. 1–27. Benjamins: Amsterdam.

Clahsen, H., S. Bartke, and S. Göllner (1997). Formal features in impaired grammars: A comparison of English and German SLI-children. *Journal of Neurolinguistics 10*(2/3), 151–171.

Clahsen, H., S. Eisenbeiss, and M. Penke (1996). Underspecification and lexical learning in early syntactic development. In H. Clahsen (Ed.), *Generative perspectives on language acquisition*, pp. 129–159. Amsterdam: Benjamins.

Clahsen, H. and C. Felser (2006). Grammatical processing in language learners. *Applied Psycholinguistics 27*, 3–42.

Clahsen, H. and M. Penke (1992). The acquisition of agreement morphology and its syntactic consequences: New evidence on German child language from the Simone-corpus. See Meisel (1992), pp. 181–223.

Clahsen, H., M. Penke, and T. Parodi (1993). Functional categories in early child German. *Language Acquisition 3*(4), 395–429.

Clark, E. V. and B. F. Hecht (1983). Comprehension, production and language acquisition. *Annual Review of Psychology 34*, 325–349.

Clements, W. and J. Perner (1994). Implicit understanding of belief. *Cognitive Development 9*, 377–397.

Conroy, A., E. Takahashi, J. Lidz, and C. Phillips (2009). Equal treatment for all antecedents: How children succeed with principle B. *Linguistic Inquiry 40*, 446–486.

Crain, S. and R. Thornton (2003). The acquisition of syntax. In *The Encyclopedia of Cognitive Science*. Macmillan: London.

Dapretto, M. and E. L. Bjork (2000). The development of word retrieval abilities in the second year and its relation to early vocabulary growth. *Child Development 71*(3), 635–648.

de Villiers, J., J. Cahilane, and E. Altreuter (2004). What can production reveal about Principle b? In *Proceedings of GALANA 2004, Generative Approachtes to Language Acquisition North America*, pp. 89–100.

de Villiers, J. and P. de Villiers (1973a). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research 2*(3), 267–278.

de Villiers, J. and P. A. de Villiers (1973b). Development of the use of word order in comprehension. *Journal of Psycholinguistic Research 2*(4), 331–341.

de Villiers, J. and V. E. Johnson (2007). The information in third-person /s/: acquisition across dialects of American English. *Journal of Child Language 34*, 133–158.

de Villiers, P. A. and J. de Villiers (1972). Early judgements of semantic and syntactic acceptability by children. *Journal of Psycholinguistic Research 1*(4), 299–310.

Demuth, K. and E. McCullough (2009). The prosodic (re)organization of children's early English articles. *Journal of Child Language 36*, 173–200.

Dick, F., E. Bates, E. Ferstl, and A. D. Friederici (1999). Receptive agrammatism in english- and german-speaking college students processing under stress. *Journal of Cognitive Neuroscience Supplement 1, 48*.

Dick, F., B. Wulfeck, M. Krupa-Kwiatkowski, and E. Bates (2004). The development of complex sentence interpretation in typically developing children compared with children with specific language impairments or early unilateral focal lesions. *Developmental Science 7*(3), 360–377.

Dispaldro, M. and B. Benelli (2012). Putting singular and plural morphlogy in context. *Journal of Child Language 39*(4), 863–884.

Eberhard, K. M. (1997). The marked effect of number on subject-verb agreement. *Journal of Memory and Language 36*, 147–164.

Eberhard, K. M., J. C. Cutting, and K. Bock (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review 112*(3), 531–559.

Eberhard, K. M., M. J. Spivey-Knowlton, J. C. Sedivy, and M. K. Tanenhaus (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research 24*(6), 409–436.

Ettlinger, M. and J. Zapf (2011). The role of phonology in children's acquisition of plural. *Language Acquisition 18*(4), 294–313.

Fenson, L., P. Dale, J. S. Reznick, D. Thal, E. Bates, J. P. Hartung, S. Pethick, and J. S. Reilly (1993). *The MacArthur Communicative Development Inventory: User's guide and technical manual*. San Diego: Paul H. Brookes.

Ferdinand, R. A. (1996). *The Development of Functional Categories: The Acquisition of the Subject in French.* Holland: Holland Institute of Generative Linguistics.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development 8*, 181–195.

Fernald, A. and G. W. McRoberts (1993, October). Effects of prosody and word position on infants lexical comprehension. Paper presented at the Boston University Conference on Language Development, Boston.

Fernald, A. and G. W. McRoberts (1996). Prosodic bootstrapping: A critical analysis of the argument and the evidence. In J. L. Morgan and K. Demuth (Eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: Lawrence Earlbaum Associates.

Fernald, A., J. P. Pinto, D. Swingley, A. Weinberg, and G. W. McRoberts (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science 9*(3), 228–231.

Fernald, A., D. Swingley, and J. P. Pinto (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development 72*(4), 1003–1015.

Fernald, A., R. Zangl, A. L. Portillo, and V. A. Marchman (2008). Looking while listening: Using eye movements to monitor spoken language comprehension in infants and young children. In I. Sekerina, E. M. Fernandez, and H. Clahsen (Eds.), *Developmental Psycholinguistics: Online methods in children's language processing*. John Benjamins, Amsterdam.

Fernandez, K. J., G. F. Marcus, J. A. D. Nubila, and A. Vouloumanos (2005). From semantics to syntax and back again: Argument structure in the third year old life. *Cognition 100*, B10–B20.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology 47*, 164–203.

Fisher, C. (2002). The role of abstract syntactic knowledge in language acquisition: a reply to Tomasello (2000). *Cognition 82*, 259–278.

Foote, R. (2011). Integrated knowledge of agreement in early and late english-spanish bilinguals. *Applied Psycholinguistics 32*, 187–220.

Foote, R. and K. Bock (2011). The role of morphology in subject-verb number agreement: A comparison of Mexican and Dominican Spanish. *Language and Cognitive Processes 27*(3), 429–461.

Fox, A. V. (2006). *TROG-D. Test zur Überprüfung des Grammatikverständnisses*. Idstein: Schulz-Kirchner Verlag.

Frank, A. and S. Düming (2007). Die Rolle der Sonorität bei spezifischer Sprachentwicklungsstörung - ein Fallbeispiel. Poster auf der 7. Jahrestagung der Gesellschaft für Aphasieforschung und -behandlung (GAB).

Fraser, C., U. Bellugi, and R. Brown (1963). Control of grammar in imitation, comprehension and production. *Journal of Verbal Learning and Verbal Behaviour 2*, 121–135.

Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language & Linguistic Theory 5*(4), 519–559.

Freedman, S. E. and K. I. Forster (1985). The psychological status of overgenerated sentences. *Cognition 19*, 101–131.

Friedman, N. and R. Novrogodsky (2004). The acquisition of relative clause comprehension in Hebrew: A study of sli and normal development. *Journal of Child Language 31*, 661–681.

Gelman, S. A., P. J. Goetz, B. W. Sarnecka, and J. Flues (2008). Generic language in parent-child conversations. *Language Learning and Development 4*(1), 1–31.

Gerken, L. (1996). Prosodic structure in young children's language production. *Language 72*(4), 683–712.

Gerken, L. (2002). Early sensitivity to linguistic form. *Annual Review of Language Acquisition 2*, 1–36.

Gerken, L. and R. N. Aslin (2005). Thirty years of research on infant speech perception: The legacy of Peter W. Jusczyk. *Language Learning and Development 1*(1), 5–21.

Gerken, L. and A. Bollt (2008). Three exemplars allow at least ome linguistic generalization: Implications for generalization mechanisms and constraints. *Language Learning and Development 4*(3), 228–248.

