Cognition 221 (2022) 104988

Contents lists available at ScienceDirect

Cognition

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

German-speaking children use sentence-initial case marking for predictive language processing at age four

Duygu Özge [a,e,*,1], Jaklin Kornfilt [b], Katja Maquate [c], Aylin C. Küntay [d], Jesse Snedeker [e]

[a] Middle East Technical University, Turkey
[b] Syracuse University, United States of America
[c] Humboldt University of Berlin, Germany
[d] Koç University, Turkey
[e] Harvard University, United States of America

ARTICLE INFO

Keywords:
Child language processing
Development of parsing abilities
Acquisition of German case marking
Predictive use of case marking
Incremental processing in children

ABSTRACT

Adults incrementally integrate multiple sources of information to predict the upcoming linguistic structure. Although we have substantial evidence that children can use lexicosemantic information triggered by the verb, we have limited information as to whether children can use morphosyntax to generate predictions during the course of processing. Previous studies show that four-year-old Turkish-speaking children can use case-marking cues predictively; however German-speaking children have been reported to fail until late in development. The present visual-world eye-tracking study provides the first evidence from four-year-old German-speaking children (mean age: 4;03) interpreting sentence initial case marking cues independent of the identity of the verb and the canonical word order to predict the thematic role of the upcoming argument. We presented children with a visual context with a stereotypical but ambiguous event, the thematic structure of which can be resolved only on the basis of the case marking cues on subject-initial and object-initial structures locating the verb sentence-finally. Children were able to use the accusative case on the non-canonical object-initial utterances to predict that the upcoming argument should have the agent role before this argument and the verb became available. This study shows that the previously reported discrepancy between these two case-marking languages (i.e., Turkish and German) is not due to the crosslinguistic differences but due to methodological differences employed across studies. These findings provide support for language acquisition theories assuming early abstractions and adult-like parsing mechanisms predictively integrating multiple sources of cues.

1. Introduction

Adults process language incrementally, which allows them to quickly predict the upcoming structure of an utterance (e.g., Staub & Clifton Jr., 2006). For instance, they can use the lexical associations between nouns to activate the knowledge of events featuring typical agents and patients (Ferretti, McRae, & Hatherell, 2001; McRae, Hare, Elman, & Ferretti, 2005) as well as word order cues (Knoeferle, Crocker, Scheepers, & Pickering, 2005). They can also use the verbs and a supportive visual context to anticipate the thematic roles of an upcoming referent (Altmann & Kamide, 1999; Altmann & Mirković, 2009; Ferretti et al., 2001; Kamide, Altmann, & Haywood, 2003; Kamide, Scheepers, & Altmann, 2003; McRae et al., 2005). Furthermore, adults are also very good at determining the thematic roles of arguments in a sentence using other linguistic cues such as case marking (cf. Kamide, Altmann, & Haywood, 2003; Kamide, Scheepers, & Altmann, 2003; Özge, Küntay, & Snedeker, 2019). Hence, incrementally-built linguistic expectations of agents and patients are not solely based on the activation of lexically-related items, conceptual expectations or the order of constituents in a sentence (e.g., SVO - Subject Verb Object) but they also arise with the parsing of structural cues (Kamide, Altmann, & Haywood, 2003; Kamide, Scheepers, & Altmann, 2003). That is, as adults, we are able to form thematic relations in an utterance using our conceptual representations and structural abstractions. However, for children, there is still limited evidence about whether or not they share similar processing mechanisms as adults and whether they are as effective as adults in

* Corresponding author at: Middle East Technical University, Language and Cognitive Development Lab, Turkey.
E-mail address: duyguo@metu.edu.tr (D. Özge).
[1] Duygu Özge was a Marie Curie Fellow at Harvard University Lab for Developmental Studies during the design and data collection of this study.

https://doi.org/10.1016/j.cognition.2021.104988
Received 24 July 2020; Received in revised form 14 August 2021; Accepted 5 December 2021
Available online 23 December 2021

Table 1
Comparison of Turkish and German case systems.

	German	Turkish
Type of morphological marking	Weakly inflecting	Agglutinating (strong inflection)
Means of expression of case	Independent function word	Inflectional
Location where it is marked	On the determiner and modifier (and sometimes additionally on the head noun)	On the head noun
Length and ambiguity of inflected forms	Short: Multiple grammatical features (e.g., case, gender, number) are fused together as one form resulting in shorter but more ambiguous forms	Each grammatical feature is encoded with a different inflectional morpheme resulting in longer but more reliable (because more consistent and more transparent) inflected forms

incrementally integrating linguistic and nonlinguistic cues during the course of spoken language interpretation (Snedeker, 2009; Snedeker & Huang, 2015).

Children, by the age of four, already seem to comprehend most of the linguistic input (Snedeker, 2009; Snedeker & Huang, 2015). Similar to adults, they can incrementally use phonological information to predict upcoming lexical items (e.g., Fernald, Swingley, & Pinto, 2001; Swingley, Pinto, & Fernald, 1999), they can use the verb's selectional restrictions to resolve structural ambiguities (Snedeker & Trueswell, 2004; Trueswell, Sekerina, Hill, & Logrip, 1999), and they can rapidly anticipate upcoming arguments on the basis of lexical associations (Borovsky, Elman, & Fernald, 2012; Mani & Huettig, 2012; Nation, Marshall, & Altmann, 2003). We also know from languages like Turkish that children can use case marking as a cue to transitivity (Göksun, Küntay, & Naigles, 2008), they can incrementally react to case marking ambiguities and process morphosyntactic dependencies triggered by case marking cues (Özge, Marinis, & Zeyrek, 2015). Recently, Turkish-speaking four-year-old children have been shown to incrementally parse nominative and accusative case markers to predict the thematic role of the upcoming argument independent of the verb both in canonical and noncanonical word orders (Özge et al., 2019), suggesting that children may rely on early and effective mapping between the case markers and thematic roles.

However, it is not clear whether this is a phenomenon we can cross-linguistically generalize to other case-marking languages. We know, for example, from previous studies that German case markers are acquired late and German-speaking children cannot use case markers for thematic interpretation independent of the word order cues until age six or seven (Dittmar, Abbot-Smith, Lieven, & Tomasello, 2008). This difference poses a cross-linguistic puzzle. This could be due to cross-linguistic differences in the reliability of the case marking system in these two languages so children acquiring a language with a more reliable system (e.g., Turkish) may be forming abstractions earlier than children acquiring a less reliable system (e.g., German) (Dittmar et al., 2008; Slobin & Bever, 1982). However, this could also be due to the methodological differences in the previous online studies. The present study aims to solve this puzzle using the same task as the one used in Özge et al. (2019) with German-speaking children to test whether German children can integrate case marking incrementally to predict the role of the second argument before the verb becomes available in verb-final structures. This will allow us to understand the source of the differences observed so far in the literature. If the source of these differences is crosslinguistic in nature (i.e., relatively less transparent case system in German), then we expect to observe difficulty on the part of German children with respect to incremental interpretation of case marking in object-initial (vs. subject-initial) orders. On the other hand, if the pattern of development is the same across languages and the previous disparities observed in Turkish and German are due to methodological differences, we expect German children to interpret the case markers predictively. In the rest of the introduction, we first discuss the case marking system in German versus Turkish, we review previous studies on the acquisition of case marking in German with a focus on their methodology, and we present the aim and the theoretical motivation of our study.

1.1. Case marking in Turkish versus German

Case is a syntactic as well as morphological notion; as an abstract syntactic notion, it helps determine certain word order phenomena (see, for example, Chomsky, 1981 for a systematic proposal of syntactic Case, following Vergnaud, 1977). More traditionally, case is a morphological reflection of the grammatical relations that noun phrases have to the verb in a sentence (Chomsky, 1965; Siegel, 1974). In this paper, we will limit attention to morphological case, since it is the overt morphological marker that is heard and interpreted by the hearer.

