German children use prosody to identify participant roles in transitive sentences*

THOMAS GRÜNLOH, ELENA LIEVEN and MICHAEL TOMASELLO

Abstract

Most studies examining children's understanding of transitive sentences focus on the morphosyntactic properties of the construction and ignore prosody. But adults use prosody in many different ways to interpret ambiguous sentences. In two studies we investigated whether 5-year-old German children use prosody to determine participant roles in object-first (OVS) sentences with novel verbs (i.e., whether they use prosodic marking to overrule word order as a cue). Results showed that children identify participant roles better in this atypically ordered construction when sentences are realized with the marked, OVS-typical intonational pattern, especially in combination with case marking (Study 1). In a second study, we embedded these sentences into an appropriate discourse context. The results show that, even in the absence of any case marking, children understand participant roles when they are realized with the appropriate intonational pattern. These findings demonstrate that young children can use intonation to help identify participant roles in transitive sentences, at least in marked constructions such as the German object-first (OVS) construction.

Keywords: Competition model; prosody; intonation; word order; case marking.

1.   Introduction

Speakers of different languages must use different cues to work out argument structure and sometimes the reliance on different cues changes in the course of development. To become competent with a language young children must master many different grammatical constructions and their associated cues:

pairings between forms and their relatively complex communicative functions. A construction of particular importance is the basic transitive construction—as in *She eats the apple*—because it is the developmentally earliest construction in which children must use grammatical cues to determine the different roles of the two participants involved (in this case what is doing the eating and what is getting eaten). In most languages, the transitive construction (or its nearest equivalent) marks the roles of the two participants with multiple, redundant cues. Thus, in our English example, *she* comes before the verb and *the apple* comes after it, but in addition the pronoun *she* is in nominative form (as opposed to the accusative form *her*). To comprehend such a sentence, then, children could use either the word order or the case marking of the pronoun, or both, to figure out the basic 'who did what to whom' of the sentence (there are also other grammatical cues, such as subject verb-agreement, that are sometimes diagnostic). Children can sometimes use semantic plausibility as well, in the sense that it is much more likely that the animate being *she* is doing the eating than being eaten, whereas the reverse is the case for the inanimate object *apple*.

Bates and MacWhinney (1987, 1989) performed a series of studies designed to determine how children in different languages use different grammatical cues to comprehend transitive sentences—resulting in what they call the competition model. This model is based on estimation of the validity or information value of different cues. The idea is that, in different languages, cues are accorded different weights in the comprehension process depending on the consistency (cue reliability) and frequency (cue availability) with which they are used. The product of these two components (cue validity) differs according to the language is being spoken. For example, English children learn to rely heavily on word order, so that from a fairly early age they interpret *The fence is kicking the cow*—a sentence in which word order and semantic plausibility (based on animacy) are in competition—quite literally as the fence (agent) kicking the cow (patient). In Italian—which has a more variable word order—children go with semantic plausibility in an equivalent sentence and infer that the animate being, the cow, is kicking the inanimate object, the fence (Bates et al. 1984; see Chan et al. 2009 for similar findings with Cantonese and English children).

The Competition Model is clearly a 'usage-based' model in that it ties the development of children's grammar to particular features of the input—the relative weights of individual cues. However, how these cues interact either during online processing or in the process of development is still an open question. One possibility is that children start by relying on only the most 'valid' cue for their language, only subsequently developing sensitivity to less 'valid' cues as their strength builds up. An alternative is that children may initially rely on a sentence schema in which all, or most, of the cues are present and only

subsequently abstract the relative value of each cue. Thus in the Dittmar et al. (2008) study, to be discussed in more detail, the youngest children were only able to correctly identify the subject of the sentence when it was marked by both case and SVO word order, reflecting the 'coalitions-as-prototypes' suggestion of the Bates and MacWhinney (1987) model. This would fit with evidence that children start by learning form-meaning patterns in which child-identified meanings are connected to schemas which are only partially analyzed into the components of the adult grammar. By the time children are five—the age of the children in this study—one would expect them to have gone some way towards identifying these cues and their particular role in the construction. In addition, morphological (e.g. case-marking), intonational and syntactic constructions (e.g. 'grammatical subject') are also being gradually abstracted on the basis of form and function relationships between constructions.

If children are indeed initially learning a schematic version of constructions then it is highly likely that, in real life, prosody is an essential component. Constructions have a characteristic prosody (Taylor 2002) and anecdotal evidence suggests that children have some understanding of the prosodic contrast between imperatives and questions at a very young age. Prosody as a grammatical cue has been virtually ignored by language acquisition researchers and has largely been studied only as a phenomenon of speech perception. For example, a fair amount of research has been done on infants' use of rhythmic properties to discriminate between languages (e.g. Mehler et al. 1988) and to segment the speech stream (e.g. Johnson and Jusczyk 2003). In terms of the grammatical use of prosody, some researchers have found that it has little or no effect on children's interpretation of structurally ambiguous sentences (e.g. Vogel and Raimy 2002, Choi and Mazuka 2003; but see Snedeker and Yuan 2008, for more positive results using both action and looking time measures). But Arnold (2008) recently found that 4- and 5-year-old children use the presence or absence of sentence accent to guide their interpretation of the degree to which noun phrases are given by the discourse context. A number of studies have shown that adult listeners use prosodic cues reliably to resolve syntactic ambiguities (Schafer et al. 2000) and to find phrasal boundaries (e.g. Carlson et al. 2009; see Speer et al. 2003 for a review).

Very few studies, however, have focused on the use of prosody to assign basic participant roles, such as the agent and patient in transitive sentences. In the framework of the competition model, Bates et al. (1984) found that 3.5-year-old Italian children used sentence stress as a cue, but only in interaction with non-canonical word order (and the effect went away with older children). A language where prosody might be even more important for interpreting transitive sentences is German. While most transitive sentences in German have agent-patient word order (with the main verb in either verb-second or verb-final position), word order can be variable, with the patient sometimes

coming first. The inviolable cue for agent-patient relations is case marking, which occurs on the determiner. The problem is that the case system has been prone to much syncretism, and so sometimes case marking is ambiguous. The following examples illustrate the situation.

(1)   *Der Löwe jagt den Hund.* [word order and case both usable]
       The-masc-nom lion is.chasing-3SG the-masc-acc dog.
       'The lion (NOM) is chasing the dog (ACC).'

(2)   *Die Katze jagt die Ziege.* [case marking ambiguous]
       The-fem-nom/acc cat is.chasing-3SG the-fem-nom/acc goat.
       'The cat (NOM/ACC) is chasing the goat (NOM/ACC).'

(3)   *Den Hund VERB der Löwe.* [word order & case marking conflict]
       The-masc-acc dog is.chasing-3SG the-masc-nom lion. [lion is agent!]
       'The dog (ACC) is chasing the lion (NOM).'

In the first, prototypical example, word order and case marking both indicate the first noun phrase as the agent. In the second example case marking is ambiguous and thus it is unclear whether the first noun phrase is the patient and the second noun phrase is the agent or vice versa. In this case, word order is typically used (i.e. again identifying the first noun phrase as the agent). In the third example—a so-called patient-first sentence—case marking and word order conflict and, due to the nature of German grammar, case marking prevails (and the preverbal noun phrase is the patient). A construction like this where the first noun phrase is the patient is much less common in German, and it therefore typically occurs with prosodic stress on the first noun phrase.

Weber et al. (2006) tested German adults using an eye tracking paradigm and found that they attended to and used prosodic information to guide their comprehension of such sentences. German sentences were presented with case-ambiguous first NPs and unambiguous second NPs, e.g.:

(4)   *Die Katze jagt womöglich den Vogel.*
       'The cat (NOM/ACC) is possibly chasing the bird (ACC).'

(5)   *Die Katze jagt womöglich der Hund*
       'The cat (NOM/ACC) is possibly chased by the dog (NOM).'

