Lingua 120 (2010) 1910–1925

Contents lists available at ScienceDirect

Lingua

journal homepage: www.elsevier.com/locate/lingua

What asymmetries within comprehension reveal about asymmetries between comprehension and production: The case of verb inflection in language acquisition

Oda-Christina Brandt-Kobele *, Barbara Höhle

University of Potsdam, Department of Linguistics, Karl-Liebknecht-Straße 24-25, 14476 Potsdam-Golm, Germany

ARTICLE INFO

Article history:
Received 2 July 2009
Received in revised form 11 February 2010
Accepted 11 February 2010
Available online 10 April 2010

Keywords:
Verb inflection
Acquisition of German
Production-comprehension asymmetry
Eye tracking

ABSTRACT

Two recent studies (Johnson et al., 2005; Perez-Leroux, 2006) found that English- and Spanish-learning children do not show the ability to use verb inflection as a cue to subject number before the age of 5 to 6 years. These findings suggest an asymmetric development as verb inflections are usually correctly produced before this age.

In the present study we investigated whether German 3- to 4-year-olds take advantage of the information provided by the verb inflection in sentence comprehension. In a first study, children's looking behavior at two pictures was measured after presentation of a sentence in which the subject number was coded only by the verb inflection. The results from this study suggest that children's looks reflect correct interpretation of the sentences and thus show their ability to make use of verb inflection. In a second experiment, preferential looking was combined with an additional task in which the children had to point to the matching picture. In this case children did not perform above chance level.

Our results underline the relevance that specific task demands have on the performance of children in comprehension testing. These have to be accounted for when interpreting findings on production and comprehension asymmetries in language acquisition.

1. Introduction

Children's acquisition of verb inflection has been studied in various contexts and languages, employing a whole range of different paradigms. Much attention has been given to the acquisition of morphemes that determine finiteness of a sentence (e.g. Wexler, 1998; Poeppel and Wexler, 1993; Guasti, 2002). A further role of inflectional elements is to display properties of syntactic dependencies that the inflected element enters into (Nichols, 1986). The English 3rd person singular *-s*, for example, conveys information about the person and number of the sentence subject and is thus a marker of agreement between subject and verb. A natural question to pose when investigating the acquisition of inflectional elements is to ask when children are able to extract the information conveyed by such inflectional markings.

In trying to answer this question, we will present new evidence that German children aged 3 to 4 years are able to extract the number information conveyed by verb inflections and use this information when interpreting sentences in which the subjects are ambiguous with regard to number. This was established using an eye-tracking procedure employing a preferential looking paradigm. The results of our comprehension task provide new information about the relationship

* Corresponding author. Tel.: +49 331 977 2296; fax: +49 331 977 2095.
E-mail address: odabrandt@googlemail.com (O.-C. Brandt-Kobele).

0024-3841/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.lingua.2010.02.008

between children's comprehension and production abilities concerning verb inflections that mark number agreement. Earlier findings on children's comprehension of verb inflections have suggested an asymmetric relationship between the receptive and the productive domains, with productive abilities preceding correct comprehension (e.g. Johnson et al., 2005). This paper is structured as follows. First, we will present earlier findings on children's processing of verb inflections encoding the number contrast in production and comprehension, which altogether yielded in the assumption of an asymmetric developmental path. Then we will present two studies on the comprehension of number information provided by verb inflection in German children, thereby comparing different methodological approaches (eye tracking vs. picture selection). As our results indicate a great impact of methodology on the comprehension abilities found in children, we will discuss this matter in the context of the assumed production-comprehension asymmetry as well as possible causes of the discrepancies across methods.

The majority of studies on the acquisition of verb inflections conducted so far examine children's use of such morphemes in the productive modality. For this, different methods have been employed, for example investigating children's spontaneous speech at various ages (e.g. Brown, 1973; Clahsen, 1986) or doing elicited production tasks (e.g. Rice and Wexler, 2002). The acquisition path of verb inflection revealed by such studies differs cross-linguistically (Bittner et al., 2003). In some languages, so-called "Optional Infinitive" languages such as English, Dutch, German and French, children start out producing uninflected or non-finite forms (e.g. 'Daddy walk') as well as inflected and thus finite verb forms (e.g. 'He is big'). In the second and third year of life, the productive use of inflectional morphology increases in its frequency in obligatory contexts of adult usage (e.g. Guasti, 2002). In other languages, so-called pro-drop languages, such as Italian, Spanish and Turkish which have rich morphological systems and allow the overt subject to be unrealized, infinitival forms are extremely rare in children's early speech. Thus, children acquiring such languages start out producing inflected verb forms already at the age of 1 and a half or 2 (Guasti, 1993). To compare children's receptive and productive capacities with respect to verb inflection that marks number agreement, two 'Optional Infinitive' languages, namely English and German, are focused on in this paper. Additionally, we will consider data from the pro-drop language Spanish to rule out the hypothesis that the late comprehension found in English children is only due to the impoverished morphological system.

In his longitudinal study of the spontaneous speech of three English-speaking children, Brown (1973) found production of the 3rd person singular *-s* in 90% of obligatory contexts between the ages 2;2 and 3;10, already indicating substantial variance across children even within one language. Using elicited production data, Rice and Wexler (2002) found that English learners on a group level do not reach the 90% criterion of correct use of the 3rd person singular *-s* before the age of 4 years. One important difference between spontaneous speech and elicited production data, besides diverging task demands, lies in verb selection. Assuming verb frequency and familiarity to have an impact on children's ability to use an inflected form correctly (e.g. Aguado-Orea, 2004), rather than assuming lack of knowledge of the appropriate inflection (Brown, 1973) or the existence of a stage in development in which abstract grammatical features may be under-specified, as the Optional Infinitive Model does (Wexler, 1994), verb selection itself might be a factor explaining different ages of mastery found with different methods: if children spontaneously produce familiar and high frequent verbs, the chance of errors in spontaneous speech is lower compared to elicited production.

So far, studies about the productive onset and mastery of inflections by German-speaking children are sparse and based solely on spontaneous speech data. 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) has looked at 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 morphemes at various ages. Correct use, again in Brown's sense of more than 90% correct, for the 3rd person singular inflectional morpheme *-t* did not occur before the age of 2;11. For the 3rd person plural inflection *-n*, correct use emerged even a little later, namely at around 3;1. In a similar study, Poeppel and Wexler (1993) analyzed the spontaneous speech of one child aged 2;1. They found that in all utterances that contained a 3rd person singular subject and a finite verb the verb was correctly inflected with the *-t* marker. This lead the authors to conclude that the agreement system is basically available at age 2 (see Bittner, 2003 for further spontaneous speech data on German inflections). The discrepancies between Poeppel and Wexler's findings 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 be kept in mind that findings from spontaneous speech might overestimate children's production abilities to inflect verbs correctly. Still, we conclude that the possibly youngest age of productive mastery of German inflection is between 2 and 3 years.

Receptive abilities for inflectional morphemes in children have been tested in only a few studies so far. Using the head-turn preference paradigm,[1] Soderstrom et al. (2002) examined the early sensitivity of English-learning 19-month olds to the presence of the 3rd person singular *-s* in sentences containing singular subjects. In line with other studies reporting early sensitivity to syntactic well-formedness of structures which young children do not produce themselves (e.g. Santelmann and Jusczyk, 1998; Höhle et al., 2006), Soderstrom and colleagues found that the infants preferred the grammatical (3rd person singular *-s* present) over the ungrammatical passages (3rd person singular *-s* absent). However, the properties to which the children in such head-turn preference experiments are sensitive are unclear. It is possible that the children actually process

[1] Note that in head-turn preference studies no actual comprehension but only sensitivity to a verbal form is assessed, as only auditory stimuli are used and no visual material is presented. See, for example, Juscyzk (1996) for a detailed explanation of the method.

the number information on the subject and verb and are additionally able to check the matching of these morpho-syntactic features on a grammatical basis (Soderstrom, 2008), which could be viewed as a hint to early comprehension. In contrast, children's preference for grammatical structures might be solely based on their knowledge about distributional properties, i.e. surface properties of the input. Thus, based on these head-turn preference findings it is hard to argue that young children are actually attuned to the manifestation of correct subject-verb agreement and that they are able to infer the semantic implications of the functional morphemes involved in this agreement relation.

A more direct approach to the question of when children are sensitive to the different morpho-syntactic categories involved in verb inflection is rendered by studies that investigated whether children can make use of the information provided by the inflectional morphemes for the identification of subject number. Johnson et al. (2005) conducted a picture-selection task in which the verb inflection was the only cue to subject number. To achieve this, they used verbs that began with an *-s* consonant cluster which was coarticulated with the plural *-s* on the noun (*The duck swims on the pond* vs. *The ducks swim on the pond*). 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. Just the 5- and 6-year-old children but not the younger ones performed above chance level. But even the older children performed well below 100% correct across all conditions. In the plural condition even the 6-year-olds did not continue to display above chance-level performance. Johnson et al. (2005) concluded that, especially for the 3- and 4-year-old children, the English 3rd person singular inflection *-s* is not a transparent marker for subject number agreement. They hypothesized that the "poverty" of the English present tense agreement system, in which the 3rd person singular is the only morphological form overtly marked, may be relevant for the children's ignorance concerning the information provided by the inflectional ending.