The fact that German and the Turkish case systems are similar in that both languages have lexical cases as well as structural cases,[2] may lead one to expect that the acquisition of these cases as well as their uses for predictive purposes would be similar. However, previous acquisition studies have shown that children acquiring German lag behind children acquiring Turkish, when it comes to putting case morphology to incremental use in syntactic comprehension (e.g., Dittmar et al., 2008; Ketrez & Aksu-Koç, 2009; Kröger, Münster, Burigo, & Knoeferle, 2018; Kröger, Münster, & Knoeferle, 2017; Maquate & Knoeferle, 2021; Özge et al., 2019; Schipke, Friederici, & Oberecker, 2011; Zhang & Knoeferle, 2012). This observation may have to do with the fact that case morphology in Turkish is more transparent than case morphology in German (Dressler, 2019; Stephany & Voeikova, 2009; Xanthos et al., 2011).

Table 1 summarizes the comparison of Turkish and German case systems on the basis of the type of morphological marking, means of expression of case, location where it is marked, and length and ambiguity of the inflected forms.

In Turkish each case is expressed by a dedicated morpheme with the exception of the nominative, which is a null morpheme. Turkish case is agglutinated on the head noun of a noun phrase and it is not expressed on modifiers or determiners. Furthermore, each grammatical feature is encoded with a different inflectional morpheme (i.e., there is no fusion of multiple grammatical features into one single form) resulting in longer but more reliable inflected forms. These features render Turkish case system more transparent. The surface forms of case morphemes do not show much variation except for some cases of phonological variation such as vowel harmony or consonant alternations. We illustrate these properties via two nouns (cat/bird) in (1), with stems exhibiting different (final) vowels and differing in having or lacking a stem-final consonant; we include the indefinite article (*bir* 'a'), and an adjectival modifier (*küçük* 'small'), to show lack of concord within the noun phrase:

[2] Lexical case is the kind of case that is determined by a particular lexical item, such as a verb (e.g., the verb 'to help' licenses the dative case on its complement in Turkish and German). Lexical case remains invariant irrespective of the voice of the predicate (i.e., active or passive). This contrasts with structural case such as the accusative case that alternates with the nominative case when the voice of the predicate changes (i.e., the referent is marked in the accusative case if the predicate is in the active voice, while the same referent is marked in the nominative case when that referent appears in the passive voice both in German and Turkish).

(1)

a. Nominative

Küçük bir kedi(−ø)	ve	küçük bir kuş(−ø)	gel-di.
Small a cat	and	small a bird	come-past.3sg

'A small cat and a small bird came.'

b. Accusative

Küçük bir kedi-yi	ve	küçük bir kuş-u	sev-di-m.
Small a cat-Acc	and	small a bird-Acc	love-Past.1sg

'I stroked a small cat and a small bird.'

c. Dative

Küçük bir kedi-ye	ve	küçük bir kuş-a	mama	ver-di-m.
Small a cat-Dat	and	small a bird-Dat	food	give-Past.1sg

'I gave food to a small cat and a small bird.'

d. Genitive

Küçük bir kedi-nin	ve	küçük bir kuş-un mama-sı	bura-da.
Small a cat-Gen	and	small a bird-Gen food-Poss.3s	here-Loc

'The food of a small cat and of a small bird is here.'

e. Locative

Küçük bir kedi-de	ve	küçük bir kuş-ta	parazit	var.
Small a cat-Loc	and	small a bird-Loc	parasite	exist

'There exists parasite in a small cat and a small bird.'

f. Ablative

Küçük bir kedi-den	ve	küçük bir kuş-tan	kan	al-dı-lar.
Small a cat-Abl	and	small a bird-Abl	blood	take-Past3pl

'They took a blood sample from a small cat and a small bird.'

The case morphology in German is less transparent because it is less consistent with respect to contrasts among the case morphemes and multiple grammatical categories (e.g., case, gender, number) are fused together as one form. Furthermore, the case morpheme is not always placed on the head noun of a noun phrase. For example, when a neuter noun heads a noun phrase, only the genitive case morpheme is realized on the head noun (in addition to also being reflected on the determiner and an adjectival modifier); all other cases are realized on the determiner and the modifier. In addition, and importantly for our purposes, the nominative and the accusative morphology on the determiner and the modifier are identical, when the head noun is a neuter noun (cf. Haider, 2010: 237).

(2)

a. Nominative

Ein-e	klein-e	Katze	und	ein	klein-er	Vogel
a-Fem. Nom	small-Fem. Nom	cat	and	a-Masc. Nom	small-Masc. Nom	bird

sind	gekommen
have	come

'A small cat and a small bird have come.'

b. Accusative

Ich	habe	ein-e	klein-e	Katze	und	ein-en
I	have	a-Fem.Acc	small-Fem.Acc	cat	and	a-Masc.Acc

klein-en	Vogel	gestreichelt
small–Masc.Acc	bird	petted

'I have petted a small cat and a small bird.'

c. Genitive

Das	Futter	der	klein-en	Katze und des
The	food	the-Fem.Gen	small-Fem.Gen	cat and the-Masc.Gen

klein-en	Vogel-s	sind	hier.
small-Masc.Gen	bird-Masc.Gen	are	here

'The food(stuff) of the small cat and of the small bird are here.'

d. Locative and Ablative: German has no morphological ablative or locative on noun phrases; the meanings of those cases are expressed via prepositions.

In (2a), the noun phrase headed by the feminine noun Katze 'cat' is the subject, while in (2b), the same noun phrase is the direct object although the shape of the determiner and of the adjectival participle are identical in the two examples. The same is also true for neuter nouns (3).

(3)

a. Nominative

[*Das*	faszinierende	Buch]	kostet	nur	zehn	Dollar.
the.Neut.Nom	fascinating-Nom	book	costs	only	ten	dollar

'The unbelievably fascinating book costs only ten dollars.'

b. Accusative

Ich	habe	[*das*	faszinierende	Buch]	gelesen.
I	have	the.Neut.Acc	fascinating-Acc	book	read

'I have read the fascinating book.'

Only when the head noun is masculine do the determiner and the modifier differ in their appearance, depending on whether they express nominative or accusative case, as in (4).

(4)

a. Nominative

[*Der*	faszinierende	Mann]	ist	mein	Nachbar.
the.Mas.Nom	fascinating-Nom	man	is	my	neighbor

'The unbelievably fascinating man is my neighbor.'

b. Accusative

Ich	habe	[*den*	faszinierenden	Mann]	kennengelernt.
I	have	the.Mas.Acc	fascinating-Acc	man	met

'I have met the unbelievably fascinating woman.'

However, even here, there is no total transparency and differentiation, because the masculine modifier in the nominative is identical to the modifier of a feminine noun as well as to the modifier of a neuter noun, in both nominative and accusative contexts.

1.2. Previous studies on case marking in different word orders in German children

Even though German word order is considered to be relatively free and both object and subject-initial sentences in German are grammatical, subject-initial sentences are canonical and more frequent while object-initial sentences are non-canonical and less frequent (Bornkessel, Schlesewsky, & Friederici, 2002; Hoberg, 1981; Kempen & Harbusch, 2005) (Table 2).

Adults can understand and interpret unambiguous object-initial sentences with ease, yet when object case marking is ambiguous between nominative and accusative, sentences are initially interpreted as having a subject-initial (vs. object-initial) structure, hence showing a clear subject-initial preference (Hemforth & Konieczny, 2000). German children have consistently been reported to have difficulty assigning thematic roles in object-initial sentences until late in acquisition. For instance, Dittmar et al. (2008) tested two-, four-, and seven-year-old children regarding their abilities to assign agent and patient roles.