Gerken, L., B. Landau, and R. E. Remez (1990). Function morphemes in young children's speech perception and production. *Developmental Psychology 26*(2), 204–216.

Gerken, L. and B. J. McIntosh (1993). Interplay of function morphemes and prosody in early language. *Developmental Psychology 29*(3), 448–457.

Gerken, L. and M. Shady (1996). The picture selection task. In D. McDaniel, C. McKee, and H. S. Cairns (Eds.), *Methods for assessing children's syntax*, pp. 125–145. MIT Press.

Gerken, L., R. Wilson, and W. Lewis (2005). Infants can use distributional cues to form syntactic categories. *Journal of*

Child Language 32, 249–268.

Gertner, Y., C. Fisher, and J. Eisengart (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science 17*, 684–691.

Gibson, E. (1992). On the adequacy of the Competition Model. *Language 68*(4), 812–830.

Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition 1*, 3–55.

Gleitman, L. R. and E. Wanner (1982). Language acquisition: The state of the state of the art. In L. R. Gleitman and E. Wanner (Eds.), *Language acquisition: The state of the art*. New York, NY: Cambridge University Press.

Goldberg, A. E. (2004). But do we need universal grammar? comment on lidz et al. (2003). *Cognition 94*, 77–84.

Goldin-Meadow, S., M. Seligman, and R. Gelman (1976). Language in the two-year old. *Cognition 4*, 189–202.

Golinkoff, R. M. and K. Hirsh-Pasek (2008). How toddlers begin to learn verbs. *Trends in Cognitive Science 12*(10), 397–403.

Golinkoff, R. M., K. Hirsh-Pasek, K. M. Cauley, and L. Gordon (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language 14*, 23–45.

Gómez, R. (2002). Variablity and the detection of invariant structure. *Psychological Science 13*(5), 431–436.

Gómez, R. (2007). Statistical learning in infant language development. In M. G. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics*. Oxford University Press.

Gómez, R. and L. Gerken (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition 70*, 109–135.

Gómez, R. and J. Maye (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy 7*(2), 183–206.

Goodluck, H. (1996). The act-out task. In D. McDaniel, C. McKee, and H. S. Cairns (Eds.), *Methods for assessing children's syntax*, pp. 147–162. MIT Press.

Grimm, H. and H. Doil (2000). *ELFRA: Elternfragebögen für die Früherkennung von Risikokindern, Handanweisung*. Hogrefe Verlag: Göttingen.

Grinstead, J. (1994). The emergence of nominative case assignment in child Catalan and Spanish. Unpublished master's thesis, University of California Los Angeles UCLA.

Grodzinsky, Y. and T. Reinhart (1993). The innateness of binding and coreference. *Linguistic Inquiry 24*(1), 69–101.

Grüter, T. and A. Fernald (2011, March). Can 2-year olds use the singular-plural distinction in online processing? Poster presentend at CUNY 2011: Conference on Human Sentence Processing, Stanford.

Gualmini, A., S. Maciukaite, and S. Crain (2002). Children's insensitivity to contrastive stress in sentences with ONLY. In *PWPL 9.1.: Proceedings of the 26th Annual PLC*, University of Pennsilvania.

Guasti, M. T. (1993). Verb syntax in Italian child grammar: Finite and nonfinite verbs. *Language Acquisition 3*(1), 1–40.

Guasti, M. T. (2002). *Language acquisition: The growth of grammar*. MIT Press.

Harris, T. and K. Wexler (1996). The optional-infinitive stage in child English. In H. Clahsen (Ed.), *Generative perspectives on language acquisition*, pp. 1–42. Amsterdam: Benjamins.

Haskell, T. R. and M. C. MacDonald (2005). Constituent structure and linear order in language production: Evidende from subject-verb agreement. *Journal of Experimental Psychology 31*(5), 891–904.

Heine, A., V. Thaler, S. Tamm, S. Hawelka, M. Schneider, J. Torbeyns, B. D. Smedt, L. Verschaffel, E. Stern, and A. M. Jacobs (2010). What the eyes already 'know': Using eye movement measurement to tap into children's implicit numerical magnitute representations. *Infant and Child Development 19*, 175–186.

Hendriks, P. and C. Koster (2010). Production/comprehension asymmetries in language acquisition. *Lingua 120*, 1887–1897.

Hendriks, P. and J. Spenader (2006). When production precedes comprehension: an optimization approach to the acquisition of pronouns. *Language Acquisition 13*(4), 319–348.

Hirsh-Pasek, K. and R. M. Golinkoff (1996). The intermodal preferential looking paradigm: A window onto emerging language comprehension. In C. McDaniel, D. McKee, and H. S. Cairns (Eds.), *Methods for assessing children's syntax*, pp. 105–124. Cambridge, MA: MIT Press.

Hochmann, J.-R., A. D. Endres, and J. Mehler (2010). Word frequency as a cue for identifying function words in infancy. *Cognition 115*, 444–457.

Hoekstra, T. and N. Hyams (1998). Aspects of root infinitives. *Lingua 106*, 81–12.

Hohenstein, J. and N. Akhtar (2007). Two-year-olds' productivity with verb inflection. *Journal of Child Language 34*, 861–873.

Höhle, B., F. Berger, A. Müller, M. Schmitz, and J. Weissenborn (2009). Focus particles in children's language: Production and comprehension of auch 'also' in German learners from 1 year to 4 years of age. *Language Acquisition 16*, 36–66.

Höhle, B., M. Schmitz, L. M. Santelmann, and J. Weissenborn (2006). The recognition of discontinuous verbal dependencies in German 19-month-olds: Evidence for lexical and structural influences on children's early processing capacities. *Language Learning and Development 2*(4), 277–300.

Höhle, B. and J. Weissenborn (2003). German-learning infant's ability to detect unstressed closed-class elements in continous speech. *Developmental Psychology 6*(2), 122–127.

Höhle, B., J. Weissenborn, D. Kiefer, A. Schulz, and M. Schmitz (2004). Functional elements in infant's speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy 5*(3), 341–353.

Hollich, G. J., K. Hirsh-Pasek, and R. M. Golinkoff (1998). Introducing the 3-D Intermodal Preferential Looking Paradigm: A new method to answer an old question. In

C. Rovee-Collier (Ed.), *Advances in Infancy Research*, Volume 5, pp. 355–373. Norwood, NY: Ablex Publishing Co.

Hoover, J. R., H. L. Storkel, and M. L. Rice (2011). The interface between neighborhood density and optional infinitives: normal development and specific language impairment. *Journal of Child Language 39*(4), 835–862.

Hornby, P. A. and W. A. Haas (1970). The use of contrastive stress in children. *Journal of Speech and Hearing Research 13*, 395–399.

Houston-Price, C., E. Mather, and E. Sakkalou (2007). Discrepancy between parental report of infant's receptive vocabulary and infants' behavior in a preferential looking task. *Journal of Child Language 34*, 701–724.

Hsieh, L., L. B. Leonard, and L. Swanson (1999). Some differences between english plural noun inflections and third singular verb inflections in the input: The contributions of fequency, sentence position and duration. *Journal of Child Language 26*, 531–543.

Huettig, F. and G. T. Altmann (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition 96*, B23–B32.

Hurewitz, F., S. Brown-Schmidt, K. Thorpe, L. R. Gleitman, and J. C. Trueswell (2000). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in comprehension and production. *Journal of Psycholinguistic Research 29*(6), 597–626.

Hurewitz, F., A. Papafragou, L. R. Gleitman, and R. Gelman (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development 2*(2), 77–96.