Specifically, immediately upon hearing the first noun phrase, subjects fixated on the agent of the action (in a picture depicted by the sentence) when the nuclear accent (sentence stress) was on the verb, typically used for canonical agent-first sentences. In contrast, when the realization of the nuclear accent was on the first NP, typically indicating non-canonical patient-first sentences, participants interpreted the ambiguously case-marked, first noun phrase as the patient. These results showed that adult-listeners do use prosodic information

in the interpretation of ambiguous SVO and OVS sentences when no clear morphological information is available. Before the onset of the second NP, the patient was fixated upon more often than the agent when the intonational pattern already indicated the first NP as the agent, but not when intonation pointed to NP1 as the patient. Thus, the interpretation of word-order ambiguities was modulated by prosody and this was integrated rapidly enough to affect listeners' interpretation of grammatical function and assignment of participant roles before case information became available to clarify the ambiguity.

Dittmar et al. (2008) investigated young German children's comprehension of transitive sentences (containing nonsense verbs) that had various combinations of word order and case marking cues. They found that children as young as 2.5 years of age had a strong word order bias. They only correctly interpreted transitive sentences in which both word order and case marking indicated the first noun phrase as the agent. But when word order and case marking conflicted, as in (3), only the 7 year-olds behaved like adults by relying on case marking over word order. That is to say, the 2-year-olds and 5-year-olds most often interpreted the agent in sentences such as *Den Hund VERB der Löwe* as being the first noun phrase, whereas adults chose the second noun phrase almost 100% of the time. The problem, however, is that in this study all of the sentences were produced for the children with very similar prosody for all conditions. But patient-first sentences are not felicitous if they do not have the typical OVS-marked intonational pattern. It is therefore possible that young children are capable of understanding patient-first transitive sentences but only when the natural prosodic pattern that they hear in their everyday environment is present (as it was for the German adults in the Weber et al. 2006 experiment).

In the current study, therefore, we used a paradigm very similar to that of Dittmar et al. (2008) but systematically varied prosodic cues. In two studies, we presented five-year-old German children with transitive sentences involving nonce verbs (so that they could not use verb-specific information to interpret the sentences). Both studies employed a 2×2 design. Sentences either had ambiguous case marking or else they were marked by case on the determiner as patient-first sentences (the kind that children systematically misinterpreted in the Dittmar et al. 2008 study). Crossed with this variable, we either provided or failed to provide a rising L+H* pitch accent on the first noun phrase (of the type successfully used by German adults in the Weber et al. 2006 study). The question was thus whether children would use pitch accent on the first noun phrase in an adult-like manner to interpret transitive sentences and move away from their strong word order bias—both when case marking indicated that the first noun phrase was the patient and also when case marking was ambiguous so that pitch accentuation, in a sense, competed with word order. Our prediction was that 5-year-old children should be able to use the prosodic cue provided, and so show more skill with non-canonical, patient-first transitive sentences

than children in the Dittmar et al. 2008 study. If so, it would be the first study to our knowledge in which young children systematically use prosodic information as a grammatical cue to assign basic participant roles during sentence interpretation.

2.   Study 1

Following Dittmar et al. (2008), in the first study we used a video-pointing task to evaluate young German children's tendency to interpret transitive sentences on the basis of word order and case marking. We presented these sentences as either clearly case-marked as in (6) or ambiguous as in (7):

(6)   *Den Hund VERB der Löwe.*
      The-masc-acc dog is.verbing-3SG the-masc-nom lion.
      'The (ACC) dog is VERBING the (NOM) lion.'

(7)   *Die Katze VERB die Kuh.*
      The-fem-nom/acc cat is.verbing-3SG the-fem-nom/acc cow.
      'The (NOM/ACC) cat is VERBING the (NOM/ACC) cow.'

What was new in the study was that we either did or did not provide a prosodic cue that indicates a patient-first interpretation for adults (Weber et al. 2006). To make sure that the prosodic cue was given appropriately and consistently, all sentences were computerized and manipulated regarding their intonation. The prerecorded stimuli were presented to children over a hidden audio speaker.

2.1.   *Method*

Participants

Sixteen normally developing, monolingual German children with an average age of 4.10 years (range 4.5–5.3; 8 boys and 8 girls) were included in the study. An additional 2 children were tested but excluded from the study due to disinterest in the video clips (1) or experimenter error (1). Children were recruited from a database of parents who had volunteered to participate in psychological studies. They came from diverse socio-economic backgrounds. All children were tested in nursery schools in a medium-sized German city. As a control group, we tested 10 adults with the same procedure.

Materials

All novel verbs referred to prototypical causative transitive actions, involving direct contact between a volitional agent and an affected patient. Actions were reversible and involved either a caused change-of-state or a change-of-location.

The four novel verbs *wiefen*, *tammen*, *baffen* and *mommeln* were used to describe four novel transitive actions that were performed with four novel apparatuses. *Wiefen* was used to refer to an animal rocking another animal that was standing on an apparatus like a rocking-chair, by pushing it with its head. *Tammen* referred to an animal pushing down on another animal by jumping on its back so that the platform it was standing on, with a spring underneath, sank. *Baffen* was used to refer to an animal spinning around another animal that was standing on a disk. The fourth novel verb *mommeln* referred to an animal jumping on a platform in order to catapult an animal standing on the other side of this catapult. (For test sentences and animal pairing see Appendix A). The agents and patients of a particular event were pairs of animals with the same grammatical gender. Exactly which gender depended on the condition. All children heard the same test sentences in four conditions: In Condition 1, the Case Marking / Contrastive Intonation condition, they heard the novel verbs within an argument structure in which the patient was the first noun phrase and was case marked with the accusative, and the agent was the second noun phrase and was case marked with the nominative; for example, *Den (+accusative) Hund wieft der (+nominative) Elefant.* — 'The (+accusative) dog is weefing the (+nominative) elephant'. The intonational realization of the utterances was characterized by a strong, pitch accent on the first noun phrase. In Condition 2, the Case Marking / Neutral Intonation condition, children heard a sentence structure with the same grammatical markings as in Condition 1, but here, the construction was completely de-accented. In Condition 3 and 4, children heard the same intonational realization of target sentences as in the first and the second conditions, i.e. strong pitch accent on NP1 or de-accentuated intonational contour. But, in these conditions, children did not get any information about participant roles from word order or case marking, because they were ambiguous. For an example of the experimental conditions, see Table 2.