Table 2
Frequency of subject-initial versus object-initial structures in German.

Argument Ordering	Hoberg (1981)	Kempen and Harbusch (2005)
Subject-before-Accusative-Object	86.1%	86.1%
Accusative-Object-before-Subject	0.5%	0.2%

They manipulated the word-order by changing the case-marking cues on the nouns and created German SVO and OVS sentences with novel verbs. In the conditions used, word order and case-marking could either support each other (i.e., being unambiguous and canonical), as in (5a) or word order or case-marking could be the only cue available for the thematic role assignment, as in (5b) and (5c), respectively.

(5)

a. Condition 1: Canonical word-order & unambiguous case marking

Guck mal,	der	Löwe	wieft	den	Hund
	the.Mas.Nom	lion-Mas.Nom	weef	the.Mas.Acc	dog-Mas.Acc

'Look, the lion is weefing the dog.'

b. Condition 2: Canonical word-order & ambiguous case marking.

Guck mal,	die	Katze	wieft	die
	the.Fem.Nom/Acc	cat-Fem.Nom/Acc	weef	the.Fem.Acc/Nom
Maus				
mouse-Fem.Acc/Nom				

'Look, the cat is weefing the mouse.'

c. Condition 3: Noncanonical word-order & unambiguous case marking

Guck mal,	den	Bären	wieft	der	Hase
	the.Mas.Acc	bear-Mas.Acc	weef	the.Mas.Nom	hare-Mas.Nom

'Look, the bear is weefing the hare.'

The children in this study watched two videos played side by side. The two videos showed a potential agent and a potential patient involved in a novel action denoted by a novel verb (e.g., *wiefen*). Both videos featured the same actions but the thematic agent and patient roles were reversed. While watching the videos, children listened to an utterance in one of the three conditions and then they pointed to the video that matched the spoken sentence. Two-year-olds only correctly assigned the thematic roles when SVO word order and unambiguous case marking supported each other (Condition 1). However, they could not determine the agent and patient when either the word order (Condition 2) or the case-marking (Condition 3) were the only cues to thematic role assignment. As for the four-year-old children, they primarily relied on canonical word order (SVO) information. Their accuracy for matching the sentence with the correct video was at ceiling for the word order only condition (Condition 2) and for the condition in which both word order and case-marking supported each other (Condition 1). Yet, their accuracy only reached chance level in the case-marking only condition (Condition 3). The seven-year-olds, on the contrary, performed similarly to adults for all sentence types. The results thus indicate that German children only after around the age of seven were able to determine agent and patient roles using case-marking regardless of the constituent order.

The offline results by Dittmar et al. (2008) are supported by electrophysiological (EEG) evidence. Schipke et al. (2011)'s results suggest that German children struggle processing non-canonical OVS (vs. SVO) sentences. In an EEG study, they presented three-, four-, and six-year-old German children with grammatically correct and incorrect transitive OVS and SVO sentences. The incorrect sentences included double-nominative and double-accusative sentences, as in (6a) and (6b), respectively.

(6)

a. Condition 1: Double-nominative sentence

*Der	Tiger	küsst	der	Frosch.
the.Mas.Nom	tiger-Mas.Nom	kiss	the.Mas.Nom	frog-Mas.Nom

'The tiger kisses the frog.'

b. Condition 2: Double-accusative sentence

*Den	Tiger	küsst	den	Frosch.
the.Mas.Acc	tiger-Mas.Acc	kiss	the.Mas.Acc	frog-Mas.Acc

'The tiger kisses the frog.'

While the double-nominative sentences indicated an adult-like result pattern in all age groups (i.e., a LAN-P600 pattern in response to the second NP), the double-accusative incorrect sentences showed distinct ERP responses time-locked to the second NP compared to their grammatically correct OVS counterpart sentences for each age group: three-year-olds showed an early positivity between 200 and 500 ms, the -year-old children showed a negativity between 1200 and 1300 ms and the six-year-olds showed two negativities between 600 and 800 ms and 1300–1600 ms. The authors interpret these results as supporting the developmental process for the acquisition of the German case-marking system: ERPs indicate a shift from error detection in the three-year-olds to thematic repair strategies in the middle age group. At six years of age, children's ERP responses were similar to adults' ERP responses associated with thematic/syntactic repair strategies (Schipke et al., 2011).

A recent eye-tracking study also supports the findings that four- and five-year old children still struggle interpreting the case-marking cues to determine the agent and patient in non-canonical OVS sentences. Kröger et al. (2017) presented four- and five-year old children with ambiguous action scenes while they listened to unambiguous German SVO (e.g., *The elephant*$_{\text{subject\_agent\_NOM}}$ *draws immediately the cheetah*$_{\text{object\_patient\_ACC}}$) and OVS (e.g., *The elephant*$_{\text{object\_patient\_ACC}}$ *draws immediately the donkey*$_{\text{subject\_agent\_NOM}}$) sentences. In the scenes, actions such as drawing were depicted but did not give away specific role relations, since two characters performed identical actions and could both be the potential agent of the action denoted by the verb, e.g., the donkey was drawing the elephant and this elephant was drawing the cheetah. The scenes moreover did not contain stereotypical world knowledge regarding who is doing what with whom (i.e., donkeys do not typically draw elephants). Instead, Kröger et al. (2017) manipulated the prosodic contour of the SVO and OVS sentences to investigate if children's (and adults') thematic role assignment difficulties in OVS sentence could be alleviated by a supporting prosodic contour. For OVS sentences, the stress was put on the case-marked first NP, i.e., the patient; for SVO sentences, the stress was put on the verb (following Weber, Grice, & Crocker, 2006). Participants answered a post-trial comprehension question about who-does-what-to-whom. In contrast to adults who reliably determined the thematic roles offline, children scored at chance answering the comprehension questions, regardless of the prosodic contour. The eye-tracking data moreover revealed that on-line, neither children nor adults used prosody to determine the thematic roles of the arguments. Additionally, in the OVS sentences, children (in contrast to adults) did not anticipate the correct agent and hence were not able to assign the thematic roles in unambiguously case-marked sentences.

Taken together, these results indicate that German children under the age of six do not seem to have acquired the competence to determine the agent and patient thematic roles using only the case marking as a cue to thematic role assignment in non-canonical object initial sentences. Yet, languages are not acquired in isolation from world-knowledge and

visual contextual information. As such, language processing seldom happens in isolation. A couple of studies suggest that four- and five-year old German children can indeed determine the thematic roles of (un) ambiguously case-marked non-canonical OVS sentences where a rich visual context acts as reliable cue in addition to case-marking.

In a visual world eye-tracking study, Maquate and Knoeferle (2021) (cf., Zhang & Knoeferle, 2012 for a similar design and results) presented four- and five-year old German children and adults with agent – patient – agent scenes and unambiguous OVS sentences about who-is-doing-what-to-whom. The scenes depicted two potential agents and a patient, i.e., a happy-looking mermaid, a neutral-looking elephant and a grumpy-looking eagle. In half of the scenes, the potential agents (i.e., the mermaid and the eagle) would perform a visually depicted action, i.e., giving a present and handcuffing, towards the patient. In the other half of the trials, no actions were depicted and the three characters simply stood next to each other. In addition to the action cue, participants were also primed either with a happy or a sad looking speaker face. After 2 s of scene preview time, participants listened to the positively emotionally valenced and unambiguously case-marked German OVS sentence describing the scene (e.g., *The elephant*$_{\text{object\_patient\_ACC}}$ *happily hugs the mermaid*$_{\text{subject\_agent\_NOM}}$) while looking at the scene. Participants answered post-trial comprehension questions about who-does-what-to-whom. At issue was if and how visually depicted actions and emotional speaker facial expressions can alleviate children's (and adults') real-time processing difficulties with unambiguous non-canonical German OVS sentences. The results showed that emotional speaker face did not influence children's anticipatory fixations and post-trial answers. However, when the actions denoted by the verbs were (vs. were not) depicted in the scenes, children (and adults) quickly anticipated the correct agent (i.e., the subject) of the sentence. Moreover, children answered significantly more comprehension questions correctly when actions were (vs. were not) depicted. When no actions were depicted, they scored below chance. These results demonstrate that children can assign thematic roles during real-time processing in unambiguously case-marked OVS sentences when the visual context is clearly supporting the linguistic context (e.g., as is the case for depicted actions and verbs denoting the depicted action vs. less directly referential cues such as emotional facial expressions or prosodic cues).