Hüttner, T., H. Drenhaus, R. van de Vijver, and J. Weissenborn (2004). On the acquisitin of the german focus particle auch 'too': Comprehension does not always precede production. In A. Brugos, L. Micciulla, and C. E. Smith (Eds.), *Proceedings of the 28th Annual Boston University Conference on Language Development*.

Hyams, N. (1999). Underspecification and modulariy in early syntax: a formalist perspective on language acquisition. In M. Darnell, K. Wheatley, F. Newmayer, M. Noonan, and E. Moravcsik (Eds.), *Functionalism and formalism in linguistics*, Volume 1. Amsterdam: John Benjamins.

Ingham, R. (1998). Tense without agreement in early clause structure. *Language Acquisition 7*(1), 51–81.

Ionin, T. and K. Wexler (2002). Why is 'is' easier than '-s'?: acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research 18*, 95–136.

Jian, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics 25*, 603–634.

Johnson, V. E. (2005). Comprehension of third person singular /s/ in aae-speaking children. *Journal of Language, Speech and Hearing Services in Schools 36*, 116–124.

Johnson, V. E., J. de Villiers, and H. Seymour (2005). Agreement without understanding? The case of third person singular -s. *First Language 25*(3), 317–330.

joo Song, H. and C. Fisher (2005). Who's 'she'? discourse prominence influences preschoolers comprehension of pronouns. *Journal of Memory and Language 52*, 29–57.

Jusczyk, P. W. (1997). *The Discovery of Spoken Language*. MIT Press.

Jusczyk, P. W. and R. N. Aslin (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology 29*, 1–23.

Kail, M. (2004). On-line grammaticality judgements in french children and adults: a crosslinguistic perspective. *Journal of Child Language 31*, 713–737.

Kedar, Y., M. Casasola, and B. Lust (2006). Getting there faster: 18- and 24-month-old infants' use of function words to determine reference. *Child Development 77*(2), 325–338.

Keeney, T. J. and N. Smith (1971). Young children's imitation and comprehension of sentential singularity and plurality. *Language and Speech 14*, 372–383.

Keeney, T. J. and J. Wolfe (1972). The acquisition of agreement in

English. *Journal of Verbal Learning and Verbal Behaviour 11*, 698–705.

Kidd, E., A. J. Stewart, and L. Serratrice (2010). Children do not overcome lexical biases where adults do: the rolte of the referential scene in garden-path recovery. *Journal of Child Language 38*(1), 222–234.

Kilborn, K. and A. Cooreman (1987). Sentence interpretation strategies in adult dutch-english bilinguals. *Applied Psycholinguistics 8*, 415–431.

Kouider, S., J. Halberda, J. Wood, and S. Carey (2006). Acquisition of English number marking: The singular-plural distinction. *Language Learning and Development 2*(1), 1–15.

Leech, R., J. Aydelott, G. Symons, J. Carnevale, and F. Dick (2007). The development of sentence interpretation: effects of conceptual, attentional and semantic interference. *Developmental Science 10*(6), 794–813.

Legendre, G., I. Barriere, L. Goyet, and T. Nazzi (2010). Comprehension of infrequent subject-verb agreement forms: Evidence from French-learning children. *Child Development 81*(6), 1859–1875.

Legendre, G. and P. Smolensky (2012). On the asymmetrical difficulty of acquiring pronoun reference in French: Production versus comprehension. *Journal of Logic, Language and Information 21*(1), 7–30.

Leonard, L. B., S. M. Camarata, M. Pawlowska, B. Brown, and M. N. Camarata (2007). Tense and agreement morphemes in the speech of children with specific language impairment during intervention: Phase 2. *Journal of Speech, Language and Hearing Research 49*, 749–770.

Leonard, L. B., C. M. Caselli, U. Betolini, K. K. McGregor, and L. Sabbadini (1992). Morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition 2*(2), 151–179.

Leonard, L. B., J. A. Eyer, L. M. Bdore, and B. G. Grela (1997). Three accounts of the grammatical morpheme difficulties of english-speaking children with specific language impairment. *Journal of Speech, Language and Hearing Research 40*, 741–753.

Leonard, L. B., C. A. Miller, and A. J. Owen (2000). The comprehension of verb agreement morphology by english-speaking children with specific language impairment. *Clinical Linguistics and Phonetics 14*(6), 465–481.

Lew-Williams, C. and A. Fernald (2007). Young children learning spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science 18*(3), 193–198.

Li, P., T. Ogura, D. Barner, S.-J. Yang, and S. Carey (2009). Does the conceptual distinction between singular and plural sets depend on language? *Developmental Psychology 45*, 1644–1653.

Lidz, J. L., H. Gleitman, and L. R. Gleitman (2003). Understanding how input matters: verb learning and the footpint of universal grammar. *Cognition 87*, 151–178.

Love, T., M. Walenski, and D. Swinney (2009). Slowed speech input has a differential impact on on-line and off-line processing in children's comprehension of pronouns. *Journal of Psycholinguistic Research 38*, 285–304.

Lukács, A., L. B. Leonard, B. Kas, and C. Plé (2009). The use of tense and agreement by Hungarian-speaking children with language impairment. *Journal of Speech, Language and Hearing Research 52*, 98–117.

Lukyanenko, C. and C. Fisher (2010, November). Where are the cookies? 3-year-olds use number-marked verbs to anticipante upcoming nouns. Talk presented at BUCLD 35.

Ma, W., R. M. Golinkoff, D. M. Houston, and K. Hirsh-Pasek (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development 7*, 185–201.

MacWhinney, B., E. Bates, and R. Kliegl (1984). Cue validity and sentence interpretation in English, German and Italian. *Journal of Verbal Learning and Verbal Behaviour 23*, 127–150.

Mani, N. and K. Plunkett (2010a). In the infant's mind's ear: Evidence for implicit naming in 18-month-olds. *Psychological Science 21*(7), 908–913.

Mani, N. and K. Plunkett (2010b). Twelve-month-olds know their cups from their keps and tups. *Infancy 15*(5), 445–470.

Marchman, V. A. and E. Bates (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language 21*, 339–366.

Marchman, V. A., K. Plunkett, and J. C. Goodman (1997). Overregularization in English plural and past tense inflectional morphology: a response to Marcus (1995). *Journal of Child Language 24*, 767–779.

Marcus, G. F. (1995). Children's overregularizatoins of english plurals: A quantitative analysis. *Journal of Child Language 22*(2), 447–459.

Marshall, C. and H. K. J. van der Lely (2006). A challenge to current models of past tense inflection: The impact of phonotactics. *Cognition 100*, 302–320.

McDaniel, D. and H. S. Cairns (1996). Eliciting judgements of grammaticality and reference. In D. McDaniel, C. McKee, and H. S. Cairns (Eds.), *Methods for assessing children's syntax*, Chapter 11, pp. 233–254. MIT Press.

McDaniel, D., C. McKee, and H. S. Cairns (Eds.) (1996). *Methods for assessing children's syntax*. MIT Press.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgement performance by late second language learner. *Journal of Memory and Language 55*, 381–401.

McDonald, J. L. (2008a). Differences in the cognitive demands of word order, plural, and subject-verb agreement constructions. *Psychonomic Bulletin & Review 15*(5), 980–984.

McDonald, J. L. (2008b). Grammaticality judgements in children: The role of age, working memory and phonological ability. *Journal of Child Language 35*, 247–268.

Meints, K., K. Plunkett, and P. L. Harris (1999). When does an ostrich become a bird? the role of typicality in early word comprehension. *Developmental Psychology 35*(4), 1072–1078.

Meisel, J. (Ed.) (1992). *The acquisition of verb placement*. Kluwer Academic Publishers.