The two intonation patterns employed are transcribed using GToBI-annotations (Grice and Baumann 2002; Grice et al. 2005). In GToBI, tonal values can be described as either high (H) or low (L) targets which are anchored to prominent elements. The association with prominence is characterized by the association of tones to the lexically stressed syllable of the designated word. Pitch accents are marked with a star '*' following the tone, e.g. H*. Additionally, the intonational contour can be described with the 'leading tone' i.e. the pitch immediately before the pitch accent. According to this, the intonational realization of the utterances in Conditions 1 and 3 was characterized by a strong, rising L+H* pitch accent on the first nominal phrase (see Figure 1a). All stimuli were recorded by a female native speaker. She was asked to utter the sentences with as much emphasis as possible in the Contrastive Intonation conditions or as flat as possible in the Neutral Intonation conditions (see Figure 1b). If necessary, the recordings were later edited and manipulated
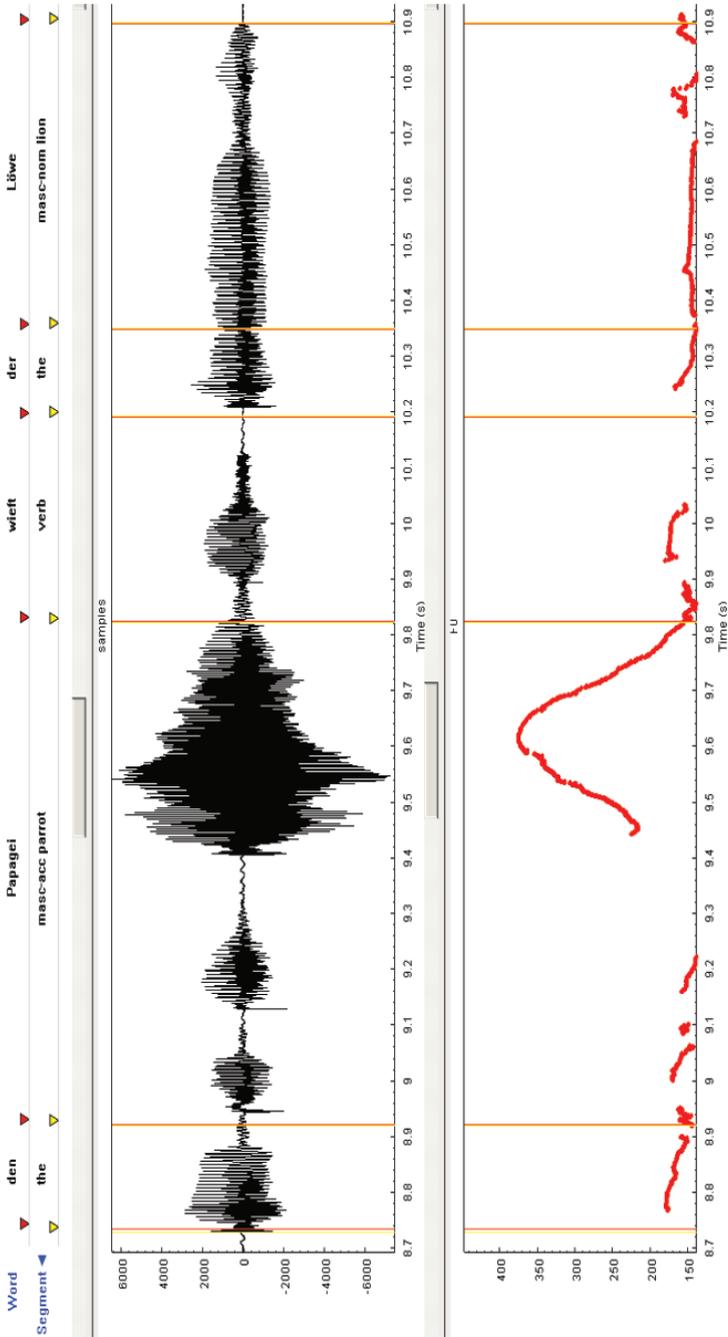
Figure 1a.   *Example of the intonation of the target utterance in the Contrastive Intonation condition. The contour bears a L+H* pitch accent on the first Nominal Phrase.*
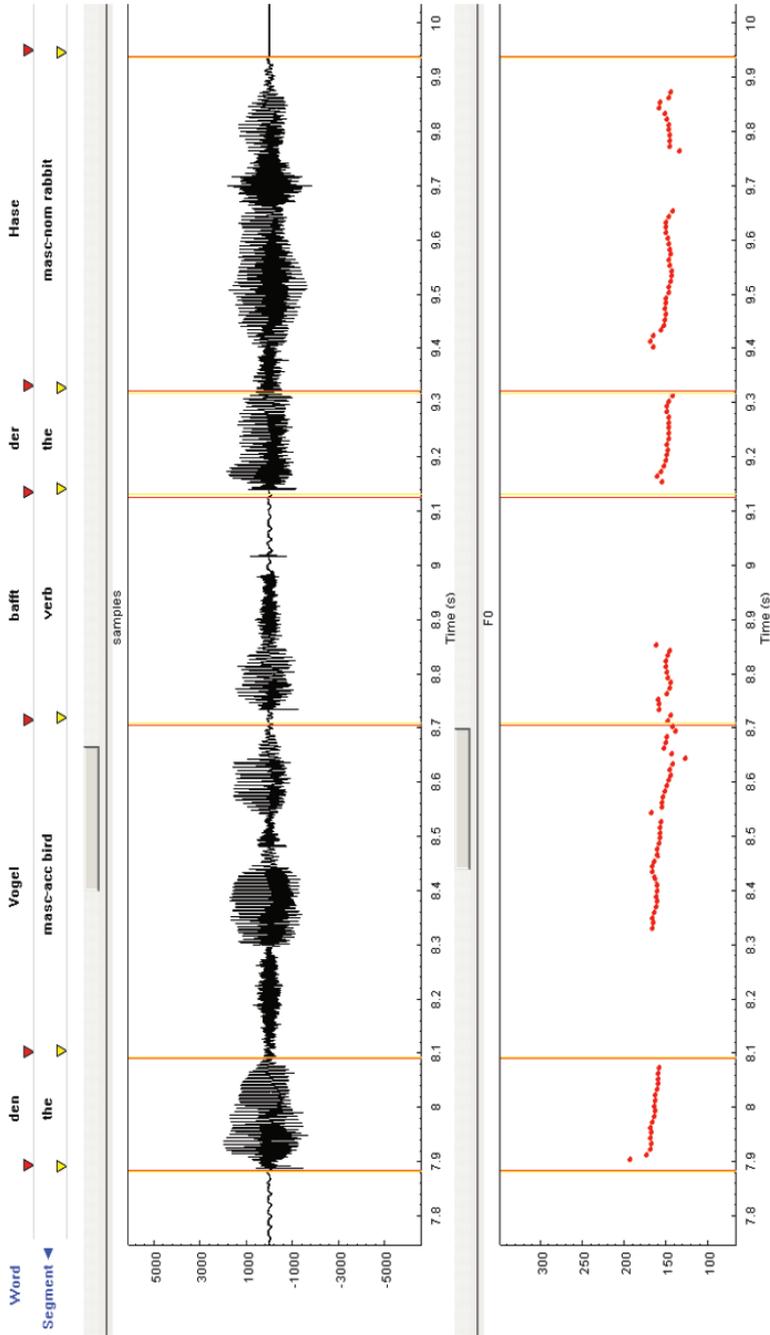
Figure 1b.   *Example of the monotonised intonation of the target utterance in the Neutral Intonation condition.*

Table 1.   *Acoustic properties of the stimuli. The table shows the mean minimum and maximum fundamental frequency (F0) and the pitch range in semitones (st) of NP1 and the whole utterance plus the standard deviation in parentheses.*[1]

	NP1			Utterance		
	Minimum F0 in Hertz	Maximum F0 in Hertz	Pitch range in st	Minimum F0 in Hertz	Maximum F0 in Hertz	Pitch range in st
Contrastive Intonation	131,53 (38,1)	384,25 (39,5)	19,27 (6,5)	105,85 (34,2)	384,26 (39,5)	23,08 (6,2)
Neutral Intonation	150,19 (7,3)	187,26 (20,2)	3,7 (1,14)	133,81 (26,7)	202,14 (25,1)	7,44 (4,0)

Table 2.   *Examples of the four test conditions containing the four novel transitive actions.*

	Contrastive Intonation	Neutral Intonation
Case Marking	*Den PAPAGEI wieft der Elefant* 'The (ACC) PARROT is weefing the (NOM) elephant.	*Den HUND bafft der Löwe* 'The (ACC) DOG is baffing the (NOM) lion'
No Case Marking	*Das ZEBRA tammt das Eichhörnchen* 'The (NOM/ACC) ZEBRA is tamming the (NOM/ACC) squirrel'	*Die KATZE mommelt die Ziege* 'The (NOM/ACC) CAT is mommeling the (NOM/ACC) goat'

by a speech analyst and ToBI-expert. He ensured that the stimuli were as natural as possible.

In Condition 3, the No Case Marking / Contrastive Intonation Condition, the German case marking was ambiguous (because only those animals were used that take the German feminine or neuter gender, which does not decline in the nominative or accusative case, e.g. *Die Katze wieft die Ziege.* — 'The cat is weefing the goat.') and thus it was unclear whether the patient was the first noun phrase and the agent was the second noun phrase or vice versa. But, as in Condition 1 intonation was characterized by a strong, contrastive L+H* accent on the first noun phrase, which indicates NP1 as the patient.

Finally, in Condition 4, the No Case Marking / Neutral Intonation structure, the children heard a sentence structure with the same grammatical markings, but with monotonised intonation. Each of the four conditions was tested with each of the four novel verbs; therefore each child heard 16 test sentences (see Table 2).