Another eye-tracking study using depicted actions and wiggling agent characters as cues but ambiguously case-marked German OVS sentences underlines this assumption. Kröger et al. (2018) presented children and adults with visual scenes depicting three characters. The middle character was role ambiguous; it could be the agent or the patient of an action. One outer character was another potential agent of an action and the other outer character was a potential patient of an action (cf. Kröger et al., 2017). Participants listened to ambiguously case-marked German OVS sentences (e.g., *Das*$_{\text{DET\_Acc/Nom}}$ *Käferchen* $_{\text{N\_ Acc/Nom}}$ *schubst*$_{\text{Verb}}$ *gerade der* $_{\text{DET\_Nom.}}$ *Stier* $_{\text{N\_ Nom}}$, transl.: 'The little bug is pushing the bull.') while they looked at the visual scenes. At the onset of the verb region in the sentence, either the action denoted by the verb in the sentence would appear between the agent and the patient on the screen, or the agent character would wiggle up and down, or both the action appeared and the agent wiggled. The presentation of the visual cues was short-lived, i.e., the cues were only presented for the duration of the verb and disappeared again thereafter. Participants answered post-trial comprehension questions for who-is-doing-what-to-whom. Results suggested that children's preference to anticipate the agent (vs. the patient) was boosted by the short-lived wiggle but not by the short-lived action depiction. However, the post-trial answers indicated no advantage in determining thematic role relations in ambiguously case-marked OVS sentences for either the depicted action or the wiggling agent.

The reviewed results hence suggest that four- and five-year old German children still need additional non-linguistic cues in order to determine thematic roles by means of case-marking as the only linguistic cue. In these studies, the verb together with the depicted action denoted by the verb provided a powerful (non-)linguistic cue for determining thematic roles. Even though nouns can prime prototypical noun role fillers in adults (cf. Ferretti et al., 2001; Ferretti, Kutas, & McRae, 2007) and children seem to be able to draw on direct referential cues (depicted actions) for thematic role assignment, it is less clear if German children can also make thematic predictions independent of the verb and word order cues.

1.3. Aim and theoretical motivation of the present study

The present study addresses this question and tests whether German children can determine thematic roles in non-canonical unambiguous object initial sentences when (i) the verb is placed at the end of the sentence and (ii) the case marking on the first noun is the only cue along with a visual context relying on children's world knowledge as to who could be plausible agents or patients.

A recent visual-world study has shown that this is possible for Turkish children (Özge et al., 2019). In this study, four-year-old Turkish-speaking children viewed a visual scene with three referents (e.g., a mouse, a cat, and a piece of cheese) who could act on each other as plausible agents (i.e., a cat, a mouse) or plausible patients (a mouse, a piece of cheese) while listening to an utterance that is manipulated with respect to the case marking on the first and the second noun, as in (7). Children anticipated the second referent by shifting their gaze to the correct referent during the modifier region (i.e., 'over there'), which is a sign that they incrementally interpreted the case marker independent of the verb information.

(7)

a. Fare **Mouse-Nom**	birazdan shortly	şuradaki over there	peynir-**i** cheese-**Acc**	bulacak. find-Fut

'The mouse will shortly find the cheese over there.'

b. Fare-**yi** mouse-**Acc**	birazdan shortly	şuradaki over there	kedi cat-**Nom**	bulacak. find-Fut

'The cat over there will shortly find the mouse.

Understanding whether young children can assign an interpretation to morphosyntactic cues during online processing is crucial to our understanding of the nature of language development. If children begin the task of language acquisition with early conceptual and structural abstractions, they would rapidly establish mappings between the morphosyntactic cues and their meaning (henceforth: early abstraction accounts). To be able to do that, children might be innately endowed with basic conceptual expectations (e.g, every event has an agent or a cause) and the structural rules that are common to all languages (e.g., syntactic categories and the ways in which they may compose) (Fisher, 2002; Gleitman, 1990; Pinker, 1984). However, they would also be able to reach broad abstractions with innate conceptual representations and statistical abilities that derive abstractions from the input regularities to derive structural abstractions. On the other hand, if children did not start with early abstractions and if they learned the morphosyntax from the input in an item-based manner, their interpretation of morphosyntax would largely be verb-dependent and construction-dependent until they have gathered enough evidence from the input to form generalizations about the meaning of these structures (Tomasello, 2000) (henceforth: late abstraction accounts). The study with Turkish children showed that children were adultlike in their predictive interpretation of case markers as early as four years of age, which was taken as support for the early abstraction accounts (Özge et al., 2019). However, given the previous findings in German it is not clear whether these results can be generalized cross-linguistically. Late abstraction accounts would expect different patterns of acquisition depending on the cross-linguistic features of the linguistic system to be acquired. For instance, Turkish word order is much more flexible and dependent on case-marking compared

to the German, and the Turkish case system is more transparent, which might facilitate the early acquisition and processing of case marking.

On the other hand, the previously noted discrepancies between German and Turkish children may also be due to methodological differences in the relevant studies, rather than cross-linguistic differences. Previous studies in German come either (i) from offline studies that depend heavily on storing the information of who did what to whom in short term memory; (ii) from online studies that rely on violation paradigms combining grammatical structures with ungrammatical ones; (iii) from online studies that rely heavily on additional visual contextual information such as action depictions, and (iv) from studies that used non-stereotypical events. Both ungrammatical and non-stereotypical utterances deviate from what children usually hear in real life.

In the present study, we employed the same task as in the Turkish study by Özge et al. (2019) that situated children within a visual context with a more stereotypical event (e.g., cats finding mice, mice finding cheese) and an event structure puzzle that could only be solved with the interpretation of the case markers, as in (7) above. If German-speaking children behaved similarly to Turkish-speaking children and interpreted the case marker on the sentence initial noun to determine its thematic role without the verbal cue, they should be able to launch predictive looks to the second noun after the first noun and before the second one and the verb. This would then be a further support for early abstraction accounts and lead us to suggest that the previous discrepancies between Turkish and German children were largely due to methodological differences. On the other hand, if German-speaking children failed to show predictive effects in the critical time windows, then this would support the late abstraction accounts and lead us to suggest that the previous discrepancies were due to cross-linguistic differences.

2. Experiment

2.1. Method

2.1.1. Participants

We tested 20 four-year old German-speaking children (9 female, age range: 4;0–4;06, mean age: 4;03). Data of one of the participants was excluded from analysis as this participant looked at the empty parts of the screen throughout the experiment. This left us with 19 participants whose data we included in the analysis. Our participants were monolingual and they attended a kindergarten in Bielefeld, Germany. All children were typically developing with no history of language impairments. All had normal or corrected-to-normal vision and attended a kindergarten in Bielefeld.

The parents gave informed consent and the children were also asked for their consent at the time of testing. The children were tested in the childcare facility they were attending. The study was approved by the University Ethics Committee.