Miller, K. (2007). *Variable Input and the Acquisition of Plurality in Two Varieties of Spanish*. Ph. D. thesis, Michigan State University.

Miller, K. (2012). Not all children agree: Acquisition of agreement when the input in variable. *Language Learning and Development 8*(3), 255–277.

Miller, K. and C. Schmitt (2009). Variable vs. consistent input: Comprehension of plural morphology and verbal agreement in children. In J. M. Brucart, A. Gavarró, and J. Sola (Eds.), *Merging Features: Computation, Interpretation and Acquisition*. Oxford University Press.

Mintz, T. H. (2006). Finding the verbs: distributional cues to categories available to young learners. In K. Hirsh-Pasek and R. M. Golinkoff (Eds.), *Action Meets Word: How Children Learn Verbs*, pp. 31–63. New York: Oxford University Press.

Münte, T. F., M. Matzke, and S. Johannes (1997, May). Brain activity associated with syntactic incongruencies in words and pseudo-words. *J. Cognitive Neuroscience 9*, 318–329.

Naigles, L. G. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition 86*, 157–199.

Naigles, L. G. and S. A. Gelman (1995). Overextensions in comprehension and production revisited: preferential-looking in a study of dog, cat, and cow. *Journal of Child Language 22*, 19–46.

Naigles, L. G. and E. Hoff-Ginsberg (1998). Why are some verbs are learned before other verbs? effects of input frequency and structure on children's early verb use. *Journal of Child Language 25*, 95–120.

Naigles, L. R. (2004). Comprehension matters: a commentary on 'a multiple process solution to the logical problem of language acquisition'. *Journal of Child Language 31*, 936–940.

Naigles, L. R., E. L. Bavin, and M. A. Smith (2005). Toddlers recognize verbs in novel situations and sentences. *Devel-*

opmental Science 8(5), 424–431.

Nazzi, T., I. Barriere, L. Goyet, S. Kresh, and G. Legendre (2011). Tracking irregular morphophonological dependencies in natural language: Evidence from the acquisition of subject-verb agreement in French. *Cognition 120*, 119–135.

Nederstigt, U. (2003). *Auch and noch in Child and Adult German*. Berlin: Mouton de Gruyter.

Nelson, D. G. K., P. W. Jusczyk, D. R. Mandel, J. Myers, A. Turk, and L. Gerken (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development 18*, 111–116.

Nevins, A., B. Dillon, S. Malhotra, and C. Phillips (2007). The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Research 1164*, 81–94.

Nicol, J. L., K. I. Forster, and C. Veres (1997). Subject-verb agreement processes in comprehension. *Journal of Memory and Language 36*, 569–587.

Norbury, C. F., D. V. M. Bishop, and J. Briscoe (2001). Production of English finite verb morphology: A comparison of sli and mild-moderate hearing impairment. *Journal of Speech, Language and Hearing Research 44*, 165–178.

Oetting, J. B. and P. A. Hadley (2008). The morphosyntactic deficits in children with developmental disabilities. In R. Schwartz (Ed.), *The Handbook of Child Language Disorders*, pp. 341–364. New York: Psychological Press.

Osterhout, L. and L. A. Mobley (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language 34*, 739–773.

Ott, S. (2011). *Feld-fällt-fehlt: Untersuchungen zur Phonologie-Morphosyntax-Schnittstelle bei Kindern und Erwachsenen*. Ph. D. thesis, Universität Potsdam.

Pawlowska, M., L. B. Leonard, S. M. Camarata, B. Brown, and M. N. Camarata (2008). Factors accounting for the ability of children with sli to learn agreement morphemes in intervention. *Journal of Child Language 35*, 25–53.

Pearlmutter, N. J., S. M. Garnsey, and K. Bock (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language 41*, 427–456.

Pérez-Leroux, A. T. (2005). Number problems in children. In C. Gurski (Ed.), *Proceedings of the 2005 Canadian Linguistics Association Annual Conference*.

Phillips, C. (1995). Syntax at age two: Crosslinguistic differences. In C. Schütze, K. Broihier, and J. Ganger (Eds.), *Papers on Language Processing and Acquisition*, Volume 26 of *MIT Working Papers in Linguistics*. MIT Press.

Pine, J. M., E. V. Lieven, and C. F. Rowland (1998). Comparing different models of the development of the english verb category. *Linguistics 36*(4), 807–830.

Pinto, J. P., A. Fernald, G. W. McRoberts, and S. Cole (1998). Reliability and validity in infant auditory preference procedures. In Rovee-Collier, Lipsitt, and Hayne (Eds.), *Advances in Infancy Research*, Volume 12, pp. 221–236. Norwood, NY: Ablex Publishing Co.

Pizutto, E. and M. C. Caselli (1992). The acquisition of Italian morphology: implications for models of language development. *Journal of Child Language 19*(03), 491–557.

Pizutto, E. and M. C. Caselli (1994). The acquisition of italian verb morpholoy in a cross-linguistic perspective. In Y. Levi (Ed.), *Other children, other languages*. Hillsdale, N.J.: Erlbaum.

Plunkett, K. and G. Schafer (1998). Rapid word learning by 15-month-olds under tightly controlled conditions. *Child Development 69*, 185–205.

Poeppel, D. and K. Wexler (1993). The full competence hypothesis of clause structure in early German. *Language 69*, 1–33.

Polisenska, D. (2010). *Dutch children's acquisition of verbal and adjectival inflection*. Ph. D. thesis, LOT.

Poltrock, S. (2011). *About the relation between implicit Theory of Mind and the comprehension of complement sentences*. Ph. D. thesis, Universität Potsdam.

Pratt, A. and J. Grinstead (2007). Optional infinitives in child

Spanish. In A. B. et al. (Ed.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*. Somerville, MA: Cascadilla Proceedings Project.

Radford, A. (1990). *Syntactic Theory and the Acquisition of English Syntax.* Oxford: Blackwell.

Reznick, J. S. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics 11*, 145–166.

Rice, M. L. and K. Wexler (2001). Examiner's manual for the Test of Early Grammatical Impairment (tegi). Pearson Education, San Antonio, Texas.

Rice, M. L., K. Wexler, and P. L. Cleave (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech, Language and Hearing Research 38*, 850–863.

Rice, M. L., K. Wexler, and S. Hershberger (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language and Hearing Research 41*, 1412–1431.

Rispens, J. and P. Been (2007). Subject-verb agreement and phonological processing in delevelopmental dyslexia and specific language impairment (sli): a closer look. *International Journal of Language and Communication Disorders 42*(3), 293–305.

Rispens, J., S. Roeleven, and C. Koster (2004). Sensitivity to subject-verb agreement in spoken language in children with developmental dyslexia. *Journal of Neurolinguistics 17*, 333–347.

Rubino, R. B. and J. M. Pine (1998). Subject-verb agreement in Brazilian Portugese: what low error rates hide. *Journal of Child Language 25*, 35–59.

Saffran, J. R., R. N. Aslin, and E. L. Newport (1996). Statistical learning by 8-month-old infants. *Science 274*, 1926–1928.

Santelmann, L. M. and P. W. Jusczyk (1998). Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition 69*, 105–134.

Sauerland, U., J. Anderssen, and K. Yatsushiro (2005). The plural is semantically unmarked. In S. Kepser and M. Reis (Eds.), *Linguistic evidence*, pp. 413–434. Berlin: Mouton de Gruyter.

Scott, R. M. and C. Fisher (2012). 2.5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition 122*, 163–180.

Seidl, A., G. Hollich, and P. W. Jusczyk (2003). Early comprehension of subject and object wh-questions. *Infancy 4*(3), 423–436.