---

1.   Examples of the sound files of the test stimuli can be found under http://dx.doi.org/10.1515/COGL.2011.015_suppl_1

Design

We tested each child with four different novel verbs in transitive sentence structures using a video pointing task. During the session, the children sat in front of a 23‖ TFT-screen (1920 * 1200 Pixel, aspect ratio 16:10). In the test trials, the child saw two film scenes on the computer screen, each starting simultaneously and lasting 6 s, followed by a still image of the clips. Both of these scenes involved animals enacting the same causative event and differed only in that the agent and patient roles were reversed. All children received alternating test sentences with the four different conditions and all four novel verbs were tested in one session. All children got four trials (one for each novel verb) in each of the four conditions, leading to a total of 16 test-sentences. The order of the conditions and the novel verbs were counterbalanced in a 4 * 4 Latin square. The target screen order was counterbalanced so that the patient-first scene was presented on each side (left [L] or right [R]) in eight out of 16 trials for each child (e.g. for the pairing 'dog weef lion' and 'lion weef dog', half of the children saw the patient-first scene on the right initially and the other half saw it on the left, depending on counterbalance order). A particular side was never the correct choice for the patient-first scene more than twice in a row. No child experienced a test session in which the patient-first scene alternated regularly (e.g. LRLRLRLRL). The direction of the action was also counterbalanced (e.g. in the pairing 'dog weef lion' and 'lion weef dog' half of the children saw the agent performing the action from the left side of the scene towards the right side, and for the other half they saw the reverse). In order not to take any cues from the experimenter, the test trial was conducted with a talking puppet. All auditory stimuli were prerecorded and uttered by the puppet.

Procedure

The test session lasted for approximately 15 minutes. The computer monitor was positioned on the table approximately 50cm in front of the child. All sessions were videotaped with a camera centered behind the child, recording the child's pointing behavior. The experimenter never looked at the screen during the test trials but sat behind the screen pretending to read.

Pointing practice training: To teach the children that the aim of the task was to point to one of two pictures on the computer screen, we used a very easy warm-up task with two pictures depicting objects; for example, cheese and bacon, appeared on the screen simultaneously. The children were then asked by the experimenter to point to one of the two objects (e.g. *Zeig mir das Bild: Das ist der Käse.* — 'Show me the picture: That's the cheese'). We repeated this task 10 times with different pictures and all children solved it perfectly.

Word learning training: Each of the novel verbs and the corresponding actions were presented to each child through a live performance given by the

experimenter. To show and teach the different functions of the novel apparatuses, and thus the novel verbs, the experimenter performed the novel actions using animals whose labels take the German feminine gender and are ambiguous in the nominative or accusative case (e.g. *Ziege*—'goat' and *Ente*—'duck'). Each of the four novel verbs used in the test were randomly presented one after another by the experimenter in a variety of argument structures: in the citation form with no arguments (e.g. *Das heißt wiefen.*—'That's called weefing'), as well as in a transitive argument structure with two feminine pronouns (which are identical for subject and object position in German) in three different tenses (*Sie wird sie wiefen.*—'She is going to weef her'; *Sie wieft sie.*—'She is weefing her'; *Sie hat sie gewieft.*—'She weefed her'). The child was asked to repeat the verb using a prescribed question format (e.g. *Kannst du das sagen: wiefen?*—'Can you say that: weefing?') while the experimenter performed the action.

Film familiarization trials: Following the live enactment, the puppet declared that she has designed special clips which she wants to show the child and the experimenter; the child always agreed to see them. The child then received a familiarization trial for each verb in which he or she watched one film scene on just one half of the screen, involving two animals, with German feminine or neutral gender, acting out the novel verbs. At the same time, the puppet described the scene in a scripted manner; for example, *Guck mal*, *das heißt wiefen.*—'Look, that's called weefing'; all the while the other half of the screen remained blank. The side of the screen where the children saw the first picture (left or right), the acting direction, as well as the order of the novel verbs, was counterbalanced across and within subjects. At the end of each scene, the experimenter pointed to each animal and asked the child: *Wer ist das?*—'Who's that?' The majority of the children had no problem in spontaneously naming the participating animals. If a child did not name one of the animals, the experimenter told the child the name and asked him or her to repeat it, which nearly all of the children then did.

Test trial: The puppet then told the child and the experimenter that she had even more films that she would like to show. The experimenter then said that unfortunately he needed to read something and had no time to watch these clips with the child and puppet. He then sat behind the screen, and ran the computer program. Shortly afterwards, a red dot focused the child's attention on the center of the computer screen. The test trial then began and the child watched two scenes simultaneously (see Figure 2), which were accompanied by a prerecorded linguistic stimulus, explaining who was present in the clips and what they were doing; for example: *Guck mal, das Schwein und das Pferd. Das heisst wiefen!*—'Look, the pig and the horse. That's called weefing!'.

After the videos had stopped, the prerecorded voice of the puppet asked the child to point to the correct (still) picture by asking the target sentence accord-

Figure 2.   *Example of the visual setup used in the test trial. In the left scene the horse is 'weefing'*
the pig, in the right scene the pig is 'weefing' the horse.

ing to one of the four conditions; for example, *Zeig mir das Bild: Das Schwein wieft das Pferd!*—'Show me the picture: The (+nominative/accusative) pig weefs the (+nominative/accusative) horse!'. If the child did not point to one of the two film scenes, the puppet repeated the question a second time; however, she never asked the child to point again once she/he had already done so. Once the child had pointed to one of the two pictures, the next test trial began, preceded once more by the red dot.

2.2.   *Coding and reliability*

For every test trial, we coded responses for whether participants pointed to the picture in which the postverbal, second noun in the sentence was the agent. This was, of course, correct in the Case marking conditions, but either picture choice was possible in the No Case marking conditions. The question of interest is whether the addition of intonation would influence the children's choices. If a child did not choose either scene (this was the case in 2 trials), we coded those trials as 'wrong' (an alternative analysis in which these cases were excluded had no effect on the pattern of the results). All children were coded by the first author, and an additional coder coded 25% of all trials for testing reliability (= complete session of four randomly selected children). This revealed a perfect agreement with the first author (Cohen's Kappa = 1.0).

2.3.   *Results and discussion*

We tested for the proportion of times the NP occurring after the verb was identified as the agent of the action out of four. The data were analyzed using a
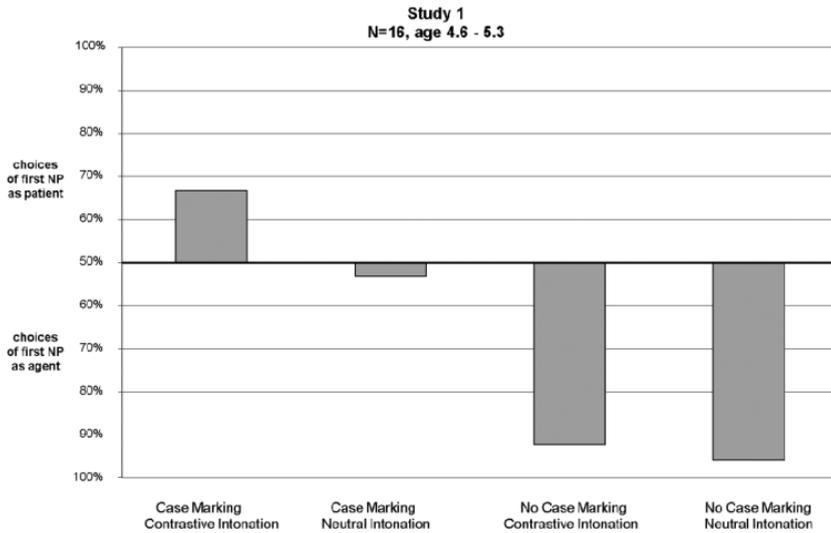
Study 1
N=16, age 4.6 - 5.3

Figure 3.   *Results of Study 1 for children in the four conditions. The diagram shows percentages of judging NP1 as either object-first or subject-first constructions as compared with chance, 50%.*

2 (Intonation) × 2 (Case Marking) repeated measures analysis of variance (ANOVA)[2]. There were main effects for both Intonation, $F(1, 15) = 4.88.4$, $p = .043$ and Case Marking $F(1, 15) = 42.8$, $p < .001$, but there was no significant Intonation × Case Marking interaction, $F(1, 15) = 3.608$, $(p = 0.061)$ (see Figure 3).