2.1.2. Stimuli

The stimuli were adapted from Turkish into German, to match the stimuli used in Özge et al. (2019) as much as possible. However, due to the ambiguity between neuter and feminine gender with nominative and accusative case marking in German, we replaced all nouns that, translated from Turkish to German, were neuter or feminine with masculine nouns (e.g., *die Möhre,* 'the carrot', became *der Kohl,* 'the cabbage'). Additionally, instead of a prenominal relativizer used in the Turkish items, we used the future tense auxiliary *wird* ('will') in the sentence's second position, to place the main (untensed) verb in sentence-final position and inserted a time adverbial *im nächsten Moment* ('in the next moment') to provide enough time between the first and the second noun phrase for children to anticipate the target. We constructed 18 critical items (9 per condition) and 8 filler items. The critical items were verb-final mono-transitive simple sentences with two arguments. The first argument (i.e., the first noun; henceforth: NP1) was either in the nominative case using the subject-object-verb (SOV) order, i.e., the nominative condition (8a), or in the accusative case using the object-subject-verb (OSV) order, i.e., the accusative condition (8b) (Appendix 1 for the stimuli).

(8)

a. Condition 1: Nominative marked sentence initial noun (SOV):

Der	Hase	wird im nächsten Moment	**den**	Kohl
The.Nom	Hare-Nom	will in the next moment	**the.Acc**	cabbage-Acc

aufspüren, nicht wahr?
find, not true?
'The hare will find the cabbage in the next moment, right?'

b. Condition 1: Accusative marked sentence initial noun (OSV):

Den	Hasen	wird im nächsten Moment	**der**	Fuchs
The.Acc	Hare-Acc	will in the next moment	**the.Nom**	fox-Nom

aufspüren, nicht wahr?
find, not true?
'The fox will find the hare in the next moment, right?'

Different from most of the previous studies that used either verb-medial structures with nonsense verbs or verb-final structures with embedded structure (eye-tracking: Kröger et al., 2017, 2018; Zhang & Knoeferle, 2012; Knoeferle & Crocker, 2006; ERP: Schipke, Knoll, Friederici, & Oberecker, 2012; Strotseva-Feinschmidt, Schipke, Gunter, Brauer, & Friederici, 2019; fMRI: Knoll, Obleser, Schipke, Friederici, & Brauer, 2012; off-line methodology: Biran & Ruigendijk, 2015; Dittmar et al., 2008), the present study employed single-clausal simple verb-final utterances. Verb-final multi-clausal structures are both syntactically more complex and less frequent than verb-medial single-clausal simple structures (Bader & Häussler, 2010). However, our design countered this drawback by using simple single-clausal structures that are constructed with a modal auxiliary 'wird' (will) that occupies the sentence-medial position while dislocating the verb to the sentence-final position without increasing the syntactic complexity and reducing the frequency.[3] We believe this is the main strength of the present design that ensured natural verb-final utterances that did not have nonsense verbs or that were not syntactically complex. Another strength of the stimuli was that we did not use ungrammatical structures as it is the case in ERP studies. Finally, rather than event depictions that are less likely to be observed in everyday life and where both referents could act on each other equally possibly (e.g., *elephants drawing cheetahs* or *ladybugs giving presents to the cats*), we used utterances with plausible event structures that children are more likely to encounter in real life (e.g., *rabbits finding cabbages* or *foxes finding rabbits*). These features might facilitate incremental and predictive interpretation of case markers. In ungrammatical structures or less likely event depictions, the parser may postpone interpretation until more information is available as the utterance does not lend itself to create grammatical or contextual expectations. In verb-final structures that included embedded clauses, the parser might have difficulty processing syntactic complexity. In verb-medial simple

[3] Although our stimuli included object- versus subject-initial verb-final utterances, the structures we used are not necessarily less frequent than verb-medial subject- and object-initial sentences. Bader & Häussler, (2010) report that the frequency of the object-initial orders might increase when the utterance includes the auxiliary verb 'sein/to be'. If this was the case for future modal auxiliary 'wird/will', the present verb-final structures may not be less frequent than verb-medial structures. It is very difficult to compare the frequency of the present stimuli to those of the past studies because we know that multiple factors contribute to the word order frequencies such as animacy, definiteness, syntactic complexity, length of the noun phrases, and the semantic features of the verb as well as its grammatical marking (Bader & Häussler, 2010). This is why we cannot conclude that the present pattern can be explained on the basis of frequency.

Fig. 1. Gray-scale version of a sample visual display for the following sentences.

structures, the parser may be engaged in predictive processing but we may not be able to observe these predictive effects prior to the verb because the verb appears relatively early in the utterance.

The filler items included intransitive sentences with simple and complex subjects, as in (9a) and (9b), respectively.

(9) Sample filler items:

a. Das Vogelei wird im nächsten Moment zerbrechen, nicht wahr?

'The bird's egg will break in a moment, right?'

b. Die Ente, die der Freund der Biene ist, wird im nächsten Moment schwimmen, nicht wahr?

'The duck, who is the bee's friend, will swim in a moment, right?'

Each sentence was accompanied with a visual scene depicting three potential referents. All pictures were prepared in color by a professional artist with a resolution of 640 × 480 pixels.

Each scene depicted three referents related to its corresponding critical item sentence (see Fig. 1): (i) The topic, which is the referent of NP1 in both the SOV and OSV conditions, e.g., the hare in (8a). (ii) The plausible agent, which is an animate entity that could plausibly be the doer of an action, e.g., the fox in the OSV condition (8b). (iii) The plausible patient, which is an inanimate entity which could plausibly be acted upon by the topic but is unlikely to act on the topic or on the plausible agent, e.g., the cabbage in the SOV condition (8a).

The location of the objects on the screen was counterbalanced, such that each object appeared equally often in each position (upper middle, lower right, lower left, see Fig. 1).

Children performed a sentence-picture verification task after each item. Following an item, an animated scene of the event mentioned in the sentence was either correctly (i.e., hare finding cabbage for example 1) or incorrectly (i.e., fox finding hare for example 1) depicted (see Fig. 2). The aim of the comprehension question was not to test whether case markers are interpreted correctly but to make sure that the participants pay attention to the experiment and listen to the utterances carefully. Whether the animated scene depicted the sentence correctly or incorrectly was counterbalanced across items. Children indicated verbally whether or not the event picture accurately depicted the event they heard about in the sentence.

A female native speaker of German who is an experienced researcher in the field of child language processing recorded all sentences in a natural pace one would speak to a child and with a neutral intonation and without focus accent on the first noun.[4] This was done to avoid the

[4] We thank our colleague Prof. Pia Knoeferle for kindly recording the stimuli sentences.

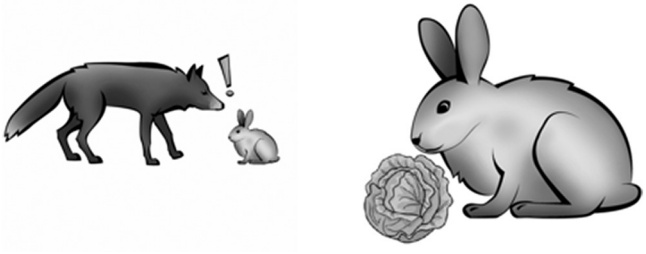

Fig. 2. Gray-scale version of sample still images shown after the animated video.

contrast of NP1 relative to NP2. We added a 500 ms silence (pause) to the beginning of every sound file. Additionally, NP1 was followed by 500 ms of pause and the sentence-final verb by 2000 ms of pause. The silence after NP1 was because of a natural prosodic break that slightly varied across all utterances. To make sure that the natural prosodic break in this critical region is the same across items, we replaced the original pause with the average pause length, which was close to 500 ms in this particular time window. The end-sentence pause of 2000 ms was added to give children sufficient time to process the utterance and to capture potential late effects. This is typically done in online studies to capture any spill-over effects as well as to control for the duration of the pause prior to the comprehension questions.