Sekerina, I., K. Stromswold, and A. Hestvik (2004). How do adults and children process referentially ambiguous pronouns? *Journal of Child Language 31*(1), 123–152.

Shady, M. (1996). *Infant's sensitivity to function morphemes*. Ph. D. thesis, The State University of New York: University at Buffalo.

Shady, M. and L. Gerken (1999). Grammatical and caregiver cues in early sentence comprehension. *Journal of Child Language 26*, 163–175.

Shi, R. and M. Lepage (2008). The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science 11*(3), 407–413.

Shi, R. and J. F. Werker (2001). Six-month-old infant's preference for lexical words. *Psychological Science 12*(1), 70–75.

Smith, C. L. and H. Tager-Flusberg (1982). Metalinguistic awareness and language development. *Journal of Experimental Child Psychology 34*, 449–468.

Smolensky, P. (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry 27*, 720–731.

Snedeker, J. and J. C. Trueswell (2004). The developing constraints on parsing decisions: The role of lexical biases and referential scenes in child and adult sentence processing. *Cognitive Psychology 49*, 238–299.

Soderstrom, M. (2002). *The acquisition of inflection morphology in early perceptual knowledge of syntax*. Ph. D. thesis, The Johns Hopkins University.

Soderstrom, M. (2008). Early perception - late comprehension? The case of verbal *-s*: a response to de Villiers & Johnson (2007). *JCL 35*, 671–676.

Soderstrom, M., K. Wexler, and P. W. Jusczyk (2002). English-learning toddlers' sensitivity to agreement morphology in receptive grammar. In B. Skarabela, S. Fish, and H.-J. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development*, pp. 643–652. Somerville, MA: Cascadilla Press.

Soderstrom, M., K. S. White, E. Conwell, and J. L. Morgan (2007). Receptive grammatical knowledge of familiar content words and inflection in 16-month-olds. *Infancy 12*(1), 1–29.

Solomon, E. S. and N. J. Pearlmutter (2005). Forward versus backward agreement processing in comprehension. Manuscript in revision.

Song, J. Y., M. Sundara, and K. Demuth (2009). Phonological constraints on children's production of English 3rd person singular -s. *Journal of Speech, Language and Hearing Research 52*(3), 623–642.

Southgate, V., A. Senju, and G. Csibra (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science 18*(7), 587–592.

Stromswold, K. (1996). Analyzing spontaneous speech. In D. McDaniel, C. McKee, and H. S. Cairns (Eds.), *Methods for assessing children's syntax*, pp. 23–54. MIT Press.

Sundara, M., K. Demuth, and P. K. Kuhl (2011). Sentence-position effects on children's perception and production of English 3rd person singular -s. *Journal of Speech, Language and Hearing Research 54*, 55–71.

Sutter, J. C. and C. J. Johnson (1990). School-age children's metalinguistic awareness of grammaticality in verb form. *Journal of Speech and Hearing Research 33*, 84–95.

Swingley, D. and R. N. Aslin (2000). Spoken word recognition and lexical representation in very young children. *Cognition 76*, 147–166.

Swingley, D., J. P. Pinto, and A. Fernald (1999). Continuous processing in word recognition at 24 months. *Cognition 71*, 73–108.

Szendroi, K. (2004). Focus and the interaction between syntax and pragmatics. *Lingua 114*(3), 229–254.

Tanenhaus, M. K., J. S. Magnuson, D. Dahan, and C. Chambers (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., M. J. Spivey-Knowlton, K. M. Eberhard, and J. C. Sedivy (1995). Integration of visual and linguistic information in spoken language comprehension. *Science 268*, 1632–1634.

The R Foundation for Statistical Computing (2009). R: A language and environment for statistical computing. URL: http://www.R-project.org.

Theakston, A. L., E. V. Lieven, and M. Tomasello (2003). The role of the input in the acquisition of third person singular verbs in english. *Journal of Speech, Language and Hearing Research 46*, 863–877.

Thornton, R. (1996). Elicited production. In D. McDaniel, C. McKee, and H. S. Cairns (Eds.), *Methods for assessing children's syntax*, pp. 77–102. MIT Press.

Thornton, R. and M. C. MacDonald (2003). Plausability and grammatical agreement. *Journal of Memory and Language 48*, 740–759.

Thorpe, K. and A. Fernald (2006). Knowing what a novel word is not: Two-year-olds 'listen through' ambiguous adjectives in fluent speech. *Cognition 100*, 389–433.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition 74*, 209–253.

Tomasello, M., M. Carpenter, and U. Liszkowski (2007). A new look at infant pointing. *Child Development 78*(3), 705–722.

Trueswell, J. C. (2008). Using eye movements as a developmental measure within psycholinguistics. In I. Sekerina, E. M. Fernandez, and H. Clahsen (Eds.), *Developmental*

Psycholinguistics: Online methods in children's language processing, pp. 73–96. John Benjamins, Amsterdam.

Trueswell, J. C. and L. R. Gleitman (2004). Children's eye movements during listening: Developmental evidence for a constraint-based theory of sentence processing. In J. M. Henderson and F. Ferreira (Eds.), *The Interface of Language, Vision and Action: Eye movements and the visual world*. New York: Psychology Press.

Trueswell, J. C. and L. R. Gleitman (2007). Learning to parse and its implications for language acquisition. In G. Gaskell (Ed.), *Oxford Handbook of Psycholinguistics*. Oxford University Press.

Trueswell, J. C., I. Sekerina, N. M. Hill, and M. L. Logrip (1999). The kindergarden-path effect: studying on-line sentence processing in children. *Cognition 73*, 89–134.

Trueswell, J. C. and M. K. Tanenhaus (1994). Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In C. Clifton and L. Frazier (Eds.), *Perspectives on Sentence Processing*, pp. 155–179. Lawrence Earlbaum Associates.

Tyler, L. (1983). The development of discourse mapping processes: The on-line interpretation of anaphoric expressions. *Cognition 13*, 309–341.

van der Lely, H. (1998). Sli in children: Movement, economy and deficit in the computational-syntactic system. *Language Acquisition 7*(2–4), 161–192.

Van Heugten, M. and E. K. Johnson (2010). Linking infant's distributional learning abilities to natural language acquisition. *Journal of Memory and Language 63*, 197–209.

Van Heugten, M. and E. K. Johnson (2011). Gender-marked determiners help Dutch learner's word recognition when gender information itself does not. *Journal of Child Language 38*, 87–100.

Van Heugten, M. and R. Shi (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science 12*(3), 419–425.

van Heugten, M. and R. Shi (2010). Infants' sensitivity to non-adjacent dependencies across phonological phrase boundaries. *Journal of Speech, Language and Hearing Research 128*(5), EL223–EL228.

Verrips, M. and J. Weissenborn (1992). Routes of verb placement in early German and French: The independence of finiteness and agreement. In J. Meisel (Ed.), *The acquisition of verb placement*. Kluwer Academic Publishers.

Vigliocco, G. and J. Franck (2001). When sex affects syntax: Contextual influences in sentence production. *Journal of Memory and Language 45*, 368–390.

von Berger, E., B. Wulfeck, E. Bates, and N. Fink (1996). Developmental changes in real-time sentence processing. *First Language 16*, 193–222.

Wagers, M. W., E. F. Lau, and C. Phillips (2009). Agreement attraction in comprehension: representations and processes. *Journal of Memory and Language 61*, 206–237.

Wängler, H. H. (1963). *Rangwörterbuch hochdeutscher Umgangssprache*. Marburg: Elwert.

Warlaumont, A. S. and L. Jarmulowicz (2011). Caregivers' suffix frequencies and suffix acquisition by language impaired, late talking and typically developing children. *Journal of Child Language* (DOI:10.1017/S0305000911000390).