This assumption was tested in a study conducted by Perez-Leroux (2005) with Spanish-speaking children aged 3 to 6. In contrast to English, Spanish has a richer set of verb inflections, with specific inflectional endings for all person and number forms of a verb. Perez-Leroux adopted the material and the procedure used by Johnson et al. (2005). She 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 (1).

1	a.		Nada	en el charco.
		(The duck-SG)	swim-3SG	on the pond
		The duck	*swims*	*on the pond.*
	b.		Nadan	en el charco.
		(The duck-PL)	swim-3PL	on the pond
		The ducks	*swim*	*on the pond*

Despite the morpho-syntactic differences across these two languages, the data from the Spanish-speaking children were remarkably similar to those from the English-speaking children tested by Johnson et al. While 3 and 4-year-old children's responses did not differ from chance performance in either number condition, the 5 and 6-year-old children reached performance levels better than chance (but only in the plural condition).

When comparing the findings from the production and the sentence comprehension studies, a puzzling picture emerges. The production studies suggest that latest by age 4 English children have mastered the subject-verb-agreement system, indicating that they can process the number information 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.

A 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 et al. (1963) examined the singular-plural distinction as marked by inflection or by the auxiliaries *is* and *are* (e.g. *The deer is running* vs. *The deer are running*). To avoid doubly marking the number information only nouns without overt plural marking like *sheep* or *deer* were used. They tested 3;1- to 3;7-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 asking the children to name the same pictures. Comprehension was found to be ahead of production such 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. Thus, the findings by Fraser et al. (1963) cannot be considered to have revealed an asynchronous relationship between children's production and comprehension of grammatical morphemes.

The question remains as to whether the production of correct number agreement between subject and verb in fact precedes the comprehension of number information given by verb inflection in sentence interpretation, and thus whether production and comprehension of verb inflection is a further area of language acquisition in which the development of comprehension seems to lag behind the development of productive skills. Such a counterintuitive pattern has been found in the acquisition path of other linguistic structures, namely English pronouns (e.g. Hendriks and Spenader, 2006; Sekerina et al., 2004), the German focus particle *auch* ('also') (e.g. Hüttner et al., 2004; Höhle et al., 2009) and restrictive noun phrase modifiers (e.g. Hurewitz et al., 2000).

The comprehension studies on verb inflection cited above all employed some variant of the picture-selection task, which requires explicit decisions to be made by the children tested. This places demands on non-linguistic cognitive skills and one may hypothesize that these cognitive skills are not fully developed in children and are thus hindering them from revealing their actual language comprehension abilities (e.g. Höhle et al., 2009). When searching for a method for assessing children's early interpretation of grammatical morphemes that puts only low task demands on children and does not require an explicit choice among a set of pictures, preferential looking comes into play. This method has been widely used to assess young children's lexical and syntactic comprehension abilities (e.g. Hirsh-Pasek and Golinkoff, 1996a; Naigles, 2002; Golinkoff et al., 1987). While children are presented with two visual stimuli and a linguistic stimulus describing one of these, they typically fixate the matching visual stimulus.

Based on these findings and assumptions, the aims of our study are threefold. First, we want to examine whether German-learning children aged 3 to 4 years show comprehension of verb inflection, such that they can infer the number of a sentential subject from the verbal inflection alone. Second, we want to compare the impact of different methodological approaches on children's receptive morpho-syntactic abilities by comparing the outcomes of experiments using different methods. Finally, we want to add information to what is known about the relationship between early production and comprehension of verb inflection in children.

In our first experiment children were presented with sentences in which the verb inflection was the only cue to the number of the subject. Each sentence was presented with a picture pair that depicted the action described in the sentence either conducted by one or by two actors. Thus, the materials used were similar to those of the previous studies in English and Spanish. But instead of using a picture selection task children's eye gaze at these two pictures was tracked before and after the presentation of the stimulus sentences. If German 3- to 4-year-old children were able to infer the correct subject number via the processing of inflectional information, we expected longer looking times at a picture depicting one actor after hearing a singular inflected verb and longer looking times at a picture showing two actors after hearing a plural-inflected verb.

2. Experiment 1

2.1. Methods

2.1.1. 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 this study.

The parents were asked to fill out a short questionnaire to enable detection of any language deficits in their children. This 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 (1st person singular to 3rd person plural, all present tense regular inflection, e.g. *ich gehe* 'I go', *du gehst* 'you go', *er geht* 'he goes', etc.) were provided and the parents had to mark the forms their child had already produced. These data should provide us with a rough estimation of our participant's production of verb inflection and thus extend beyond the spontaneous speech data reported in the literature (e.g. Clahsen, 1986). According to parental reports, 77% of the children produce 3rd person singular inflections, and 51% of the children produce 3rd person plural inflections. We conclude from these answers that clearly not all children participating in our study are already producing verb inflections.

2.1.2. Material and design

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 3rd person singular female (*sie*) and 3rd person plural (*sie*) are homophones, making the sentence temporarily ambiguous until the inflection marker of the verb has been parsed. The verbs were either inflected for 3rd person singular (*-t*) or 3rd person plural (*-n*), providing the number variation in the experimental design (2).[2]

2	a.	Sie	fütter-t	einen Hund.
		Pronoun-3SG	feed-3SG	a dog
		She	*is feeding*	*a dog.*
	b.	Sie	fütter-n	einen Hund.
		Pronoun-3PL	feed-3PL	a dog
		They	*are feeding*	*a dog.*

[2] 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 adult. We consider it as very unlikely that young children are confused by this ambiguity, as the politeness form is very rarely used in the speech directed to children and children start producing it late (some not even in their elementary school years). Since the politeness form is 2. Pers. Sg. and thus not felicitous when describing pictures, no interference is expected.

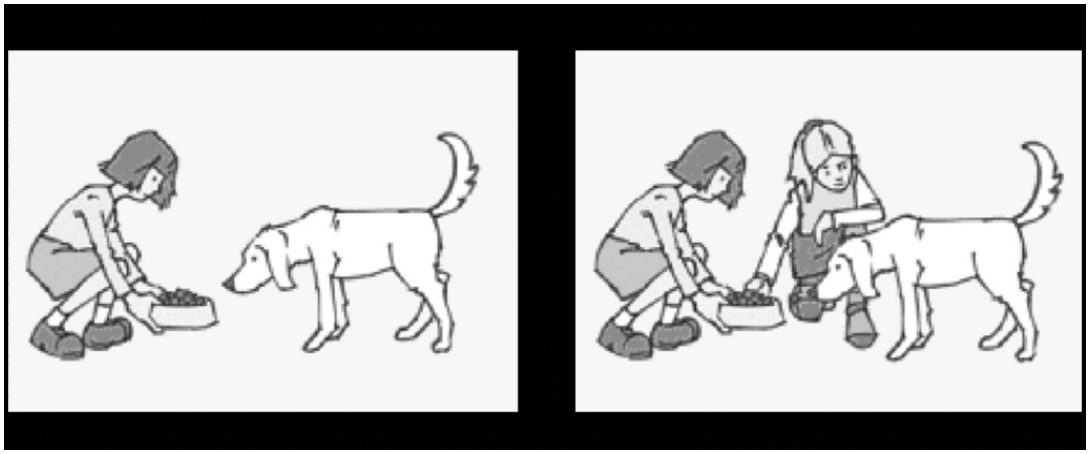


Fig. 1. Visual material used in Experiment 1 and 2. 1-Actor picture (left) and 2-actor picture (right). Example: *Sie füttert einen Hund/Sie füttern einen Hund*. 'She is feeding a dog'/'They are feeding a dog'.

Four different verbs were used, which were disyllabic when inflected for either number, depictable, and known to young German children.[3] Each verb was combined with two different objects each to create eight experimental sentences. The object NPs always contained an indefinite article (*einen*, *eine*, 'a'). The verbs and objects used are provided in Appendix A.

An additional verb again combined with two different objects was used for practice. These combinations were presented in both number conditions, yielding four practice trials. All verbal stimuli were recorded by a female native speaker of German in a child-directed manner. Mean length of sentences was 1581 ms (range: 1392–1832 ms).

For each sentence a simple coloured drawing of the described situation was made up. The two pictures 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 condition, only one person was performing the action mentioned in the sentence (e.g. one girl feeding a dog (6a)). In the 2-actor condition, two persons were performing the same action together (e.g. two girls feeding a dog together (6b)). See Fig. 1. To conform to the female pronoun in the singular condition, all depicted characters were girls. The experimental sentences are provided in Appendix B.