We manipulated the case marking on NP1 within subjects and created 2 counterbalanced lists. Each list contained 4 items in the OSV, 5 items in SOV condition and 5 items in the SOV, 4 items in OSV condition, respectively. Each item appeared in both conditions across both lists. The lists were pseudorandomized so that two critical items in the same condition never directly appeared back to back. The fillers were the same in both lists. For the comprehension questions, half of all trials in each condition were followed by correct event depictions and half of all trials were followed by incorrect event depictions. Sentences from both conditions were followed equally often by correct event depictions on half of the trials and by incorrect event depictions on the other half of the trials.

2.1.3. Procedure

Participants' eyes were tracked using a Tobii T60 eye tracker. Children gave verbal consent and sat in a comfortable chair in front of the eye tracker. We used Tobii Studio to calibrate the children's eyes. The trials were gaze contingent, so that a trial only started when the child was fixating on a fixation dot. If participants left their seat or moved their head a lot during the study, calibration was repeated. Prior to each trial, the objects that would appear for each item appeared on the screen individually and were introduced with their corresponding names. This was done to ensure that all children knew the names of the objects and that they had enough time to view the pictures. Following this naming phase, each trial began with the visual display of the scene showing the same three objects on the screen. The sentences were played over external speakers and started 500 ms after scene onset. 2000 ms after sentence offset, the scene disappeared and a simple animation for the comprehension question began. This animation ended in a still picture (see Fig. 2). The children looked at the screen while listening to the sentences. After seeing the animation and while looking at the still picture, they were asked whether or not the still picture they saw matched the sentence they heard. The experimenter coded the children's answers during the experiment. An experimental session lasted approximately 20 min.

3. Results

The data and the codes of the analyses of the present study is accessible at Özge, Kornfilt, Katja, Aylin, and Snedeker (2021, December

19). We first report the performance in the end-sentence comprehension task (i.e., evaluation of whether the end-sentence animation depicted the utterance correctly) for each condition. Our participants showed 96.47% success in the nominative condition while showing 95% success in the accusative condition, which demonstrates that the children paid attention to the task and listened to the utterances carefully.

Fig. 3 shows the eye-gaze patterns of the participants during the course of each condition (nominative, accusative) that reveals how the rate of agent preference changed through the course of the utterances (see Appendix 2 for the graphs showing how the looks to each referent – topic, plausible agent, plausible patient – changed through the course of the utterance). For these graphs, the fixations were divided into 100-ms time windows from the onset until the end of each utterance. Our eye-tracker samples position of raw gaze at 60 Hz so the sampling interval is 16.67 milliseconds, which means 3.6 samples per second and around 6 samples per 100 ms. Thus the number of data points per time window varied depending on the duration of each time window (e.g., for the sentence initial Pause 1 which lasted 500 ms the number of datapoints varied between 1 and 30 samples depending on the participant; for the longest time window that lasted around 1666.8 ms -i.e., TW3 that hosted the auxiliary and the adverbial, the number of datapoints varied between 1 and 100 samples). We used an R script that checks the referent the participants looked at each time point in each 100-ms time window. We assigned each 100 ms time window to a larger window based on when it occurred relative to the critical words in the sentence. This left us with the following time-windows: a 500-millisecond-pause we added preceding the first noun (Pause 1) (TW0), the first noun (TW1), a natural pause following the first noun (Pause 2) (TW2), the auxiliary (*wird*) and the adverbial (*im nächsten moment*) (TW3), a natural pause preceding the second noun (Pause 3) (TW4), the second noun (TW5), a natural pause following the second noun (Pause 4) (TW6), the verb (TW7), a natural pause following the verb (TW8), the tag question (*nicht wahr?) (TW9),* 2000-millisecond-end-sentence pause (TW10). For the sake of simplicity, in the figures below, the region after the verb is omitted as the same pattern is maintained until the end of the utterance. Each time window began with the onset of the relevant word and ended at the onset of the next word. We did not offset the time windows by 200 ms to eliminate any interference of possible phonological cues from the upcoming word and to make sure the effects we observe reflects the processes driven only within the related string.

We used as our dependent variable a binary variable '*agent preference'* which was derived by recoding AOIs as Agent and Patient: Samples (henceforth: looks) to the target in the accusative condition was recoded as Agent (i.e., 1 for the *'agent preference'* variable) and looks to the target in the nominative condition was recoded as Patient (i.e., 0 for the '*agent preference'* variable). Following Degen, Kursat, and Leigh (2021), we did not pre-aggregate data by subject and item since we used generalized logistic regression to model binary data. All other looks to the Topic or to the empty screen were also excluded from the analysis (for the predictive time windows the amount of Topic or screen looks did not differ by condition: TW3: $z = -0.29$, $p = .7$; TW4: $z = -0.33$, $p = .7$; TW5: $z = -0.11$; $p = .9$).

We were particularly interested in whether the case marking would influence the agent looks during the predictive region between NP1 and NP2. We would expect greater agent preference in the accusative condition compared to the nominative condition if case was used predictively. We know from the second experiment in Özge et al. (2019) that Turkish-speaking children showed the expected predictive effects towards the end of the adverbial (i.e., not right after NP1). However, we might see predictive effects earlier in German children as the case marking information is marked as a separate lexical item (article) preceding NP1. In line with our expectations, Fig. 3 shows that our participants show greater agent preference in the accusative condition compared to the nominative condition during Pause 2 (TW2), which is right after NP1 (the topic) (TW1). We do not see a clear preference during the next time window that has the auxiliary and the adverbial (i.e., *will in the next moment*) (TW3) but the agent preference becomes larger again during Pause 3 (TW4) before NP2 (TW5), this difference gets larger during the Pause 4 after NP2 (TW6) and it remains the same until the end of the utterance.

To see whether the agent preference changed over course of the time windows, we fit a generalized logistic regression model (Baayen, Davidson, & Bates, 2008) where the estimates were based on Laplace Approximation of the likelihood. Having a binary response variable (i.e., Agent Preference), we used a binomial (logit) family. Our fixed effects were the Case Marking on the first noun (accusative or nominative) and

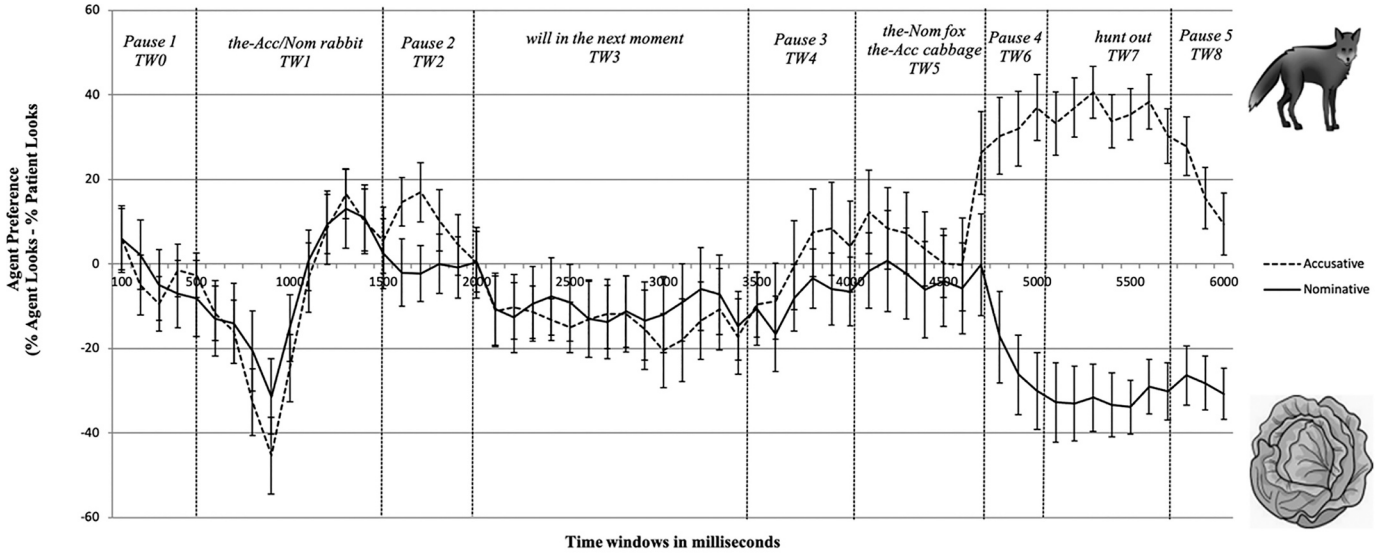

Fig. 3. Gaze patterns of agent preference in each time window in each condition. Error bars indicate the standard error of the mean. Values above zero indicate a preference to look at the potential agent, values below zero indicate a preference to look at the potential patient. Total number of data points from 19 participants and 18 items (9 per condition): Accusative: 16387, Nominative: 15709.