Waxman, S. R. and J. L. Lidz (2006). Early word learning. In D. Kuhn and R. Siegler (Eds.), *Handbook of Child Psychology: Vol. 2. Cognition, Perception and Language*. New York: Wiley.

Weisleder, A. and S. R. Waxman (2010). What's in the input? frequent frames in child-directed speech offer distributional cues to grammatical categories in Spanish and English. *Journal of Child Language 37*, 1089–1108.

Wenzlaff, M. and H. Clahsen (2004). Tense and agreement in German agrammatism. *Brain & Language 89*, 57–68.

Wexler, K. (1994). Optional infinitives, head movement, and the economy of derivations. In D. Lightfoot and N. Hornstein (Eds.), *Verb Movement*. New York, NY: Cambridge University Press.

Wexler, K. (1998). Very early parameter setting and the Unique Checking Constraint: a new explanation for the Optional Infinite stage. *Lingua 106*, 23–79.

Wiese, R. (2000). *The Phonology of German (2nd corrected edition)*. Oxford: Oxford University Press.

Wilson, S. (2003). Lexically specific constructions in the acquisition of inflection in English. *Journal of Child Language 30*, 75–115.

Wojtecka, M., C. Koch, A. Grimm, and P. Schulz (2011). Production and comprehension of sentence negation in child German. In A. Grimm, A. Müller, C. Hamann, and E. Ruigendijk (Eds.), *Production-comprehension asymmetries in child language*, Volume 43 of *SOLA Studies on Language Acquisition*. De Gruyter, Berlin.

Wood, J., S. Kouider, and S. Carey (2009). Acquisition of singular-plural morphology. *Developmental Psychology 45*(1), 202–206.

Woodward, A. L., E. M. Markman, and C. M. Fitzsimmons (1994). Rapid word learning in 13 and 18-month-olds. *Developmental Psychology 30*, 553–566.

Wulfeck, B. (1993). A reaction time study of grammaticality judgements in children. *Journal of Speech and Hearing Research 36*, 1208–15.

Wulfeck, B., E. Bates, and R. Capasso (1991). Cross-linguistic study of grammaticality judgements in Broca's aphasia. *Brain & Language 41*, 311–336.

Wulfeck, B., E. Bates, M. Krupa-Kwiatkowski, and D. Saltzman (2004). Grammatical sensitivity in children with early focal brain injury and children with specific language impairment. *Brain & Language 88*, 215–228.

Xanthos, A., S. Laaha, S. Gillis, U. Stephany, A. Aksu-Koc, A. Christofidou, N. Gagarina, G. Hrzica, F. N. Ketrez, M. Kilani-Schoch, K. Korecky-Kröll, M. Kovacevic, K. Laalo, M. Palmovic, B. Pfeifer, M. D. Voeikova, and W. U. Dressler (2011). On the role of morphological richness in the early development of noun and verb inflection. *First Language 31*(4), 461–479.

Yoshida, K. A., C. T. Fennell, D. Swingley, and J. F. Werker (2009). Fourteen-month-old infants learn similar-sounding words. *Developmental Science 12*(3), 412–418.

Zangl, R. and A. Fernald (2007). Increasing flexibility in children's online processing of grammatical and nonce determiners in fluent speech. *Language Learning and Development 3*, 199–231.

Zapf, J. A. and L. B. Smith (2007). When do children generalize the plural to novel nouns? *First Language 27*(1), 53–73.

Zapf, J. A. and L. B. Smith (2008). Meaning matters in children's plural productions. *Cognition 108*, 466–476.

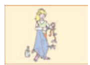
A Appendix

A.1 Verbal and visual material used in Experiment 2, 3 and 4

Introduction

Schau mal. Das sind Julia, Anna und Sarah. Sie machen verschiedene Sachen, mal alleine und mal zusammen. Gleich wirst du sehen, was.

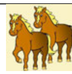
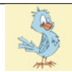
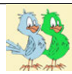
'Hey look. These are Julia, Anna and Sarah. They are going to do a lot of things, sometimes by themselves and sometimes together. You are going to see now, what they do.'

Singular sentence	**Plural sentence**	1-actor	2-actor
Sie fütter-t einen Hund Pronoun-SG feed-3SG a dog 'She is feeding a dog'	*Sie fütter-n einen Hund* Pronoun-PL feed-3PL a dog 'They are feeding a dog'		
Sie fütter-t ein Pferd Pronoun-SG feed-3SG a horse 'She is feeding a horse'	*Sie fütter-n ein Pferd* Pronoun-PL feed-3PL a horse 'They are feeding a horse'		
Sie angel-t einen Fisch Pronoun-SG catch-3SG a fish 'She is catching a fish'	*Sie angel-n einen Fisch* Pronoun-PL catch-3PL a fish 'They are catching a fish'		
Sie angel-t einen Schuh Pronoun-SG catch-3SG a shoe 'She is catching a shoe'	*Sie angel-n einen Schuh* Pronoun-PL catch-3PL a shoe 'They are catching a shoe'		
Sie öffne-t ein Fenster Pronoun-SG open-3SG a window 'She is opening a window'	*Sie öffne-n einen Hund* Pronoun-PL open-3PL a window 'They are opening a window'		
Sie öffne-t ein Geschenk Pronoun-SG open-3SG a present 'She is opening a present'	*Sie öffne-n ein Geschenk* Pronoun-PL open-3PL a present 'They are opening a present'		
Sie bastel-t einen Drachen Pronoun-SG make-3SG a kite 'She is making a kite'	*Sie bastel-n einen Drachen* Pronoun-PL make-3PL a kite 'They are making a kite'		
Sie bastel-t einen Hund Pronoun-SG make-3SG a hat 'She is making a hat'	*Sie bastel-n einen Hund* Pronoun-PL make-3PL a hat 'They are making a hat'		
*Sie streichel-t eine Katze** Pronoun-SG pet-3SG a cat 'She is petting a cat'	*Sie streichel-n eine Katze** Pronoun-PL pet-3PL a cat 'They are petting a cat'		
*Sie streichel-t ein Baby** Pronoun-SG pet-3SG a baby 'She is feeding a baby'	*Sie streichel-n ein Baby** Pronoun-PL pet-3PL a baby 'They are feeding a baby'		

* used in practice trials

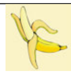
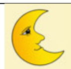
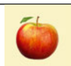
A.2 Verbal and visual material used in Experiment 1

Subject nouns and subject pictures

Nr.	singular noun	plural noun	1-actor pic	2-actor pic
1.	*das Baby* 'the baby'	*die Babys* 'the babies'		
2.	*der Bär* 'the bear'	*die Bären* 'the bears'		
3.	*der Hase* 'the rabbit'	*die Hasen* 'the rabbit'		
4.	*der Hund* 'the dog'	*die Hunde* 'the dogs'		
5.	*das Pferd* 'the horse'	*die Pferde* 'the horses'		
6.	*das Schaf* 'the sheep'	*die Schafe* 'the sheep'		
7.	*das Schwein* 'the pig'	*die Schweine* 'the pigs'		
8.	*der Vogel* 'the bird'	*die Vöegel* 'the birds'		
9.*	*der Affe* 'the monkey'	*die Affen* 'the monkeys'		
10.*	*der Junge* 'the boy'	*die Jungen* 'the boys'		