Because the chance level for our dependent variable was always 50%, we also investigated in which conditions the children were above chance in choosing the first noun as patient. The results show that the children were only above chance in the Case Marking / Contrastive Intonation condition (Condition 1; one sample t-test: $t[15] = 2,2$, $p = 0.044$). In contrast, in the Case Marking / Neutral Intonation, the children were approximately at chance level (Condition 2; $t[15] = −0.355$, $p = 0.728$) and in the No Case Marking / Contrastive Intonation (Condition 3) as well as in the No Case Marking / Neutral Intonation Condition (Condition 4), children were below chance (both $t[15] < −14$, both $p > 0.01$), i.e. they were significantly more likely to choose the first noun as agent. A comparison between the two conditions Case Marking / Contrastive

---

2.   Additionally, the data has been analyzed using a General Linear Mixed Model. These results revealed the same overall pattern of results, i.e. significance values of interactions and main effects.
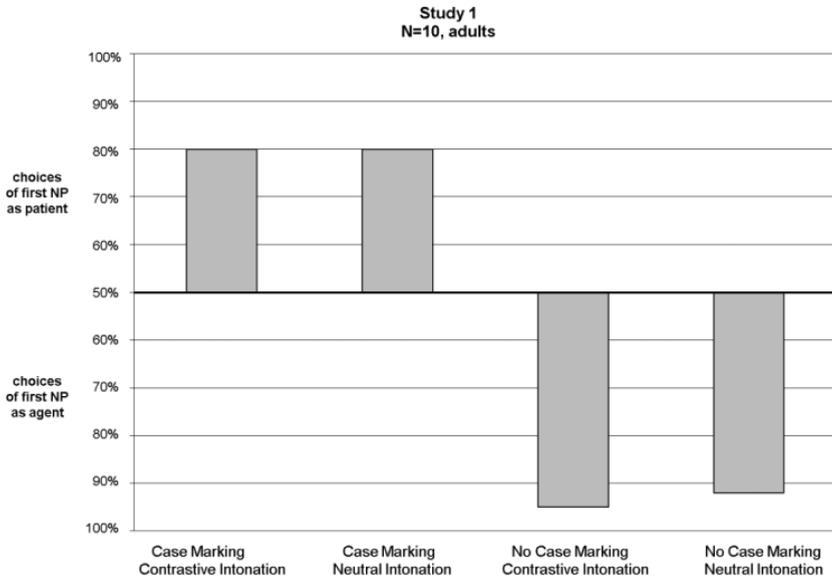
Figure 4. *Results of Study 1 for adults in the four conditions. The diagram shows percentages of judging NP1 as either object-first or subject-first constructions as compared with chance, 50%.*

Intonation and Case Marking / Neutral Intonation showed that children were significantly better in judging participant roles when intonation was available (paired-sample t-test: t[15] = 2.36, 0.032). Choices in the two conditions No Case Marking / Contrastive Intonation and No Case Marking / Neutral Intonation were not significantly different (t[15] = 0.368, p = 0.718).

2.4.   Adult control group

For the adult control group, we found main effects for Case Marking, F(1, 9) = 50.08, p < .001, but not for Intonation and no significant interaction between these two. (see Figure 4).

For a further analysis, we compared the results from the children with those of the adults. The data were analyzed using a three-way mixed analysis of variance (ANOVA) with two repeated factors (Case Marking and Intonation) and one between-subjects factor (Age). There were main effects for Case Marking, F(1, 24) = 96,72, p < 0.01, but not for Intonation, F(1, 24) = 3.12, p = 0.09. There was a marginally significant interaction between Case Marking and Age, F(1, 24) = 4.49, p = 0.045, but no significant difference between Intonation and Age (F[1, 24] = 1.9, p = 0.180), between Case Marking and Intonation

(F[1, 24] = 1.48, p = 0.235) or between a three-way-interaction (F[1, 24] = 2.66, p = 0.115). A comparison between conditions of the two studies only revealed a significant difference between children and adults in the Case Marking / Neutral Intonation conditions (t[9] = −3.35, p = 0.008).

What these results show is that the children are using case marking when it is available, and word order when it is not, to interpret the roles of the NP's in transitive sentences. Thus, children moved strongly away from choosing NP1 as the agent when case marking indicated this as the patient. The two conditions without case marking show that intonation by itself is not sufficient for these young children to identify a transitive construction combined with the appropriate OVS-intonation pattern. They instead rely heavily on the word order cue, choosing therefore the first noun as the agent. Comparison of the two conditions with case marking however, suggests that the intonation and case marking together provide a stronger cue than case marking alone. This was not the case with adults who could use case marking alone to select NP1 as the patient. This shows that children can use intonation in order to glean extra information when it is used redundantly with other cues. This finding is broadly consistent with the findings of Dittmar et al. (2008) that German children best comprehend transitive sentences with multiple, redundant cues. In their study the two cues that reinforced one another were word order and case marking, and in the current study they were case marking and intonation (pitch accentuation).

3.   Study 2

The sentences in Study 1 were presented to children outside of any meaningful discourse context. If intonational highlighting serves in many cases to contrast the stressed item with something in the previous discourse, then one could argue that presenting sentences in isolation does not provide children with a natural interpretive context and is, in fact, contrary to the principles of a usage-based approach. Indeed, in the adult literature, it has been argued on several occasions that experimenters should present intonationally contrastive sentences in more natural discourse contexts (e.g. Albritton et al. 1996). In the second study, therefore, we used the same linguistic materials and same basic method as in Study 1, with one crucial difference. Each test sentence was preceded by a discourse context in which a speaker described a scene incorrectly by misidentifying the patient using a normal, agent-first transitive sentence (e.g. 'The dog is weefing the frog', when in fact he is weefing the lion). The test sentence was then a patient-first transitive sentence, uttered as a correction, with stress on the patient (in very loose translation, 'No, it is the LION that's getting weefed.'). This is arguably something close to the "natural home" of patient-first transitive sentences in everyday German discourse, and should

give young children a better opportunity to show even more skills in using intonation to interpret patient-first transitive sentences.

3.1.   *Method*

Participants

Sixteen normally developing, monolingual German children with an average age of 4;10 years (range 4;6–5;3; 10 boys and 6 girls) were included in the study. Children were recruited from a database of parents who had volunteered to participate in psychological studies. They came from diverse socio-economic backgrounds. All children were tested in nursery schools in a medium-sized German city.

Materials & Design

Materials and design were the same as in Study 1 with the exception that the instructions for the test trials did not come from just one puppet, but instead were communicated in a conversation between two puppets. Whereas one of the puppets was the same character as in Study 1 (P1), the other puppet (P2) was introduced as an unreliable character, because he was too young to know the names of the animals or not able to remember the novel verbs. Instead, he said everything wrongly and was therefore corrected by P1. Thus, the target instruction in the form of the transitive OVS utterance (using the same stimuli as in Study 1) was embedded in a contrastive context.

   All children heard the same test sentences (see Appendix B) in a transitive OVS structure. The same four novel verbs were used in the same four conditions as in Study 1: Case Marking / Contrastive Intonation, Case Marking / Neutral Intonation, No Case Marking / Contrastive Intonation, No Case Marking / Neutral Intonation. Before the child heard the target sentence, P2 uttered a transitive SVO sentence, in which the patient was always wrong as in (8). P2 was then corrected by P1 using an utterance of the target sentence in transitive OVS structure, as in (9).

(8)   *Der Löwe VERB den Frosch!*
      The-masc-nom lion is.verbing-3SG the-masc-acc frog!
      'The lion (NOM) is VERBING the frog (ACC).'

(9)   *Nicht den Frosch VERB der Löwe, sondern **den Hund VERB der Löwe**!*[3]
      Not the-masc-acc frog is.verbing-3SG the-masc-nom lion,
      but the-masc-acc dog is.verbing-3SG the-masc-nom lion!

---

3.   The second NP, printed in bold, was the same auditory stimuli used in Study 1. Except from that, all other auditory stimuli in this study were natural and unmanipulated.

'It's not the frog (ACC) that's VERBING the lion (NOM), it's the dog (ACC) that's VERBING the lion (NOM)!'

An example of the first part of the correcting utterance as in (9) can be seen in Figure 5. The stimuli were recorded by the same female native speaker as in Study 1. She was invited to utter the sentences as naturally as possible, leading to a L+H* accent on NP1. Other than the second part of the utterance (the target OVS-sentence), which was recycled from Study 1, the speech material was not manipulated.

Procedure

The procedure of this study was the same as in Study 1 with the exception that the instructions did not come solely from one puppet but were embedded in a conversation between two puppets, as described previously.