Table 3
Model comparisons (The shaded cells reflect the significant effects).

Effect Added	AIC	Df	Chi Sq.	*p* value
Null model (Random intercepts only)	38,675.7	3	–	–
+ Case Marking	36,105.8	4	2571.9	<0.001***
+ Time Window	35,596.3	12	525.46	<0.001***
+ Case * Time Window	31,545.2	20	4067.2	<0.001***

the Time Window (from the onset of NP1 (TW1) until the offset of tag question (TW9)) as well as the interaction of these variables, and we added Participants and Items as our random intercepts. We did not include fixed effects as random slopes for Participants and Items since the model failed to converge. The predictors were added into the model in the order stated above, and their contribution was analyzed using the likelihood-ratio-tests. This measure indicated that the best model was the one with both of the fixed effects (i.e., Case + Time Window) and the interaction of the two (i.e., Case * Time Window) (Table 3). The final code we used was as follows:

Table 5
Model comparisons (The shaded cells reflect the significant effects).

Effect Added	AIC	Df	Chi Sq.	p value
Null model (Random intercepts only)	17,617	3	–	–
+ Case Marking	17,556	4	62.5	<0.001***
+ Time Window	17,484	7	78.09	<0.001***
+ Case * Time Window	17,445	10	45.25	<0.001***

window with the auxiliary *wird* and the adverbial *im naechsten Moment* (TW3), and the natural prosodic pause preceding the second noun (Pause 3) (TW4). We included NP1 because in German, the case marking information is provided in the article preceding the noun, so this region might be the first region we could see predictive effects of case marking. We excluded the regions starting from the second noun because these regions do not inform us about predictive processing as the identity of the second noun reveals how the thematic roles are distributed between the two entities taking part in the event. The model comparisons suggested that the best model was again the one including the both of the main effects (i.e., Case + Time Window) and the interaction of the two (i.e., Case * Time Window) (Table 5).

$$\text{glmer}(\text{AgentPreference} \sim \text{Case} + \text{TimeWindow} + \text{Case*TimeWindow} + (1|\text{Participant}) + (1|\text{Item}), \text{family} = \text{binomial}, \text{control} = \text{glmerControl}(\text{optimizer} = \text{“bobyqa”}))$$

The coefficient parameter estimates of this analysis are presented in Table 4. This revealed that there was an effect of case [X^2 (4) = 2571.9, $p < .001$]: there were more looks to the potential agent in the accusative condition than in the nominative condition (Table 4). Second, there was an effect of time window [X^2 (12) = 525.46, $p < .001$]: reflecting an increase in the agent preference across condition as the sentence progressed. Finally, there was an interaction between the case and the time window [X^2 (20) = 4067.2, $p < .001$] indicating that processing patterns differed between the two sentence types as the utterance unfolded.

To focus more directly on the predictive processing, we conducted a parallel analysis of predictive time windows. This region included NP1 (TW1), the natural prosodic pause after NP1(Pause 2) (TW2), time

The coefficient parameter estimates are shown in Table 6. The effect of case persisted in this analysis [X^2 (4) = 62.5, $p < .001$], with greater agent preference in the accusative condition than in nominative condition. In this predictive time window, the effect of time window [X^2 (7) = 78.09, $p < .001$] and case by time window interaction [X^2 (10) = 45.25, $p < .001$] were also significant. The effect of time window was due to the fact that the agent preference during the auxiliary and the adverbial region (TW3) was greater than that of NP1 region (TW1). The interaction was because there was a greater agent preference in the accusative compared to the nominative condition in Pause 2 (TW2) and Pause 3 (TW4) regions while this pattern was not significant during NP1 (TW1) and the auxiliary/adverbial region (TW3).

Also, because participants often look at the same object for a second

Table 4
Coefficient parameter estimates with %95 confidence intervals (CI) (The predictive time windows are NP2 (TW2), Pause 3 (TW3), Auxiliary and Adverbial (TW4); the shaded cells reflect the significant effects).

	Estimate	Std. Error	Lower CI	Upper CI	z value	Pr (>|z|)
Intercept	−0.02	0.17	−0.36	0.31	−0.16	>0.5
Case Marking	−0.49	0.06	−0.63	−0.36	−7.27	<0.001***
Time Window, TW2	0.12	0.06	0.01	0.24	2.11	<0.05*
Time Window, TW3	−0.24	0.04	−0.32	−0.16	−5.81	<0.001***
Time Window, TW 4	0.08	0.06	−0.05	0.21	1.22	>0.5
Time Window, TW 5	0.31	0.04	0.22	0.39	7.15	<0.001***
Time Window, TW 6	0.71	0.08	0.55	0.87	8.87	<0.001***
Time Window, TW 7	1.15	0.08	0.98	1.33	12.9	<0.001***
Time Window, TW 8	1.25	0.10	1.04	1.45	12.06	<0.001***
Time Window, TW 9	0.41	0.05	0.3	0.52	7.41	<0.001***
Case:TW2	−0.72	0.12	−0.96	−0.48	−6.01	>0.05
Case:TW3	−0.06	0.08	−0.22	0.09	−0.79	<0.001***
Case:TW4	−0.51	0.13	−0.77	−0.24	−3.78	<0.001***
Case:TW5	−0.63	0.08	−0.8	−0.46	−7.35	<0.001***
Case:TW6	−2.79	0.16	−3.11	−2.48	−17.31	<0.001***
Case:TW7	−5.72	0.17	−6.07	−5.37	−31.95	<0.001***
Case:TW8	−5.43	0.20	−5.83	−5.02	−26.21	<0.001***
Case:TW9	−2.94	0.11	−3.16	−2.72	−26.27	<0.001***

Table 6
Coefficient parameter estimates with %95 confidence intervals (CI) (The shaded cells reflect the significant effects).

	Estimate	Std. Error	Lower CI	Upper CI	z value	Pr (>|z|)
Intercept	−0.06	0.19	−0.44	0.31	−0.33	>0.05
Case	−0.25	0.07	−0.40	−0.11	−3.50	<0.001***
Time Window TW2	0.11	0.06	−0.01	0.23	1.79	>0.05
Time Window TW3	−0.29	0.04	−0.37	−0.21	−6.82	<0.001***
Time Window TW4	0.07	0.06	−0.06	0.2	1.03	>0.05
Case: Time Window TW2	−0.72	0.12	−0.96	−0.47	−5.84	<0.001***
Case: Time Window TW3	−0.04	0.08	−0.21	0.11	−0.55	>0.05
Case: Time Window TW4	−0.47	0.12	−0.74	−0.20	−3.41	<0.001***

Table 7
Coefficient parameter estimates with %95 confidence intervals (CI) (The shaded cells reflect the significant effects).