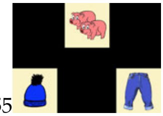
* used in practice trials

Object nouns and object pictures

Nr.	object 1	object 2	referent pic 1	referent pic 2
1.	*die Bürste* ‘the brush’	*die Flasche* ‘the bottle’		
2.	*die Wurst* ‘the sausage’	*die Banane* ‘the banana’		
3.	*der Mond* ‘the moon’	*der Stern* ‘the star’		
4.	*der Ball* ‘the ball’	*der Schuh* ‘the shoe’		
5.	*der Käse* ‘the cheese’	*der Kuchen* ‘the cake’		
6.	*die Tür* ‘the door’	*die Dose* ‘the can’		
7.	*die Hose* ‘the pants’	*die Mütze* ‘the hat’		
8.	*der Tisch* ‘the table’	*der Stuhl* ‘the chair’		
9.*	*der Apfel* ‘the apple’	*der Keks* ‘the cookie’		
10.*	*das Buch* ‘the book’	*das Messer* ‘the knife’		

* used in practice trials

Combined presentations on eye-tracker screen (example)



55

Test sentences (*used in practice trials)

Nr.	Subject (SG / PL)	Verb (3SG / 3 PL)	Object 1	Objekt 2
1.	das Baby / die Babys	schüttelt / schütteln	eine Bürste	eine Flasche
	the baby / the babies	shakes / shake	a brush	a bottle
2.	der Bär / die Bären	futtert / futtern	eine Banane	eine Wurst
	the bear / the bears	eats / eat	a banana	a sausage
3.	der Hase / die Hasen	bastelt / basteln	einen Mond	einen Stern
	the rabbit / the rabbits	makes / make	a moon	a star
4.	der Hund / die Hunde	liefert / liefern	einen Ball	einen Schuh
	the dog / the dogs	delivers / deliver	a ball	a shoe
5.	das Pferd / die Pferde	knappert / knabbern	einen Käse	einen Kuchen
	the horse / the horses	nibbles / nibble (on)	a cheese	a cake
6.	das Schaf / die Schafe	öffnet / öffnen	eine Tür	eine Dose
	the sheep / the sheep	opens / open	a door	a can
7.	das Schwein / die Schweine	angelt / angeln	eine Hose	eine Mütze
	the pig / the pigs	fishes / fish	(a) pants	a hat
8.	der Vogel / die Vögel	zeichnet / zeichnen	einen Tisch	einen Stuhl
	the bird / the birds	paints / paint	a table	a chair
9.*	der Affe / die Affen	will / wollen	einen Keks	einen Apfel
	the monkey / the monkeys	wants / want	a cookie	an apple
10.*	das Kind / die Kinder	sucht / suchen	ein Buch	ein Messer
	the child / the children	looks for / look for	a book	a knife

A.3 Statistical analyses

This appendix provides further results of the statistical analyses performed in this thesis. Significant results (on the level of $p<0.05$) are reported in the result sections of the experiment chapters. Results that display an F-value higher than 1.0, but fail to reach significance are reported here. If no interaction or main effect is reported in either section, the corresponding F-value was lower than 1.0.

A.3.1 Experiment 1 (SV-Online)

Looking behaviour in Phase 1

- 2x2x4 ANOVA comparing Grammaticality, Subject number and Group
  - Main effect for Grammaticality: $F(1,69)=3.919$, $p=0.062$
  - Interaction Group*Grammaticality*Subject number: $F(3,69)=1.392$, $p>0.1$

Proportion of looks to target picture (3000 ms)

- 2x2 ANOVA comparing Grammaticality and Subject number in Adults
  - Main effect for Grammaticality: $F(1,24)=1.657$, $p>0.1$)
- 2x2x3 ANOVA comparing Grammaticality, Subject number and Group in Children
  - Interaction Group*Grammaticality*SubjectNumber: $F(2,43)=1.187$, $p>0.1$
  - Interaction Group*Grammaticality: $F(2,43)=1.118$, $p>0.1$
- post-hoc 2x2 ANOVA, 3-year olds:
  - Main effect of Subject number: $F(1,14)=2.573$, $p>0.1$
- post-hoc 2x2 ANOVA, 4-year olds:
  - Main effect of Subject number: $F(1,14)=2.205$, $p>0.1$
- post-hoc 2x2 ANOVA, 5-year olds:
  - Main effect of Grammaticality: $F(1,15)=1.995$, $p>0.1$
  - Main effect of Subject number: $F(1,15)=1.972$, $p>0.1$
- post hoc one-way ANOVA comparing Grammaticality in 4-year olds:
  - singular trials: $F(1,14)=2.424$, $p=>0.1$

Proportion of looks to target picture (1200 ms)

- 2x2x3 ANOVA comparing Grammaticality, Subject number and Group in Children
  - Interaction Group*Grammaticality: $F(1,40)=1.118$, $p>0.1$

Proportion of looks to subject picture (3000 ms)

- 2x2 ANOVA comparing Grammaticality and Subject number in Adults
  - Main effect for Grammaticality: $F(1,24)=2.864$, $p>0.1$)
- 2x2x3 ANOVA comparing Grammaticality, Subject number and Group in Children
  - Interaction Grammaticality*Subject number: $F(1,43)=1.025$, $p>0.1$
  - Interaction Group*Subject number: $F(2,43)=1.629$, $p>0.1$
  - Main effect of Subject number: $F(1,43)=2.786$, $p>0.1$
- post-hoc 2x2 ANOVA, 3-year olds:
  - Main effect of Grammaticality: $F(1,14)=1.859$, $p>0.1$
- post-hoc 2x2 ANOVA, 4-year olds:
  - Main effect of Subject number: $F(1,13)=1.598$, $p>0.1$
- post-hoc 2x2 ANOVA, 5-year olds:
  - Main effect of Subject number: $F(1,13)=1.41$, $p>0.1$

Proportion of looks to subject picture (1200 ms)

- 2x2 ANOVA comparing Grammaticality and Subject number in Adults
  - Main effect for Grammaticality: $F(1,24)=2.867$, $p>0.1$)
  - Interaction Grammaticality*Subject number: $F(1,24)=2.449$, $p>0.1$
- 2x2x3 ANOVA comparing Grammaticality, Subject number and Group in Children
  - Main effect for Grammaticality: $F(1,40)=1.162$, $p>0.1$)

Fixation durations on target picture

- 2x2 ANOVA comparing Grammaticality and Subject number in Adults
  - Main effect of Grammaticality: $F(1,20)=1.036$, $p>0.1$
  - Interaction Grammaticality*Subject number: $F(1,20)=1.558$, $p>$
- 2x2x3 ANOVA comparing Grammaticality, Subject number and Group in 3-year olds
  - Main effect of Grammaticality: $F(1,28)=2.177$, $p>0.1$
  - Interaction Group*Grammaticality: $F(2,28)=1.698$, $p>0.1$
  - Interaction Grammaticality*Subject number: $F(2,28)=1.437$, $p>0.1$
  - Interaction Group*Grammaticality*Subject number: $F(2,28)=2.181$, $p>0.1$

Fixation durations on subject picture

- 2x2x3 ANOVA comparing Grammaticality, Subject number and Group in 3-year olds
  - Main effect of Grammaticality: $F(1,29)=2.502$, $p>0.1$
  - Interaction Group*Subject number: $F(2,29)=1.784$, $p>0.1$
  - Interaction Group*Grammaticality*Subject number: $F(2,29)=1.331$, $p>0.1$

First fixation latency (target picture)

- 2x2 ANOVA comparing Grammaticality and Subject number in Adults
  - Main effect of Grammaticality: $F(1,20)=2.739$, $p>0.1$

A.3.2 Experiment 2 (SV-Pron)

Proportion of looks to target picture

- 2x2 ANOVA comparing Phase and Number condition in Adults
  - Interaction Phase*Number: $F(1,16)=1.526$, $p>0.1$
- 2x2 ANOVA comparing Phase and Number condition in 3-year olds
  - Main effect for Phase: $F(1,27)=1.646$, $p>0.1$