Pictures were presented pairwise in the experiment. Each pair consisted of the two pictures showing the same action with the same object, thus the pictures of each pair differed only with respect to the number of the actors depicted. Each child 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.[4] 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 in a row. Each experimental session contained four practice trials without feedback, and eight experimental trials, four for each of the number conditions.

In each trial, one pair of pictures was presented side-by-side on the eye-tracking monitor for 3 s, accompanied by an attention getting phrase (*Schau mal*!–'Look here!') to direct the child's attention to the screen (baseline phase). The baseline phase was included to control for initial picture preferences or biases. After that, the screen turned black for 2 s, during which 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 testing 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 3 s (testing phase), which was again followed by a black screen for 1 s. Then, the trial ended automatically. Thus, a trial had a duration of 10 s. The inter-trial interval lasted about 2 s, during which the screen was blank. See Fig. 2 for a schematic description of a trial course.

After four experimental trials, a short clip (e.g. Elmo jumping up and down) was presented to redirect the child's attention to the screen. The children's eyes were recorded throughout the whole experiment, which lasted about 3 min.

To measure children's eye gaze, a tabletop Tobii 1750 eye-tracking system (Tobii Technology AB, Sweden) was used, which tracks eye position every 20 ms with a resolution of 50 Hz. 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.

[3] 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.

[4] In principle, a 3rd person singular sentence matches the 2-actor picture as well as the 1-actor picture (since in a set of people, a single person is always included). To ensure that the 2-actor picture is a less felicitous match for a singular inflected verb, we ensured that the action is always performed by the two depicted girls together.

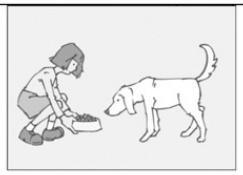
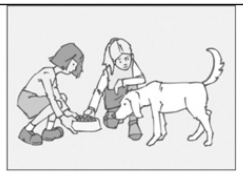
Trial Phase	Duration	Auditory stimuli	Left side of monitor	Right side of monitor
Baseline Phase	3 s	*Schau mal* 'Look at that'		
Sentence Presentation	2 s	*Sie füttert einen Hund.* 'She is feeding a dog'		
Testing Phase	3 s			
	1 s			

Fig. 2. A schematic example of the trial procedure.

The children were sitting in the reclined chair, at approximately a 60 cm distance from the monitor, watching the pictures and listening passively to the sentences. They were only instructed to sit still and listen to the sentences as well as to watch the screen. The parent was sitting behind the child in the corner of the testing 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 child was sitting comfortably and was attentive, a 5-point calibration procedure was performed. Before starting the experiment, the child was presented with a short story accompanied by an introductory picture (three girls standing side by side). The story alerted the child that she 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 C. After this, the four practice trials were presented to acquaint the child with the experimental procedure. If necessary, some children were reminded to watch the screen or not to talk in-between. After the practice trials, the eight testing trials were presented.

2.1.3. Data analysis

Eye movements were analyzed automatically by the ClearView-Software employing the standard setting. The output data we used for further analysis was one text file per participant providing information about the specific time course of the experiment (e.g. onset of each trial), the accurate position of the eye gaze (as *X*–*Y*-coordinates) at each time point, as well as number and duration of fixations according to ClearView default settings.

To analyze the looking behavior in relation to the verbal and visual stimuli presented, we defined two spatial areas of interest (AoI). Each AoI was 400 × 286 pixels in size, corresponding to the size of each of the pictures presented on the monitor. In addition, two temporal regions of interest (RoI) were defined, one corresponding to the baseline and the other to the testing phase.

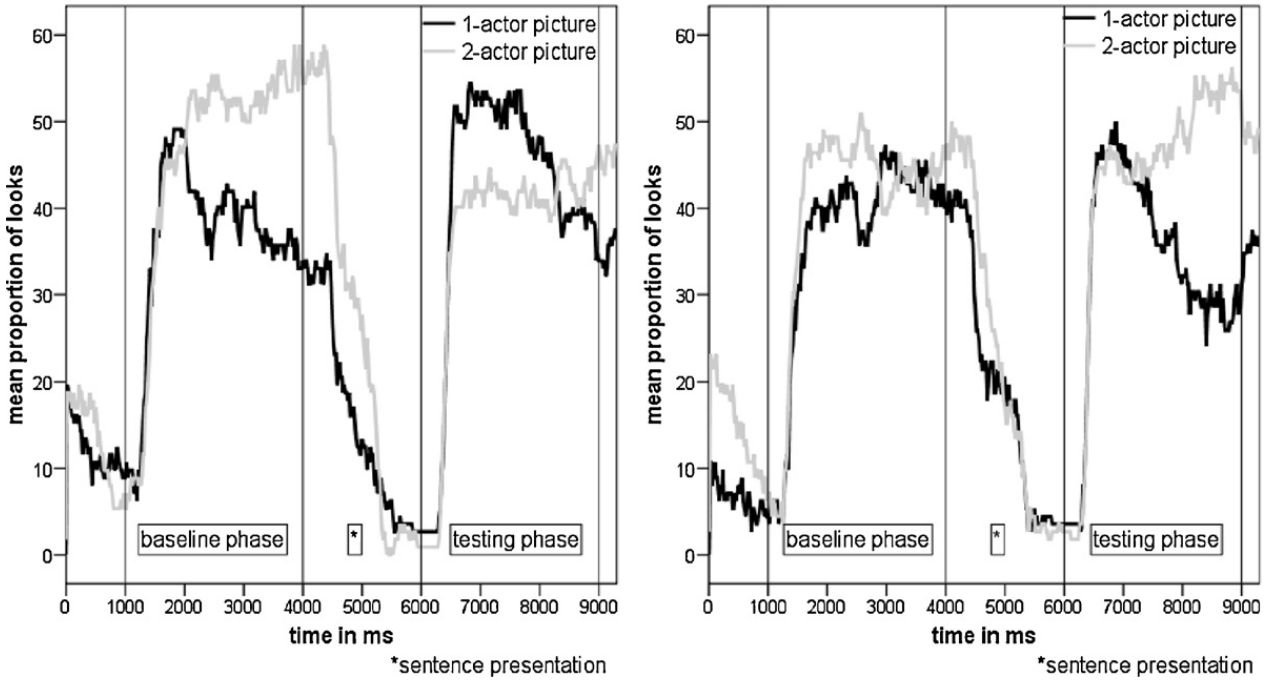


Fig. 3. (a) Children's proportions of looks in the singular number condition over course of a trial (Experiment 1). (b) Children's proportion of looks in the plural number condition over the course of a trial (Experiment 1).

The looking behavior was analyzed in two different ways: First, the looks to the target and distractor picture for every given time-stamp (every 20 ms) were averaged over trials and participants in each number condition for the purpose of a first rough inspection of children's looking behavior (see *Trial Course Analysis*). Second, a fixation duration analysis was performed, in which only fixations with a minimum duration of 100 ms on a maximal radius of 30 pixels were used. For each participant, the fixation durations to the target and the distractor picture were added within and over trials separately for the baseline and the testing phase as a function of the number condition. These summed fixations were averaged over participants. In the following, the children's summed fixation durations are labelled as looking time. Mean looking times entered into the statistical analysis, as they were subjected to a $2 \times 2 \times 2$ ANOVA, with Number and Picture Type and Phase as repeated factors (see *Fixation Duration Analysis*).[5]

2.2. Results and discussion

2.2.1. Trial course analysis

For a first visual inspection of children's looking behavior over the course of a trial, looks to 1-actor and 2-actor picture are plotted separately as a function of number condition (Fig. 3). Presented is the mean proportion of looks to the 1-actor and 2-actor picture during the course of the trial – at any given time-stamp – averaged over the four experimental trials in the singular number condition (3a), and in the plural number condition (3b). These proportions of looks are averaged over the 28 participants.[6]

The figure depicting the time course shows that children looked at the pictures in the baseline and in the testing phase. While the test sentence is presented (4–6 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 testing phase, the proportions of looks to the 1-actor and 2-actor picture seem to vary as a function of the number condition (Fig. 3a and b). While there appears to be a higher proportion of looks to the 1-actor picture in the singular number condition (Fig. 3a), the proportion of looks to the 2-actor picture seems to outrank those to the 1-actor picture in the plural number condition (Fig. 3b). To test whether children's looks to the 1-actor and 2-actor picture differed reliably as a function of number condition, further eye-tracking data analyses as well as a statistical analysis were performed.