Pointing practice training & Word learning training: Pointing practice training & Word learning training were the same as in Study 1.

Film familiarization trials: Following the live enactment of the word learning training, the child then saw a familiarization trial for each verb in which he or she watched each of the two film scenes in sequence and heard the two puppets describing them. In this description, P2 was always wrong because he was too young to remember the novel verbs and was thus corrected by P1; for example:

P2: *Guck mal, das heißt lemmen.*—'Look, that's called lemming.'
P1: *Nein P2, das heißt nicht lemmen, sondern wiefen. Das heißt wiefen.*—'No, P2, that's not lemmen. That's weefing! That's called weefing.'

During these film familiarization trials, only one clip was visible on the screen while the other half of the screen remained blank. The side of the screen where the children saw the first picture (left or right) as well as the order of the novel words was counterbalanced across and within subjects. At the end of each film scene, the experimenter pointed to both animals and asked the child: *Wer ist das?*—'Who's that?'. The majority of the children had no problem spontaneously naming the participating animals. If a child did not name one of the animals, the experimenter told the child its name and asked him or her to repeat it, which nearly all of the children then did.

Test trial: The test trial procedure was the same as in Study 1, except for the second puppet. At the moment where the attention-getter disappeared, P2 declared that he probably knows what happens in the next clips by saying a transitive SVO sentence, involving the novel verb and the right agent, but the wrong patient, as in (8). After finishing this sentence, the two clips appeared on the screen, accompanied by P1's prerecorded linguistic stimulus using the
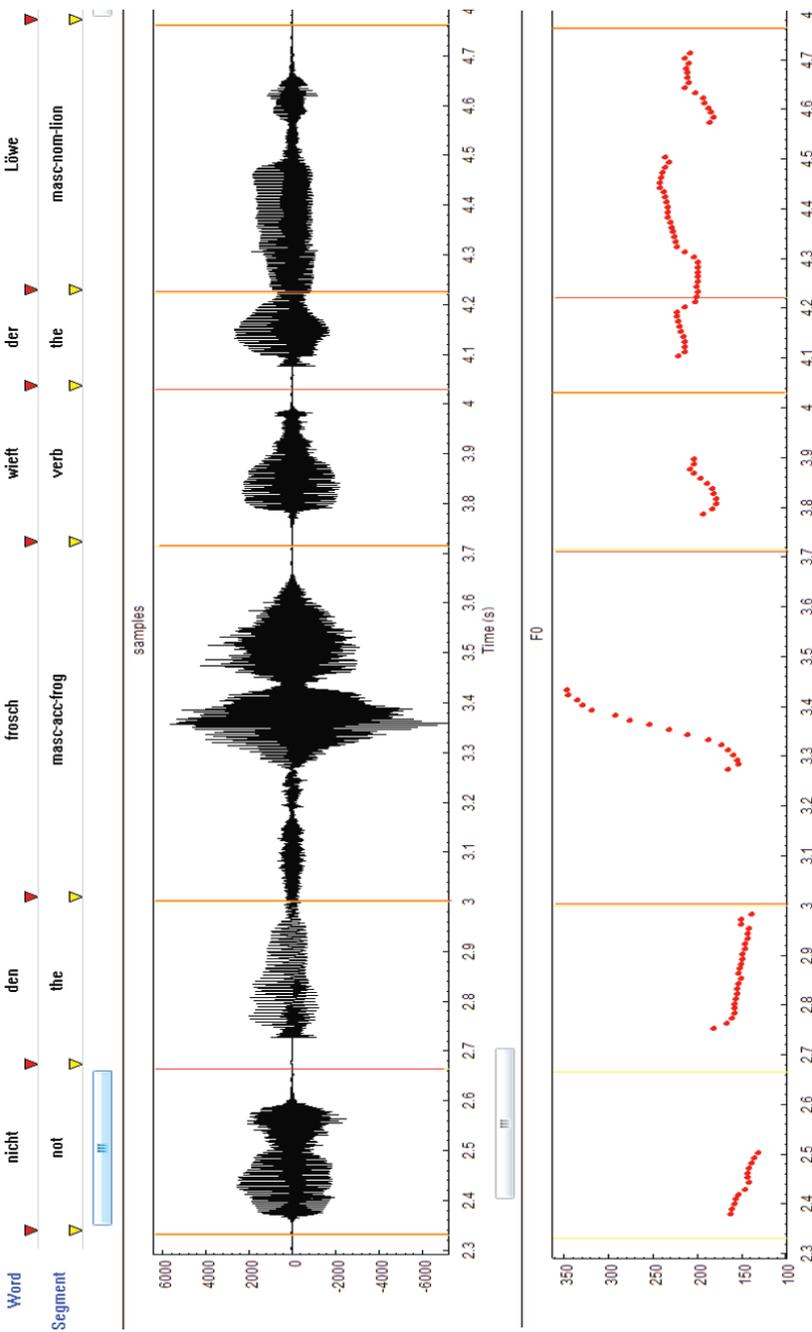
Figure 5. *Example of the intonation of the first part of the correcting utterance as in (5). The second part of the stimuli was recycled from Study 1 (see Figure 1a & Figure 1b).*

target verb in a transitive OVS argument structure, as in (9). After the videos had stopped, the prerecorded voice of the puppet asked the child to point to the correct (still) picture by asking, for example, *Zeig P2 das Bild: Den (+accusative) Löwen wieft der (+nominative) Hund!*—'Show P2 the picture: the (+accusative) lion is weefing the (+nominative) elephant'. If the child failed to point then the puppet repeated the question a second time, but she never asked the child to point again once she/he had already done so. Once the child had pointed to one of the two pictures, the next test trial began, preceded once more by the red dot.

3.2.   *Coding and reliability*

For every test trial, we coded responses for whether children pointed to the picture in which the postverbal, second noun in the sentence was the agent. If a child did not choose either scene (3), we coded those trials as 'wrong' (an alternative analysis in which these cases were excluded had no effect on the pattern of the results). For one participant, 6 trials were missing because of technical failure. In order to give all participants' data the same weight in the analyses, the remaining pointing values (= 10) were coded as the total score (= 100%) of this participant. All children were coded by the first author, and an additional coder coded 25% of all trials for reliability, revealing a high agreement with the first author (Cohen's Kappa = 0.969)

3.3.   *Results and discussion*

We again tested for the proportion of times the NP occurring after the verb was identified as the agent of the action out of four. The data were analyzed using a 2 (Intonation) × 2 (Case Marking) repeated measures analysis of variance (ANOVA). There were main effects for both Intonation, $F (1, 15) = 5.8$, $p = 0.029$ and Case Marking $F (1, 15) = 14.4$, $p = 0.002$, but no significant Intonation × Case Marking interaction ($F [1, 15] = 1.13$, $p = 0.304$) (see Figure 6).

   Because the chance level for our dependent variable was always 50%, we also investigated in which conditions the children were above chance. The results show that the children were above chance in the Case Marking / Contrastive Intonation condition ($t[15] = 4.0$, $p < 0.001$) as well as in the Case Marking / Neutral Intonation condition ($t[15] = 2.2$, $p = 0.044$). In the No Case Marking / Contrastive Intonation, children chose agents and patients at chance level ($t[15] < 0.001$, $p = 1.0$), whereas children in the No Case Marking / Neutral Intonation Condition relied solely on word order ($t[15] = −2.53$, $p = 0.023$).