Fixation durations

- 1-actor picture, 2x2 ANOVA comparing Phase and Number condition in Adults
  - Main effect for Phase: $F(1,16)=2.22$, $p>0.1$
- 1-actor picture, 2x2 ANOVA comparing Phase and Number condition in 3-year olds
  - Interaction Phase*Number: $F(1,27)=2.553$, $p>0.1$
  - post-hoc singular trials: ME Phase $F(1,27)=1.752$, $p>0.1$
  - post-hoc plural trials: ME Phase $F(1,27)=1.614$, $p>0.1$

A.3.3 Experiment 3 (SV-Point)

Proportion of looks to target picture

- 2x2 ANOVA comparing Phase and Number condition in 3-year olds
  - Interaction Phase*Number: $F(1,27)=1.452$, $p>0.1$

Fixation duration

- 1-actor picture, 2x2 ANOVA comparing Phase and Number condition in 3-year olds
  - Main effect Phase: $F(1,25)=1.786$, $p>0.1$
  - Main effect Number: $F(1,25)=1.321$, $p>0.1$

Correctly pointed trials, fixation duration

- 1-actor picture, 2x2 ANOVA comparing Phase and Number condition in 3-year olds
  - Main effect Phase: $F(1,21)=2,736$, $p>0.1$
- 2-actor picture, 2x2 ANOVA comparing Phase and Number condition in 3-year olds
  - Main effect for Number: $F(1,21)=1.775$, $p>0.1$

Incorrectly pointed trials, fixation duration

- 1-actor picture, 2x2 ANOVA comparing Phase and Number condition in 3-year olds
  - Main effect Phase: $F(1,18)=1.511$, $p>0.1$
  - Main effect Number: $F(1,18)=1.263$, $p>0.1$

A.3.4 Experiment 4 (SV-PronOnline)

Proportion of looks to target picture

- 2x2 ANOVA comparing Phase and Number condition in Adults
  - Interaction Phase*Number: $F(1,19)=2.184$, $p>0.1$
- 2x2 ANOVA comparing Phase and Number condition in 3-year olds
  - Main effect for Phase: $F(1,29)=1.38$, $p>0.1$

Fixation duration

- 2-actor picture, 2x2 ANOVA comparing Phase and Number condition in Adults
  - Main effect Phase: $F(1,13)=3.417$, $p=0.087$
- 2-actor picture, 2x2 ANOVA comparing Phase and Number condition in 3-year olds
  - Main effect Phase: $F(1,25)=3.771$, $p=0.064$

Correctly pointed trials

- 2-actor picture, 2x2 ANOVA comparing Phase and Number condition in 3-year olds
  - Main effect Phase: $F(1,17)=2.168$, $p>0.1$
  - Main effect Number: $F(1,17)=1.093$, $p>0.1$

Incorrectly pointed trials

- 1-actor picture, 2x2 ANOVA comparing Phase and Number condition in 3-year olds
  - Main effect Phase: $F(1,16)=3.849$, $p=0.067$
  - Main effect Number: $F(1,16)=3.822$, $p=0.068$
- 2-actor picture, 2x2 ANOVA comparing Phase and Number condition in 3-year olds
  - Main effect Phase: $F(1,16)=2.642$, $p>0.1$
  - Main effect Number: $F(1,16)=1{,}436$, $p>0.1$

Spektrum Patholinguistik

herausgegeben vom Verband für Patholinguistik (vpl) e. V.

Zuletzt erschienene Ausgaben:

Band 7 Schwerpunktthema: Hören – Zuhören – Dazugehören: Sprachtherapie bei Hörstörungen und Cochlea-Implantat | 2014
ISBN 978-3-86956-294-0
URN urn:nbn:de:kobv:517-opus-70629

Band 6 Schwerpunktthema: Labyrinth Grammatik: Therapie von syntaktischen Störungen bei Kindern und Erwachsenen | 2013
ISBN 978-3-86956-270-4
URN urn:nbn:de:kobv:517-opus-67659

Band 5 Schwerpunktthema: Schluck für Schluck: Dysphagietherapie bei Kindern und Erwachsenen | 2012
ISBN 978-3-86956-199-8
URN urn:nbn:de:kobv:517-opus-59877

Band 4 Schwerpunktthema: Lesen lernen: Diagnostik und Therapie bei Störungen des Leseerwerbs | 2011
ISBN 978-3-86956-145-5
URN urn:nbn:de:kobv:517-opus-53146

Band 3 Schwerpunktthema: Von der Programmierung zur Artikulation: Sprechapraxie bei Kindern und Erwachsenen | 2010
ISBN 978-3-86956-079-3
URN urn:nbn:de:kobv:517-opus-46134

Band 2 Schwerpunktthema: Ein Kopf - Zwei Sprachen: Mehrsprachigkeit in Forschung und Therapie | 2009
ISBN 978-3-940793-89-8
URN urn:nbn:de:kobv:517-opus-30451

Spektrum Patholinguistik – Schriften

herausgegeben vom Verband für Patholinguistik (vpl) e. V.

Zuletzt erschienene Ausgaben:

Band 8 Heike Herrmann: Zum Erwerb syntaktischer Aspekte von positiven und negativen W-Fragen im unauffälligen und auffälligen Spracherwerb des Deutschen | 2014
ISBN 978-3-86956-293-3
URN urn:nbn:de:kobv:517-opus-70606

Band 7 Schultheiss, Corinna: Die Bewertung der pharyngalen Schluckphase mittels Bioimpedanz : Evaluation eines Mess- und Diagnostikverfahrens | 2014
ISBN 978-3-86956-284-1
URN urn:nbn:de:kobv:517-opus-69589

Band 6 Oda-Christina Brandt-Kobele: Comprehension of verb inflection in German-speaking children | 2014
ISBN 978-3-86956-216-2
URN urn:nbn:de:kobv:517-opus-62046

Band 5 Anja Bethmann: Die supramodale Verarbeitung individueller Konzepte am Beispiel menschlicher Stimmen und visuell präsentierter Comicfiguren : eine fMRT-Studie der Temporallappen | 2012
ISBN 978-3-86956-206-3
URN urn:nbn:de:kobv:517-opus-59856

Band 4 Susan Ott: Feld - fällt - fehlt : Untersuchungen zur Phonologie-Morphosyntax-Schnittstelle bei Kindern und Erwachsenen | 2012
ISBN 978-3-86956-161-5
URN urn:nbn:de:kobv:517-opus-57798

Band 3 Birgit Herold: Prosodische Verarbeitung und lexikalische Entwicklung sehr untergewichtiger Frühgeborener während des ersten Lebensjahres | 2011
ISBN 978-3-86956-107-3
URN urn:nbn:de:kobv:517-opus-48517

One standard assumption regarding child language development is that comprehension of a linguistic structure either precedes or temporarily coincides with production of that particular structure. Studies on the acquisition of verb inflection morphology have pointed to a reverse order indicating production preceding comprehension. The present work provides a thorough and comprehensive overview of the acquisition of number inflection in children as well as innovative research on the processing of verb inflection morphology. In a series of experiments recording children's eye movements the comprehension and processing of inflectional number information in German-speaking children has been investigated. Results indicate that the comprehension does in fact not lag behind production. The results further illustrate the critical impact of methodological considerations on language acquisition theories.

Oda-Christina Brandt-Kobele studied Patholingusitics at the University of Potsdam. The present study was conducted in the Language Acquistion Lab at the Linguistics Department at the University of Potsdam.

ISSN 1869-3822
ISBN 978-3-86956-216-2
9 783869 562162