[5] Note that all trials presented to the children entered analysis. In only a few trials however (or even just the baseline or testing phase of a given trial), children fixated neither of the two areas of interest (but rather looked elsewhere). In this experiment, one child did not fixate either picture in the baseline phase of 2 trials and in the testing phase of 2 other trials. Another child did not fixate either picture in both phases in 2 trials. These trials or phases still entered the fixation duration analysis, but with a duration value of 0 ms. In such cases we do not know whether the child was looking off-screen, on-screen but not at the AoIs, or whether the eye tracker simply failed to track the child's eye gaze, despite it being at the AoIs. Because of this uncertainty regarding the underlying reason for the 'missing data', we included all trials, accepting that some trials delivered a duration value of 0 ms.

[6] 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 proportion of looks directed at either picture a or b in Fig. 2 is less than 100%.

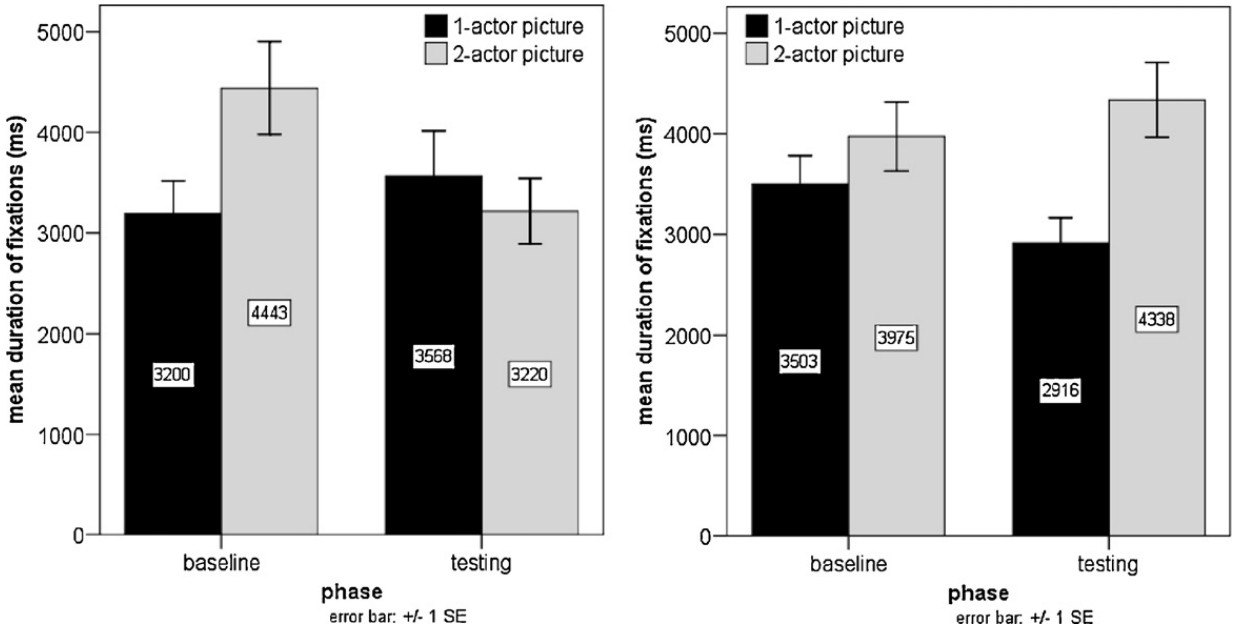


Fig. 4. (a) Children's mean looking times in the singular number condition (Experiment 1). (b) Children's mean looking times in the plural number condition (Experiment 1).

2.2.2. Fixation duration analysis

For the statistical analysis, the looking times to the 1-actor and the 2-actor pictures during the baseline and the testing phase were added for each participant. This looking time was then averaged across participants, separately for the singular and plural number conditions. The mean looking time for the singular number condition during the baseline and the testing phase is plotted in Fig. 4a, while the mean looking time for the plural number condition is plotted in Fig. 4b.

Only F1-analyses were performed on account of the low number of experimental trials in the experiment. A $2 \times 2 \times 2$ ANOVA comparing Phase (baseline vs. testing), Number (singular vs. plural) and Picture Type (1-actor vs. 2-actor) revealed a significant three-way interaction between these factors ($F(1,27) = 8.03$, $p < 0.01$). There were significant main effects for Phase ($F(1,27) = 6.39$, $p < 0.05$) and Picture Type ($F(1,27) = 5.63$, $p < 0.05$). No other effects reached significance, neither the interaction between Phase and Number ($F(1,27) = 1.19$, $p = 0.285$) nor the interaction between Number and Picture Type ($F(1,27) = 2.25$, $p = 0.145$). All other F's < 1.0.

In the following analysis, the data from the two number conditions were analyzed separately by 2×2 ANOVAs with Phase and Picture Type as independent variables. In the singular number condition, the interaction between Phase and Picture Type reached significance ($F(1,27) = 5.78$, $p < 0.05$) as well as the main effect for Phase ($F(1,27) = 6.93$, $p < 0.05$). The main effect for Picture Type did not reach significance ($F(1,27) < 1$). Separate analysis for the two phases revealed significantly longer looking time at the 2-actor picture than at the 1-actor picture in the baseline phase ($F(1,27) = 6.44$, $p < 0.05$). Such a significant difference in looking time at the 1-actor vs. 2-actor picture could not be found in the testing phase ($F(1,27) < 1$). Thus, children showed a preference for the 2-actor picture in the baseline phase, which vanished in the testing phase after hearing a sentence with a singular inflected verb.

In the plural number condition, only a marginal interaction between Phase and Picture Type was obtained ($F(1,27) = 3.84$, $p = 0.06$), but a strong main effect for picture emerged ($F(1,27) = 12.56$, $p < 0.01$). There was no main effect for Phase ($F(1,27) < 1$). Separate analysis for the two phases revealed a different pattern compared to the singular condition. While no significant differences in looking time to the two pictures were found in the baseline phase of the plural condition ($F(1,27) = 1.88$, $p = 0.181$) a reliable main effect for Picture Type emerged in the testing phase, with longer looking times at the 2-actor picture than at the 1-actor picture ($F(1,27) = 14.22$, $p < 0.01$).

In a further analysis we tested whether the presentation of the test sentences had a differential effect on children's preferences for the 2-actor picture as evidenced in the baseline phase. Thus we compared the looking times at the 2-actor picture in the baseline-phase with the looking times at the same picture in the testing-phase as a function of the grammatical number of the test sentence. This analysis showed that the children had a significant decrease in their looking times at the 2-actor picture from the baseline phase to the testing phase ($F(1,27) = 11.46$, $p < 0.01$) when presented with a singular sentence while no significant decrease in their looking times at the 2-actor-picture from the baseline to the testing phase occurred when a plural sentence was presented ($F(1,27) = 1.75$, $p = 0.197$).

2.2.3. Discussion

Summarizing our results, the following picture emerges. If the children were presented with a singular sentence, their general preference for looking at the 2-actor picture, which they had shown before hearing a sentence, vanished. This was

not the case when hearing a plural sentence. In this case, the children showed an increase in looking time to the 2-actor picture in the testing phase that led to longer fixation durations on the 2-actor picture compared to the one-actor picture.

This pattern suggests that – without a verbal stimulus that is systematically related to one of the pictures – the children tend to look 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 attracts a longer period of attention. The fact that we already found differences in the looking times at the two pictures 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 testing phase can reveal the effects of the sentence presented.

The finding that the children's looking times at the 2-actor picture decreased after the presentation of a singular inflected verb compared to the baseline suggests that the preference to look at the 2-actor picture has been overridden by the presentation of a sentence that correctly describes the 1-actor picture, and thus that the sentence presentation drives the children's 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 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.

A similar pattern of increased looking times at the target picture was not observed when a plural sentence was presented. But in this case, the target picture corresponded to the picture that the children already had looked at longer in the baseline phase. Nevertheless, the observation that there is at least a slight increase in looking time from the baseline to the testing phase and the finding that the 2-actor picture is fixated on longer than the 1-actor picture during the testing phase, indicate 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.

Summing up, we assume that our data provide evidence that 3- to 4-year-old German children are able to infer the number of an ambiguous sentential subject from the number information of the verbal inflection. This contrasts with the findings reported above for 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 5 to 6 years (Johnson et al., 2005; Perez-Leroux, 2005). One possible explanation for the earlier comprehension evidenced by German children could be the morpho-syntactic differences across these languages. While German has a rich system of verb inflection, in English main verbs are only marked for 3rd person singular and past tense. The greater relevance of inflectional endings in German may make German children acquire the verb inflections representing number differences earlier (see Guasti, 2002). 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 (Perez-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 picture-selection tasks, children in our Experiment 1 did not have to perform any particular action but were only instructed to watch the pictures and listen to the sentences. It may be hypothesized that this difference in task demands – merely looking in the eye tracking study vs. deciding and pointing in a picture-selection task – may account for the different findings. To test this possibility, a second 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. Children were presented with the exact same verbal and visual stimuli as in Experiment 1.