   A comparison between the two conditions Case Marking / Contrastive Intonation and Case Marking / Neutral Intonation revealed no significant differ-
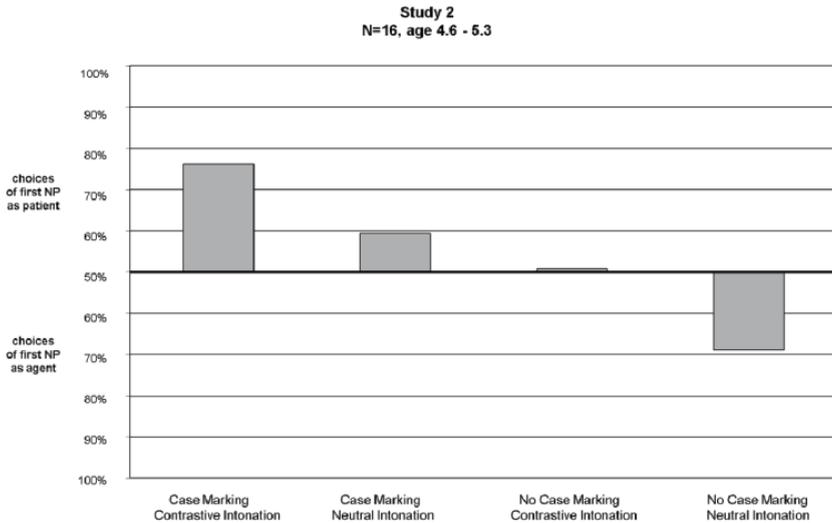
Figure 6. *Results of Study 2 in the four conditions. The diagram shows percentages of judging NP1 as either object-first or subject-first constructions.*

ence (paired-sample t-test: t[15] = 1.145, p = 0.270), whereas choices in the two conditions No Case Marking / Contrastive Intonation and No Case Marking / Neutral Intonation revealed a higher judgment of NP1 as the patient, when this interpretation was supported by intonational stress (t[15] = 3.0, p = 0.009). These results strengthen and extend those of Study 1. In this study, children used intonation, as opposed to word order, in interpreting patient-first transitive sentences. In other words, children used a high pitched accentuation of the first noun phrase to identify a patient-first transitive construction. This effect was especially clear in the two conditions without case marking, which showed that intonation by itself, in the absence of case marking, is a sufficient cue for young children to re-assess an agent-first interpretation. The two conditions with case marking, with and without intonation, did not differ, but they showed fairly high rates of success.

For a further analysis, we compared the results from Study 1 and Study 2 (see Figure 7).

The data were analyzed using a three-way mixed analysis of variance (ANOVA) with two repeated factors (Case Marking and Intonation) and one between-subjects factor (Context). There were main effects for both Intonation, (F[1,30] = 10.7, p = 0.03) and Case Marking, (F[1,30] = 52.0, p < 0.001), but no significant interaction between the two (F[1,30] = 0.3, p = 0.541). There was no significant interaction between Case Marking and Context (F[1,30] = 2.5, p = 0.118), or between Intonation and Context (F[1,30] = 0.2, p = 0.602), but
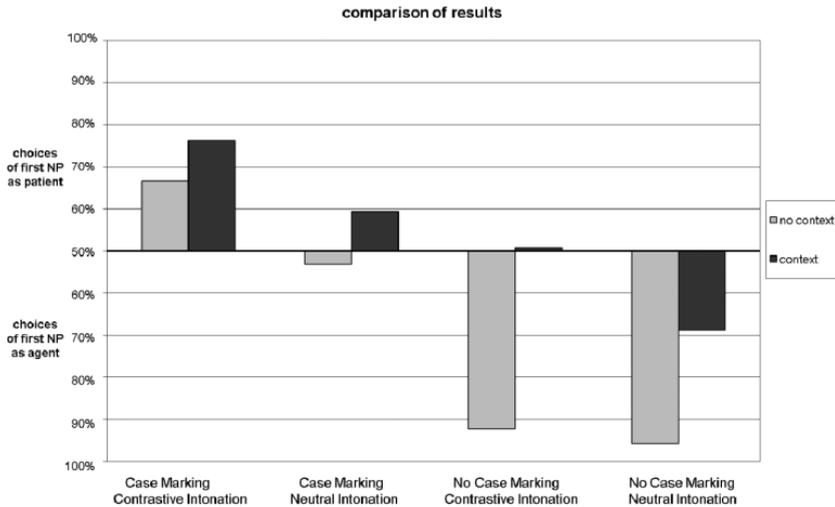
Figure 7.   *Comparison of results from Study 1 (with no context) and Study 2 (including context) in the four conditions. The diagram shows percentages of judging NP1 as either patient or agent as compared with chance, 50%.*

we found a significant interaction between all three factors, (F[1,30] = 4.4, p = 0.044). A comparison between conditions of the two studies revealed no significant difference either in the conditions Case Marking / Contrastive Intonation (paired-sample t-test: t[15] = 1.09, p = 0.285), or in the two Case Marking / Neutral Intonation conditions (t[15] = 1.72, p = 0.095). Only those choices in the two conditions No Case Marking / Contrastive Intonation (t[15] = 6.26, p < 0.001) and No Case Marking / Neutral Intonation (t[15] = 3.16, p = 0.005) revealed a significantly greater likelihood of judging NP1 as the patient, when this interpretation was supported by a combination of the prosodic pattern and the preceding context.

These results show the importance for children of a natural intonational realization in order to understand participant roles. Even when the morphosyntactic information in a construction is totally ambiguous, the intonational form of an utterance can pull children away from their strong word order bias. The results from Study 1 show that intonation is an important cue and helps children to understand agent and patient relations. But in isolation, without any help from other cues, the strong word order bias cannot be eliminated. If an appropriate context and intonational pattern are included (as for example that presented in Study 2), children can negotiate this bias and are moving towards ceiling levels when several cues are combined (i.e. case marking, intonation and discourse context).

4.   General discussion

In the current study we found that five-year-old German children recognize a high pitch accent on the initial noun phrase as a cue indicating a patient-first transitive construction. Thus, the prosodic cue is strong enough to pull children away from their strong word order bias whereby they interpret the first noun as an agent. In the first study, this effect could only be seen in combination with case marking. In those conditions where case marking was ambiguous, children still fell back on their most reliable cue—word order. This was consistent with the results from the control group. But, as opposed to adults who could also solve participant roles in sentences for which only one cue was available (Case Marking), it seems that children at the age of five could only use information provided by the intonational cue when this was supported by Case Marking. Thus there seems to be a developmental progression beyond the age of five. In the second study, where target sentences were presented in a more natural way with a combination of context and intonation, the results showed an even clearer pattern because young children used the intonational cue (in combination with case marking and context) rather than the competing cue of word order. In contrast to Dittmar et al.'s (2006) study, in which children of the same age systematically misinterpreted patient-first sentences, the children in our studies no longer depended on the most reliable cue—even in the absence of case marking. What this shows is that prosody has the power to work against this word order bias and that the information in the sound stream seems to be sufficiently rich to allow children to abstract participant roles.

The exact basis by which the children interpreted the prosodic cue remains as yet unknown. Focusing intonationally on certain words is a communicative function that serves to put emphasis on a particular part of an utterance. Varying widely across languages, it involves changes in duration, intensity, and vowel quality (e.g. Turk and White 1999; Xu and Xu 2005). However, the primary cue for perceiving focus is generally considered to be pitch variation (Dahan and Bernard 1996) and this was the cue that we controlled in our studies.

Compared to Dittmar et al.'s (2006) results, our findings from Study 1 are somewhat surprising. In the condition where case marking and word order contradicted each other, but none of the cues were reinforced by intonation (Case Marking / Neutral Intonation), our children chose participant roles at chance, whereas the children in Dittmar et al.'s (2006) study relied primarily on word order. In our opinion, this is due to the natural mechanisms of speech, both psychological as well as physiological. In our study, intonation was computerized and manipulated and thus controlled; i.e. in the neutral intonation condition, children heard a completely flat intonation pattern, whereas Dittmar et al.'s (2006) children were tested with a task in which by the experimenter uttered the target sentences in a live-situation. Even if the experimenter in that

study had concentrated on a neutral vocal production, natural tendencies like declination or macro- and micro-prosodic cues provide a minimal prosodic pattern that the children could have used to decide on the agent and patient roles. In addition, the accusative marker in our study could have been more clearly articulated (due to intonational prominence) and thus more clearly perceived, as compared to Dittmar et al.'s (2006) study.

Dittmar et al.'s (2006) corpus study of input to six children recorded for the first time at 1;8 and then again at 2;5 provides data for the frequency with which the types of sentences presented in our study occur in everyday speech around children. Overall, Dittmar et al. (2006) found 745 transitive sentences, 55% (410) of which had causative verbs. 21% of those involved conflicting (but unambiguous) case marking and word order (our Condition 1). More interestingly, only 2 sentences in the corpus appeared with an object-first order and ambiguous case marking (our Condition 3). This means that although less than 1% of all causative sentences that children hear in the input are constructions containing non-canonical word order and ambiguous case marking, the prosodic characteristics of exactly the same constructions lead children away from a word-order interpretation in our second study. In other words, despite the very low input proportions, children still manage to disambiguate these constructions when an intonational cue is present.

There are a number of possible explanations for these results, not necessarily mutually exclusive. It seems clear that the strong contextual cue in Study 2 provides the whole package in a more natural way and pulls the children towards an OVS interpretation. It is also possible that children could have learned the prosodic pattern associated with the patient-first transitive construction as a whole and abstracted a form-meaning mapping for the prosodic cue from the more frequent OVS causative constructions in the input which include case marking. However, the relatively weak results from the first study, especially in the conditions without case marking, would seem to argue against this hypothesis. It is also possible is that children are simply noticing an unusual prosodic pattern and are inferring that this suggests an unusual, marked interpretation, which they then need to guess from the various available options. One final possibility, which would give even deeper insight into the acquisition of prosodic meaning, is that children have come to understand more generally that new and 'special' information often stands in focus and receives prosodic stress. Thus it may be that by 5;0 children are in the process of abstracting a more general mapping from intonational prominence to sentential focus. This could be derived from simpler constructions. These might include OVS utterances which may well be learned as a whole together with their intonation (e.g. *DAS mag ich!*—'That I like!') as well as other syntactic constructions in which there is focal intonation such as imperatives (*Sitzt DA, nicht da!*—'Sit THERE, not there').

In line with this view, Grassmann and Tomasello (2007) demonstrated in a recent word learning study that 2-year-olds already know that those words in an utterance that correspond to contextually new referents (and are thus "special" within the discourse) are prosodically stressed. This suggests that children interpret prosodic stress in language as being iconic of the speaker's intention to refer to a salient aspect of the situation. Interestingly, we have shown, as did Grassmann and Tomasello (2007), that only a combination of newness (or salience) and stress together were effective. In our second study, where children used the prosodic cue much more effectively, the first noun phrase referred to the new participant in the situation, and critically, the contrast was with a participant who was the patient in the preceding discourse context. Furthermore, the linguistic material that is new, or in some sense contrastive, was prosodically highlighted compared to given or contextually available information. Indeed, it is not totally clear that these are separate hypotheses, as it is possible that even adult Germans use the intonation typically associated with patient-first transitive sentences in this more general way, rather than as part of the transitive construction as a whole.

In order to resolve syntactic ambiguities, children need sentences that contain multiple cues—according to Bates and MacWhinney's (1987) concept of coalitions-as-prototypes. What this means is that because sometimes several cues may indicate the same function—providing extra information—children should find it especially easy to comprehend prototypical transitive sentences, e.g. with both word order and case marking (and perhaps other cues) working in coalition.

This is consistent with our results. Whereas participants in the control group were able to solve participant roles in sentences with only one cue was available (Case Marking), it seems that children at the age of five need several, supporting cues. Our study adds the fact that children do not just use morphosyntactic cues like word order and case marking to disambiguate participant roles. Prosody, especially in combination with an appropriate context, is an important cue which in the absence of clear morphological cues can modulate subject and object assignment. Thus, the problem of processing sentences with non-canonical word order can be partially alleviated when these utterances are presented with the appropriate intonation and the appropriate context. From early on children can only interpret sentences which contain combinations of cues in the most frequently heard patterns. However, development consists in starting to identify the separate contribution of each cue. The present study indicates that, in line with usage-based approaches, both the context and sentential prosody should be treated as cues of considerable importance and investigated as such.

The present study is just a first step. It is likely that intonation interacts in complex ways with a number of different morphosyntactic cues, and indeed we

provide some evidence for this possibility. In some cases the prosodic pattern may be a part of the construction itself, whereas in other cases it may be being used more generally, for example as a contrast, in order to stress a particular noun phrase which then triggers a specific interpretation of a particular construction. But again, this may be a false dichotomy, as in many cases the distinction between these two interpretations is unclear—a good example being the English cleft construction, for example, "It was the DOG that got sick"; in this case the stress on *dog* could be interpreted by either route. In any case, the larger point is that to fully understand young children's skills at interpreting sentences online, the role of intonation and context must be taken into account.

References

Allbritton, David W., Gail McKoon & Roger Ratcliff. 1996. Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22. 714–735.

Arnold, Jennifer E. 2008. "THE BACON" not "the Bacon": How children and adults understand accented and unaccented noun phrases. *Cognition* 108. 69–99.

Bates, Elizabeth, Brian MacWhinney, Cristina Caselli, Antonella Devescovi, Francesco Natale & Valeria Venza. 1984. A cross-linguistic study of the development of sentence interpretation strategies. *Child Development 55*. 341–354.

Bates, Elizabeth & Brian MacWhinney. 1987. Competition, variation, and language learning. In Brian MacWhinney (ed.), *Mechanisms of language acquisition*. 157–193. Hillsdale, NJ: Erlbaum.

Bates, Elizabeth & Brian MacWhinney. 1989. Functionalism and the competition model. In Brian MacWhinney & Elizabeth Bates (eds.), *The crosslinguistic study of sentence processing*, 3–76. New York: Cambridge University Press.

Carlson, Katy, Lyn Frazier & Charles Clifton. 2009. How prosody constrains comprehension: A limited effect of prosodic packaging. *Lingua* 119. 1066–1082.

Chan, Angel, Elena Lieven & Michael Tomasello. 2009. Children's understanding of the agent-patient relations in the transitive construction: Cross-linguistic comparisons between Cantonese, German, and English. *Cognitive Linguistics* 20(2). 267–300.

Choi, Youngon & Reiko Mazuka. 2003. Young children's use of prosody in sentence parsing. *Journal of Psycholinguistic Research* 32. 197–217.

Dahan, Delphine & Jean-Marc Bernard. 1996. Interspeaker variability in emphatic accent production in French. *Language and Speech* 39(4). 341–374.

Dittmar, Miriam, Kirsten Abbott-Smith, Elena Lieven & Michae1l Tomasello. 2008. German children's comprehension of word order and case marking in causative sentences. *Child Development* 79(4). 1152–1167.

Grassmann, Susanne & Michael Tomasello. 2007. Two-year-olds use primary sentence accent to learn new words. *Journal of Child Language* 34. 677–687.

Grice, Martine, Stefan Baumann & Ralf Benzmüller. 2005. German Intonation in Autosegmental-Metrical Phonology. In Sun-Ah Jun (ed.) Prosodic typology: The phonology of intonation and phrasing, 55–83. Oxford: Oxford University Press.

Grice, Martine & Stefan Baumann. 2002. Deutsche Intonation und GToBI. *Linguistische Berichte* 191. 267–298.

Johnson, Elizabeth K. & Peter W. Jusczyk. 2003. Exploring possible effects of language-specific knowledge on infants' segmentation of an artificial language. In Derek Houston, Amanda Seidl, George J. Hollich, Elizabeth K. Johnson, & Ann Marie Jusczyk (eds.) *Jusczyk lab final report*, 141–148.

Mehler, Jacques, Peter Jusczyk , Ghislaine Lambertz, Nilofar Halsted, Josiane Bertoncini, & Claudine Amiel-Tison. 1988. A precursor of language acquisition in young infants. *Cognition* 29. 143–178.

Schafer, Amy J., Shari R. Speer, Paul Warren & David S. White. 2000. Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research* 29. 169–182.

Snedeker, Jesse & Sylvia Yuan. 2008. Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language* 58. 574–608.

Speer, Shari R., Paul Warren & Amy J. Schafer. 2003. Intonation and sentence processing. *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona*. Rundle Mall: Causal Productions, 95–105.

Taylor, John R. 2002. *Cognitive Grammar*. Oxford: Oxford University Press.

Turk, Alice E., & Laurence White. 1999. Structural influences on accentual lengthening. *Journal of Phonetics* 27. 171–206.

Vogel, Irene & Eric Raimy. 2002. The acquisition of compound vs. phrasal stress: The role of prosodic constituents. *Journal of Child Language* 29. 225–250.

Weber, Andrea, Martine Grice & Matthew W. Crocker. 2006. The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition* 99(2), B63–B72.

Xu, Yi, & Ching X. Xu. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33. 159–197.