3. Experiment 2

3.1. Methods

3.1.1. Participants

Twenty-eight children participated in Experiment 2. Their mean age was 3;8 years (min: 3;2 years, max: 4;4 years, 15 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 as in Experiment 1 to control for any history of language deficits and gather information about children's productive use of verb inflections. According to the parental report, 72% of the children produce 3rd person singular inflections, and 48% of the children produce 3rd person plural inflections. From this we conclude that not all children in this group produce verb inflections already and additionally that this group does not differ from the one tested in Experiment 1 with respect to production (3rd person singular inflection: $chi^2(1) = 0.29$, $p = 0.593$; 3rd person plural inflection: $chi^2(1) = 0.38$, $p = 0.537$).

3.1.2. Material and design

The same design and material was used as in Experiment 1.

3.1.3. Procedure

In Experiment 2, children had to perform a picture-selection task while their eye gaze was tracked. Thus, the procedure varied slightly. Children were instructed to point to the picture which they thought would match the presented sentence best. To monitor the children's pointing reactions, the experimenter was sitting next to the child, and if necessary, and encouraged

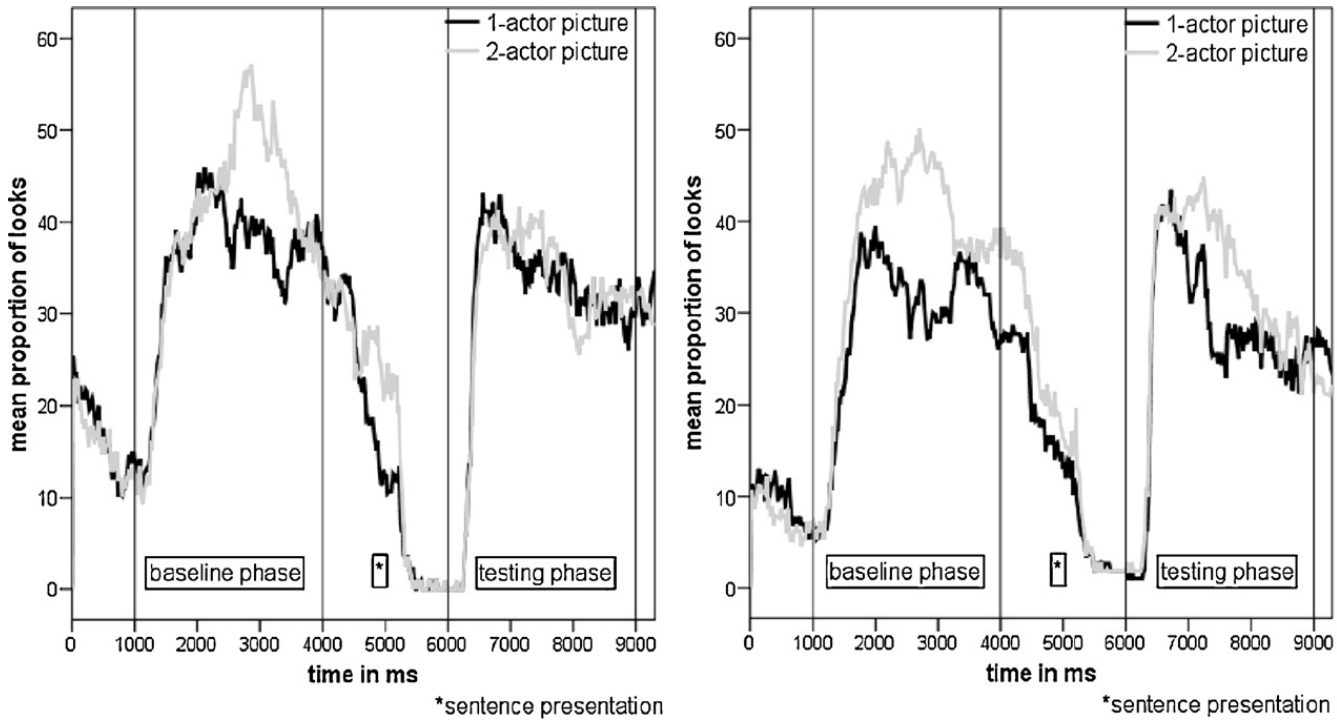


Fig. 5. (a) Children's proportions of looks in the singular number condition over course of a trial (Experiment 2). (b) Children's proportion of looks in the plural number condition over course of a trial (Experiment 2).

her to select a picture. 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 Experiment 1 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. In all further aspects the procedure was exactly as in Experiment 1. Each child was presented with four practice trials as well as eight experimental trials, four containing a singular inflected verb and four containing a plural inflected verb. Again, after every fourth trial, a short clip was presented. The same technical equipment as in Experiment 1 was used.

3.1.4. Data analysis

To keep data analysis comparable to Experiment 1 only the first three seconds of the testing phase of each experimental trial were considered in the analysis of the looking behavior. Thus, temporal RoIs as well as spatial AoIs were defined as in Experiment 1.

Again, the looking behavior of the children was analyzed in two different ways. First – for a visual inspection of the data – the mean proportion of looks to the 1-actor and the 2-actor pictures was calculated as a function of the number condition for each time-stamp over the whole trial course. Second – for statistical analyses – an analysis of the looking time was performed, with the same criteria defining looking time as in Experiment 1. The looking times at the 1-actor and the 2-actor pictures were added up per participant, separately for the singular and the plural number conditions, and averaged over participants. The mean looking time was subjected to a $2 \times 2 \times 2$ ANOVA, with Number, Picture Type and Phase as repeated factors.[7]

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 *t*-tests.

3.2. Results and discussion

3.2.1. Trial course analysis

Fig. 5 depicts the mean proportions of looks to the target and the distractor picture at any given time-stamp during the course of a trial. Fig. 5a depicts the looking behavior averaged over all trials and participants for the singular number condition, while Fig. 5b depicts the averaged looking behavior for the plural number condition.

[7] In this experiment, only one child failed to provide data points in all presented trials. In 2 trials, the child did not fixate one of the pictures in both phases, and in another trial, the child did not fixate either picture during the testing phase. Again, since we do not know whether the child was looking off-screen, on the screen but not on one of the AoIs or whether the eye tracker simply failed to track the child's eye gaze, these trials (or phases) enter analysis with a value of 0 ms.

The graphs in Fig. 5 show that children looked at the pictures in the baseline phase as well as in the testing phase. In the sentence presentation phase (4–6 s after trial onset), looks in the two spatial AoIs drop close to zero. This again can be considered as a consequence of the blank screen during that phase. In the baseline phase, children seem to look more at the 2-actor picture irrespective of the number condition. In the testing phase, however, no reliable difference seems to emerge between the looks at both picture types, either in the singular or in the plural number condition. To test whether this first impression holds, a statistical analysis of the eye-tracking data was performed.

3.2.2. Fixation duration analysis

As in Experiment 1, we analyzed the looking times at the target and the distractor pictures as a function of the number condition in the specified temporal RoIs. We followed the same rationale and the exact same way of analyzing the data as in Experiment 1. The mean looking time for the singular number condition is plotted in Fig. 6a and the looking time for the plural number condition is plotted in Fig. 6b.

As in Experiment 1, only F1-analyses were conducted on account of the low number of experimental trials in the experiment. The $2 \times 2 \times 2$ ANOVA comparing Phase (baseline vs. testing), Number (singular vs. plural) and Picture Type (1-actor vs. 2-actor) revealed a main effect for Picture Type ($F(1,27) = 11.18$, $p < 0.01$) and a marginal main effect for Phase ($F(1,27) = 3.99$, $p = 0.05$). The other interactions and effects did not reach significance (interaction between Phase and Picture Type: $F(1,27) = 2.71$, $p = 0.112$; main effect for Number: $F(1,27) = 1.46$, $p = 0.24$; all other F's < 1). Thus, children show a general preference for looking longer at the 2-actor picture than the 1-actor picture without a significant influence of Number or Phase.

For better comparison of these results with those from Experiment 1, separate analyses for the two number conditions were performed. In the 2×2 ANOVA for the singular number condition, the interaction between Phase and Picture Type did not reach significance ($F(1,27) = 2.56$, $p = 0.122$). But when the looking times at the 1-actor and the 2-actor pictures in the baseline phase of the singular number condition were directly compared, a significant preference for the 2-actor picture was found ($F(1,27) = 7.51$, $p < 0.05$), which could no longer be observed in the testing phase ($F(1,27) < 1$).

When performing a 2×2 ANOVA for the plural number condition, only a main effect for Picture Type was found ($F(1,27) = 4.41$, $p < 0.05$). The interaction between Phase and Picture Type did not reach significance ($F(1,27) < 1$), nor did the main effect for Phase ($F(1,27) = 2.72$, $p = 0.111$). The comparisons of the looking time in the baseline phase of the plural number condition revealed a marginal effect for Picture Type, since children looked longer at the 2-actor picture than at the 1-actor picture ($F(1,27) = 3.56$, $p = 0.07$). In the testing phase, as in the singular number condition, no such preference was observable ($F(1,27) = 1.54$, $p = 0.225$).

When further analyzing whether the presentation of the test sentence had a differential effect on children's preference for the 2-actor picture, a difference between the two number conditions emerged. It turned out that the decrease in looking time to the 2-actor picture from the baseline to the testing phase was significant in the singular number condition ($F(1,27) = 5.12$, $p < 0.05$), while there was no such significant change in the plural number condition ($F(1,27) = 2.63$, $p = 0.117$).

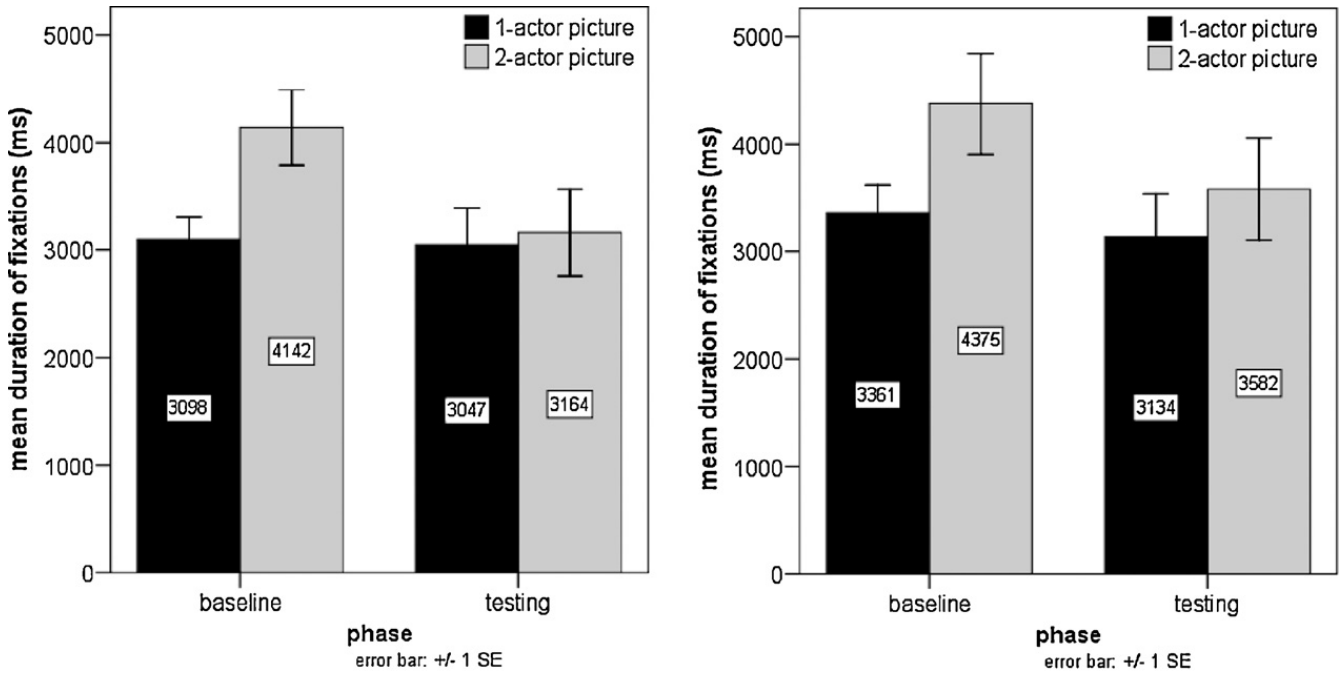


Fig. 6. (a) Children's mean looking times in the singular number condition (Experiment 2). (b) Children's mean looking times in the plural number condition (Experiment 2).

3.2.3. Pointing task

The mean percentages of the pointing reactions to the target picture as a function of the number condition were calculated.[8] In the 4 trials containing a singular inflected verb, the mean percentage of correct reactions (to the 1-actor picture) was 56.3%, (SD: 26.0%[9]) averaged over all 28 children. In the 4 plural trials, the mean percentage of correct reactions (to the 2-actor picture) was 48.5% (SD: 27.8%), averaged over all children. The mean percentage of correct pointing reactions does not differ significantly from chance-level performance in both number conditions (singular number condition: $t(27) = 1.27$, $p = 0.215$; plural number condition: $t(27) = -0.28$, $p = 0.779$). Additionally, mean percentages of correct pointings do not differ between the two number conditions ($t(27) = 0.88$, $p = 0.387$).

3.2.4. Discussion

The analysis of the looking time revealed that children fixated on the 2-actor picture longer than on the 1-actor picture during the baseline phase in both number conditions. Like in Experiment 1 we consider this as a consequence of the greater amount of visual information encoded in the 2-actor picture. As we failed to find any significant interactions between Phase and Picture Type in either of the two number conditions in this experiment, the effect of the presented sentence on children's looking behavior appears to be weaker than in Experiment 1. But separate analyses of the changes in looking times from the baseline to the testing phase in the two number conditions again revealed that hearing a sentence with a singular inflected verb is followed by a decrease in looking time to the 2-actor picture, which is not found after hearing a sentence with a plural inflected verb. This is in accordance with the results obtained in Experiment 1. Thus, there are indications that the number information provided by the verb inflection in the test sentences has an effect on children's looking behavior also in Experiment 2, even though the pattern is not as compelling as that obtained in Experiment 1. Children's looking behavior in both experiments displays the same pattern, but the effects themselves are much weaker in Experiment 2. Since the only difference between the two experiments was the presence or absence of an additional "picture-selection-by-pointing" task, we propose that this difference in tasks accounts in some way for the differences in findings across the two experiments. Possible underlying reasons for this task effect will be addressed in section 4.

Interestingly, the analysis of children's pointing reactions in Experiment 2 showed that the mean percentage of correct reactions did not differ from chance-level performance in both number conditions. This suggests that children were guessing when asked to point to the matching picture. This assumption is strengthened by the individual scores reached in the picture-selection task. Only one out of 28 children performed 100% correct in both number conditions.[10] 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. For the other 25 children, pointing reactions were distributed randomly, supporting the assumption of a guessing-strategy employed by most of the children (range of percentage correct in singular and plural number condition: 0–100% correct).

Thus we conclude that a picture-selection task like ours does not provide evidence that German-learning children aged 3 to 4 years 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; Perez-Leroux, 2005). This strengthens the assumption that it is not structural differences across the languages considered so far that account for the earlier competence found in German children (Experiment 1), but rather the different methods used.

But before concluding that eye-tracking is a more sensitive testing paradigm to demonstrate children's comprehension abilities, we have to consider the differences in the eye-tracking data across our two experiments, with less clear effects of the sentences on children's looking behavior in Experiment 2 than in Experiment 1. As we did not test the same children in Experiment 1 and Experiment 2, we cannot rule out that this difference is due to group differences between the children. But we consider this possibility as unlikely, as the groups differ neither in age nor with respect to the productive use of verb inflections. Another explanation may be that the additional pointing task in Experiment 2 interfered with children's eye gaze as an indicator of sentence interpretation, reflecting the explicit decisions the children were forced to make in the pointing task. This could lead to more switches across the pictures resulting in a smaller advantage for the target picture than in the first experiment. The fact that the gaze patterns in Experiment 2 were similar in their trends to those in Experiment 1 could provide a first evidence for that. But the relation between tasks that only involve looking 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 be the subject of future research.

4. General discussion

Our two experiments with German 3- to 4-year-old children on sentence interpretation based on verb inflection revealed discrepant findings. In Experiment 1, children did not have to perform a specific task. Instead their eye gaze was tracked to

[8] For one child, only 6 trials were scorable, as the child refused to point on the other two trials. For another child, only 7 trials were scorable. The remaining 26 children pointed on all 8 trials. The mean percentages calculated for each child are based on the number of trials that particular child performed in a scorable manner.

[9] Because there were only 4 trials per number condition and per participant, SD's are very high.

[10] Since there were only four experimental trials per number condition, even 75% correct, i.e. 3 out of 4 trials correct, cannot be considered as evidence for correct comprehension.

measure the comprehension of sentences in which the verb inflection was the only cue to subject number. In Experiment 2, children were presented with the same material, but had to perform an additional picture selection task while their eye gaze was tracked.

The findings of Experiment 1 provide clear evidence that German 3- to 4-year-old children are able to infer the number of the sentential subject solely by relying on the inflectional information. In Experiment 2 the eye-tracking data showed a similar but weaker pattern, but interestingly, in children's pointing reactions no evidence for using verb inflections in sentence comprehension was observable.

Based on the eye-tracking results we conclude that German 3- to 4-year-old children have acquired the morpho-syntactic knowledge about number agreement in the domain of the German 3rd person verb inflections, and that they can use this knowledge when determining the number of an ambiguous sentence subject. Hence, we assume that processing factors must account for children's failure to demonstrate this knowledge in the pointing task.

Still our findings raise two important questions that will be addressed in the following. First, we will be concerned with possible underlying reasons for the found within-modality asymmetry, thus with children's success in demonstrating their comprehension ability on verb inflections when tested via eye tracking versus their failure to do so when tested with a picture-selection task. Second, we will discuss what this within-modality asymmetry can tell us about asymmetries found across modalities, namely production and comprehension of verb inflections and other linguistic structures.

Concerning the first question, at least two possible explanations are conceivable. A first, rather imprecise one, would be that children's failure in the picture-selection task is due to general task demands. What these actually are is far from clear—even though this method is widely used in assessing children's language comprehension. 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 et al., 2007). 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. It may be hypothesized that these additional demands might not be fully developed in children at age 3. On the contrary, preferential looking 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. Additionally, it is closely time-locked to the processing of the incoming sentence (e.g. Trueswell and Gleitman, 2007). 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. This non-requirement of a choice might make eye tracking more adequate to employ when testing young children. It is further unclear whether the task of selecting among a set of pictures requires a different type of representation of either the linguistic or the visual stimulus.

A second explanation refers to different stages of the interpretation process that are reflected by the on-line measurements of eye tracking vs. the off-line response of picture selection. It seems probable that the latter one is more heavily influenced by heuristic processing strategies than the former one, leading to false interpretations or guessing in the later (e.g. Hurewitz et al., 2000). An effect of heuristics in sentence interpretation can even be found in adults when off-line tasks are used (Ferreira, 2003). 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. If such heuristics can overrule the information provided by the syntactic structure of a sentence even in adults, it seems reasonable to expect similar effects in children's sentence interpretation. Based on such observations, we suggest that the eye-tracking data in our study reflect the analysis of the structural properties of a sentence, which later on is shadowed by heuristics and an arising conflict between two different interpretations. The notion of different stages in the interpretation process that might be differently assessed in on-line and off-line tasks is also discussed by Trueswell and Gleitman (2007).

The second question–what our experimental findings can tell us about the relationship between early production and comprehension of verb inflections and the proposed asymmetry between these domains (Johnson et al., 2005; Perez-Leroux, 2006)—has to be faced against the background of the previous discussion. Remember that we concluded that the structural competence necessary to use the information provided by verb inflections can be assessed in 3- to 4-year-old German children when appropriate measures are employed. Thus according to our data, the comprehension ability is in place between 3- and 4-years of age, at least in German-speaking children. Concerning production of German verb inflections, we can call 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 (2002), who did not find more than 90% of children producing verb inflections before age 4. Additionally, the parental report data point to the problem of generalizing findings from spontaneous speech of very small groups of children or single case studies (e.g. Poeppel and Wexler, 1993), and that the spontaneous speech data on verb inflection production might overestimate children's abilities due to frequency and familiarity factors. Overall, 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.

Discrepancies across findings obtained with different methods might also be relevant for other linguistic areas for which asymmetric acquisition paths have been reported. One well known area is the acquisition of pronouns. While children as young as 3;0 produce pronouns correctly in spontaneous speech to express a disjoint meaning (Bloom et al., 1994), poor performance on the interpretation of these pronouns using picture-selection tasks is found until 6;6 (Chien and Wexler, 1990). Another example is the production and comprehension of the focus particle *auch* 'also' in German-learning children

(Höhle et al., 2009). Various explanations have been put forward to deal with these observations that production may be earlier in place than comprehension (Hendriks and Koster, 2010). For example, extra-grammatical and pragmatic factors have been assumed to lie at the heart of children's comprehension difficulties (Thornton and Wexler, 1999). Alternatively, grammar was viewed as a direction-sensitive system of rules in which different pairings between form and meaning can arise in production and comprehension (Hendriks and Spenader, 2006).

Interestingly, recent eye-gaze studies have found earlier comprehension than was evidenced in studies using picture-selection tasks for the above mentioned areas. Sekerina et al. (2004) have tested children's comprehension of pronouns using on-line and off-line tasks and found a dissociative pattern of performance in these two kinds of different tasks, with eye tracking revealing a more adult-like linguistic competence than indicated by picture selection. Höhle et al. (2009) conducted an eye-tracking study on children's processing of German focus particles (accented and unaccented *auch*, 'also') and found fixation patterns that indicated a higher degree of competence concerning these particles than found in other studies employing picture selection tasks (e.g. Hüttner et al., 2004 for German, Bergsma, 2002 for Dutch).

Altogether, the assumption that productive capacities for some linguistic expressions may precede their comprehension in language acquisition is mainly based on observations of non-adult-like performance in sentence comprehension tasks like picture selection and truth-value judgements. In line with other studies, our results reveal that it is not trivial to determine at which age a given linguistic structure or expression is comprehended by children. If one considers full comprehension capacity to require adult-like performance in all kinds of tasks that tap children's sentence comprehension skills, then we have to conclude that the 3-year-old children tested in our study did not show a full ability to interpret verbal inflection. On the contrary, if one argues that finding empirical evidence for children's correct interpretation in at least one of the methods available is a sufficient demonstration of the children's linguistic capacity, then variations across tasks have to be considered a consequence of specific processing constraints in children. Following this line of reasoning, the former account of the data would support the hypothesis of production preceding comprehension of number agreement, while the latter would not.

The problem of determining when a given structure has been mastered in language acquisition does not only hold for the receptive but also for the productive domain, where task dependent effects and differences in the use of specific expressions in spontaneous and elicited speech also may arise. For instance, Müller et al. (2009) found that the same children who evidenced the use of the focus particle *auch* in their spontaneous speech at the age of 2;0 produced the particle in an elicited production task in only 33% of the cases at age 2;4 and in 75% of the cases at 2;8. Similar discrepancies have been found for the use of the English 3rd person singular *-s* when comparing spontaneous speech (e.g. Brown, 1973) and elicited production data (e.g. Rice and Wexler, 2002).

This leads to the general conclusion that performance asymmetries within modalities have to be taken into account when discussing cross-modality asymmetries between production and comprehension capacities. First of all, as suggested above, the within-modality asymmetries are a challenge for determining the age of acquisition of a specific linguistic capacity either in production or in comprehension and thus for proving whether there are real asymmetries in the development of production and comprehension skills. Second, the knowledge about within-modality asymmetries in performance is relevant to decide whether production-comprehension asymmetries deserve some special attention by revealing something specific about children's underlying grammatical system that may be different for production and comprehension. Alternatively, asymmetries across modalities might prove to be nothing special, but to simply arise from the same sources as asymmetries within modalities, thereby reflecting different task demands that are put on the children and have an impact on their performance in production as well as in comprehension.

Thus, identifying areas in which production precedes comprehension needs intense empirical research that has to fulfil various methodological requirements. First, only those areas should be considered as candidates for this asymmetric pattern in which convergent findings of mastery or non-mastery of a specific linguistic structure obtained with different methods in either of the modalities are available (e.g. Hurewitz et al., 2000). Second, methods used to study production and comprehension performance should be as similar as possible with respect to the processing demands they pose on the 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. Furthermore, it would be necessary to investigate the production and comprehension abilities in the same group of children in a longitudinal fashion. To date, methodological aspects are still putting constraints on this way of comparison, since for example preferential looking and eye-tracking data are hardly interpretable on a single-subject basis unless a huge amount of trials would be used (Hirsh-Pasek and Golinkoff, 1996b). 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.

Acknowledgements

Funding for this research was partly provided by the European Commission within the collaborative research project Chlasc (*Characterizing Human Language by Structural Complexity*), Project Number 028395. Special thanks goes to Caroline Magister, Maja Stegenwallner and Katja Hummel for collecting the data. We would like to thank Tom Fritzsche for his support in preparing the eye-tracking experiments and analyzing the data, as well as for intense and helpful discussions on our findings. Further, we would like to thank Irina Sekerina and Bart Hollebrandse for their helpful comments on an earlier version of this manuscript. Last but not least, we thank all the children and their parents who participated in our study.

Appendix A. Verbs and objects used in experiment 1 and 2

		Verb	Object 1	Object 2
(1)	practice trials	streicheln	eineKatze	einBaby
		'to pet'	'a cat'	'a baby'
(2)	experimental trials	angeln	ein Fisch	ein Schuh
		'to fish'	'a fish'	'a shoe'
(3)		basteln	ein Hut	ein Drache
		'to do handicrafts'	'a hat'	'a kite'
(4)		füttern	einHund	ein Pferd
		'to feed'	'a dog'	'a horse'
(5)		öffnen	ein Geschenk	eine Tür
		'to open'	'a present'	'a door'

Appendix B. Sentences used in experiment 1 and 2

Sentences presented in practice trials (Version 1 and 2)	
(1)	Sie streichelt eine Katze. 'She pets a cat'
(2)	Sie streichelt ein Baby. 'She pets a baby'
(3)	Sie streicheln eine Katze. 'They pet a cat'
(4)	Sie streicheln ein Baby. 'They pet a baby'
Sentences presented in experimental trials (Version 1)	
(1)	Sie angelt einen Fisch. 'She fishes a fish'
(2)	Sie angeln einen Schuh. 'They fish a shoe'
(3)	Sie bastelt einen Hut. 'She makes (handicrafts) a hat'
(4)	Sie basteln einen Drachen. 'They make (handicraft) a kite'
(5)	Sie füttert einen Hund. 'She feeds a dog'
(6)	Sie füttern ein Pferd. 'They feed a horse'
(7)	Sie öffnet ein Geschenk. 'She opens a present'
(8)	Sie öffnen einen Tür. 'They open a door'
Sentences presented in experimental trials (Version 2)	
(1)	Sie angelt einen Schuh. 'She fishes a shoe'
(2)	Sie angeln einen Fisch. 'They fish a fish'
(3)	Sie bastelt einen Drachen. 'She makes (handicrafts) a kite'
(4)	Sie basteln einen Hut. 'They make (handicraft) a hat'
(5)	Sie füttert ein Pferd. 'She feeds ahorse'
(6)	Sie füttern einen Hund. 'They feed a dog'
(7)	Sie öffnet eine Tür. 'She opens a door'
(8)	Sie öffnen ein Geschenk. 'They open a present'

Appendix C. Introductory story

Schau mal, das sind Julia, Anna und Sarah. Sie machen viele tolle Sachen – mal alleine und mal zusammen. Gleich wirst du sehen, was.

'Look here, these are Julia, Anna and Sarah. They are doing all kinds of interesting things -either alone or together. Let's see what they do!'

References

Aguado-Orea, J., 2004. The acquisition of morpho-syntax in Spanish: Implications for current theories of development. Unpublished Ph.D. Thesis. The University of Nottingham, Nottingham.

Bergsma, W., 2002. (Un)stressed ook in Dutch. In: van Geenhoven, V. (Ed.), Semantics in Acquisition. Springer, Dordrecht, pp. 329–348.

Bittner, D., 2003. The emergence of verb inflection in two German-speaking children. In: Bittner, D., Dressler, W.U., Kilani-Schoch, M. (Eds.), Development of Verb Inflection in First Language Acquisition. A Cross-linguistic Perspective. de Gruyter, Berlin, pp. 53–88.

Bittner, D., Dressler, W.U., Kilani-Schoch, M., 2003. Introduction. In: Bittner, D., Dressler, W.U., Kilani-Schoch, M. (Eds.), Development of Verb Inflection in First Language Acquisition. A Cross-linguistic Perspective. de Gruyter, Berlin, pp. vi–xxxvii.

Bloom, P., Barss, A., Nicol, J., Conway, L., 1994. Children's knowledge of binding and coreference: evidence from spontaneous speech. Language 70, 53–71.

Brown, R., 1973. A First Language: The Early Stages. Harvard University Press, Cambridge, MA.

Chien, Y.-C., Wexler, K., 1990. Children's knowledge of locality conditions on Binding as evidence for modularity of syntax and pragmatics. Language Acquisition 13, 225–295.

Clahsen, H., 1986. Verb inflections in German child language: acquisition of agreement markings and the functions they encode. Linguistics 24, 79–121.

Ferreira, F., 2003. The misinterpretation of noncanonical sentences. Cognitive Psychology 47, 164–203.

Fraser, C., Bellugi, U., Brown, R., 1963. Control of grammar in imitation, comprehension and production. Journal of Verbal Learning and Verbal Behaviour 2, 121–135.

Gerken, L., Shady, M.E., 1996. The picture selection task. In: McDaniel, D., McKee, C., Smith Cairns, H. (Eds.), Methods for Assessing Children's Syntax: Language, Speech and Communication. MIT Press, Cambridge, MA, pp. 125–145.

Golinkoff, R.M., Hirsh-Pasek, K., Cauley, K.M., Gordon, L., 1987. The eyes have it: lexical and syntactic comprehension in a new paradigm. Journal of Child Language 14, 455–476.

Guasti, M.T., 1993. Verb syntax in Italian child grammar: finite and nonfinite verbs. Language Acquisition 3, 1–40.

Guasti, M.T., 2002. The Growth of Grammar. MIT Press, Cambridge, MA.

Hendriks, P., Spenader, J., 2006. When production precedes comprehension: an optimization approach to the acquisition of pronouns. Language Acquisition 13, 319–349.

Hendriks, P., Koster, C., 2010. Production/comprehension asymmetries in language acquisition. Lingua, this volume.

Hirsh-Pasek, K., Golinkoff, R.M., 1996a. The Origins of Grammar: Evidence from Early Language Comprehension. MIT Press, Cambridge, MA.

Hirsh-Pasek, K., Golinkoff, R.M., 1996b. The inter-modal preferential looking paradigm: a window onto emerging language comprehension. In: McDaniel, D., McKee, C., Smith Cairns, H. (Eds.), Methods for Assessing Children's Syntax: Language, Speech and Communication. MIT Press, Cambridge, MA, pp. 105–124.

Höhle, B., Berger, F., Müller, A., Schmitz, M., Weissenborn, J., 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., Schmitz, M., Santelmann, L.M., Weissenborn, J., 2006. The recognition of discontinuous verbal dependencies by German 19-month-olds: evidence for lexical and structural influences of children's early processing capacities. Language Learning and Development 2, 277–300.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L.R., Trueswell, J., 2000. One frog, two frog, red frog, blue frog: factors affecting children's syntactic choices in production and comprehension. Journal of Psycholinguistic Research 29, 597–626.

Hüttner, T., Drenhaus, H., van de Vijver, R., Weissenborn, J., 2004. The acquisition of the German focus particle *auch* 'too': comprehension does not always precede production. In: Proceedings of the 28th Annual Boston University Conference on Language Development. On-line supplement http://www.bu.edu/linguistics/APPLIED/BUCLD/supp.html.

Johnson, V.E., de Villiers, J.G., Seymour, H.N., 2005. Agreement without understanding? The case of third person singular /s/. First Language 25, 317–330.

Juscyzk, P.W., 1996. The Discovery of Spoken Language. MIT Press, Cambridge, MA.

Müller, A., Höhle, B., Schmitz, M., Weissenborn, J., 2009. Information structural constraints on children's early language production: the acquisition of the focus particle *auch* ('also') in German-learning 12- to 36-month-olds. First Language 29, 373–399.

Naigles, L., 2002. Form is easy, meaning is hard: resolving a paradox in early child language. Cognition 68, 157–199.

Nichols, J., 1986. Head-marking and dependent-marking grammar. Language 62, 56–119.

Perez-Leroux, A., 2005. Number problems in children. In: Gurski, C. (Ed.), Proceedings of the 2005 Canadian Linguistics Association Annual Conference, http://ling.uwo.ca/publications/CLA-ACL/CLA-ACL2005.htm.

Poeppel, D., Wexler, K., 1993. The full competence hypothesis of clause structure in early German. Language 69, 1–33.

Rice, M.L., Wexler, K., 2002. Test of Early Grammatical Impairment. Psychological Corporation, San Antonio, TX.

Santelmann, L.M., Jusczyk, P.W., 1998. Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. Cognition 69, 105–134.

Sekerina, I.A., Stromswold, K., Hestvik, A., 2004. How do adults and children process referentially ambiguous pronouns. Journal of Child Language 31, 123–152.

Soderstrom, M., Wexler, K., Jusczyk, P.W., 2002. English-learning toddler's sensitivity to agreement morphology in receptive grammar. In: Skarabella, B., Fish, S., Do, A.H.-J. (Eds.), Proceedings of the 26th Annual Boston University Conference on Language Development, 2. Cascadilla Press, Somerville, MA, pp. 643–652.

Soderstrom, M., 2008. Early perception-late comprehension of grammar? The case of verbal -s: A response to de Villiers & Johnson (2007). Journal of Child Language 35, 671–676.

Thornton, R., Wexler, K., 1999. Principle B, VP ellipsis, and interpretation. MIT Press, Cambridge, MA.

Tomasello, M., Carpenter, M., Liszkowski, U., 2007. A new look at infant pointing. Child Development 78, 705–722.

Trueswell, J.C., Gleitman, L., 2007. Learning to parse and its implications for language acquisition. In: Gaskell, G. (Ed.), Oxford Handbook of Psycholinguistics. Oxford University Press, Oxford, pp. 635–656.

Wexler, K., 1994. Optional infinitives, head movement and the economy of derivations. In: Lightfoot, D., Hornstein, N. (Eds.), Verb movement. Cambridge University Press, pp. 305–362.

Wexler, K., 1998. Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage. Lingua 106, 23–79.