	Estimate	Std. Error	Lower CI	Upper CI	z value	Pr (>|z|)
Intercept	−0.17	0.18	−0.53	0.19	−0.92	>0.05
Case	−0.41	0.05	−0.52	−0.31	−7.88	<0.001***

even more, we expected that our dependent variable would be autocorrelated across Time Windows. To address this, we constructed a model combining the time windows during the critical region starting from NP1 (TW1) ending at the offset of Pause 3 (TW4). In this analysis too, there was an effect of case in the expected direction [X^2 (1) = 62.5, $p < .001$] such that the agent preference was greater in the accusative condition compared to the nominative condition (Table 7). Furthermore, when the looks to the agent was aggregated by item within participants in these critical time windows (Dink & Ferguson, 2015), the effect of case remained significant [X^2 (1) = 5.44, $p = .01$]. Finally, to address the autocorrelation issue using an approach suggested in Mirman, Dixon, and Magnuson (2008), we conducted growth curve analyses (also see, Cho, Brown-Schmidt, & Lee, 2018). These analyses were parallel to the mixed model analyses, confirming the predictive effect of case prior to the NP2 region (Appendix 3).

To sum up, these analyses revealed there was an effect of case in the expected direction in the time windows preceding NP2 (i.e., during Pause 2 after NP1 and during Pause 3 after the aux), such that there were more looks to the plausible agent in the accusative compared to the nominative condition. This demonstrates that four-year-old German-speaking children are able to parse the case marking on the article of the first noun to generate thematic predictions about the identity of the second noun prior to the second noun as well as the verb become available.

4. Discussion

We set out to investigate whether four-year-old German-speaking children can interpret case marking information presented sentence initially on NP1 to predict the thematic role of NP2 in verb-final utterances. We used a similar visual-world eye-tracking task to the one used with Turkish speaking four-year old children (Özge et al., 2019). In this task, we presented children with a visual scene with three related referents along with a spoken utterance that started with a noun either in the nominative or the accusative case, and tracked their eye movements to see which referent they looked at during the predictive regions (i.e., after they heard NP1 and before they heard NP2). The results revealed that German-speaking children were able to shift their gaze to the target entity using the case marking on NP1 right after NP1 (i.e., before they heard NP2 and before they heard the verb). Therefore, German-speaking children behaved quite similarly to the age-matched Turkish-speaking children in employing a predictive processing mechanism. Indeed, encountering the case marking cues as a separate lexical item prior to the noun, German children showed the predictive effect of case during an earlier time window (right after NP1) than that of Turkish speaking children (i.e., right before NP2) who encountered the case marking cues as a suffix added to the noun. To the best of our knowledge, this is the first piece of evidence showing that German-speaking children as young as four years of age can attain a structural interpretation of the case markers incrementally in non-canonical word orders, which enables them to generate anticipations about the upcoming structure. In the rest of the discussion, we will discuss (i) how these findings relate to the previous findings about the interpretation of case marking in German children and (ii) how they can inform our theories of child language parsing.

4.1. Present findings in relation to previous studies on German children

As far as our knowledge of previous studies is concerned, the present study presents the earliest evidence for predictive parsing in German-speaking preschoolers. Previous studies have reported that German-speaking children cannot interpret case marking until age five unless the verb is familiar and they cannot do so until age seven unless the nouns are ordered in a canonical order (i.e., SVO vs. OVS) (Dittmar et al., 2008). These offline findings have been supported by the ERP findings revealing that children fail to detect ungrammaticalities due to case marking violations in noncanonical orders (Schipke et al., 2012) and by a series of recent eye-tracking studies reporting that four- and five-year-old German children can interpret case markers in object-initial OVS utterances only when there is an additional action cue visually depicting the actions denoted by the verbs (Kröger et al., 2017; Kröger et al., 2018; Maquate & Knoeferle, 2021; Münster, 2016). One common theoretical implication of these studies is that attaining the full competence of case marking cues that enable one to assign thematic roles in an utterance is a protracted process that is not complete until age six, therefore young children rely heavily on lexically-based cues such as the verb, world-situated cues such as the action, or distributional patterns locating the doer of the actions prior to the undergoer/patient (a la late abstraction accounts).

The pattern of predictive use of case markers we attained in the present study, however, conflict with this implication. We discussed in the introduction the idea that the contradictory patterns between German and Turkish children reported in previous studies might be due either to the crosslinguistic differences (i.e., more reliable and transparent case marking system in Turkish than in German) or to the methodological differences among the previous studies. With the present study, we rule out the first possibility, as German speaking children performed quite similarly to Turkish speaking children once they are tested with the same task. This leads us to conclude that the previous inconsistencies are more due to the methodological differences. Different from Dittmar et al.'s (2008) study, our study did not rely on an offline comprehension measure, which requires children to retain their analysis in their short-term memory until the end of the utterance. Our study allowed us to capture the thematic interpretation and anticipation during the course of the spoken utterance. Different from Knoll et al. (2012) and Schipke et al. (2012), our study did not rely on violation paradigms so our test items were not blended with ungrammatical utterances. This may be a feature that facilitates predictive parsing. In violation paradigms, the parser's structural anticipations turn out to be

wrong in half of the cases, which may be causing the parser to abandon/postpone its predictions. Different from previous visual-world eye-tracking studies (Kröger et al., 2017; Maquate & Knoeferle, 2021) that used event depictions that are less likely to encounter in everyday life (i.e., events where both referents could act on each other equally possibly; e.g., *elephants drawing cheetahs* or *ladybugs giving presents to the cats*), our study had arguably more plausible event structures we encounter in real life (e.g., *rabbits finding cabbages* or *foxes finding rabbits*). In event structures that do not reflect everyday events, the action may unfold in multiple ways. We conjecture that the parser may be forgoing its predictions until the verb/action in those cases to minimize its processing load. It may be the case that it is more economical to passively wait and see what will happen than making incorrect predictions and revising them. The same applies to the violation paradigms where half of the predictions fail. Tapping into highly plausible world knowledge as it may, our study by no means allows the participants to solve the thematic role assignment puzzle solely on the basis of this world knowledge cue. The first entity that appears in our test items could be the agent or the patient of two equally plausible event structures (e.g., the hare can be the entity finding a cabbage – *agent* or it can be the entity found by the fox – *patient*). Since the verb appears at the end of the utterance, the predictive cue to the event structure is only the case marker on NP1. Finally, locating the verb sentence-finally without extra syntactic complexity (i.e., using single-clausal simple verb-final structures) might have eased processing as well as giving a greater time window for our participants to generate their predictions. One question is whether the verb-medial structures that were predominantly employed in previous studies could be more difficult to process compared to the verb-final structures we employed in the present study. We do not think that verb-medial structures are more difficult to comprehend. However, in the verb-medial structures, there is limited time before the parser can process the case marker prior to the verb as the verb appears right after NP1. Overall, even if the verb-medial structures are more frequent in German, the time window to process the case marker before the verb is smaller compared to the verb-final structures. However, in verb-final structures the parser would have enough time to process the case marker on NP1 and program its predictions about the rest of the event structure before NP2 and before the verb. In other words, one can speculate that the parser encounters the verb as an additional cue before it can finalize its processing of the case marking on NP1 in verb-medial structures whereas it has enough time to do so in verb-final structures.[5] Similarly, the the future modality marker (*wird/will*) and an adverbial pointing to what will happen shortly (*im nächsten Moment/in the next moment/shortly*) might have facilitated the expectation-based parsing routes.

4.2. Present findings in relation to the child parsing mechanisms

Our findings in this study are consistent with the previous pattern presented by Turkish speaking children (Özge et al., 2019). For the predictive interpretation of case markers, listeners should immediately recognize the case marker on NP1 and interpret it by determining its thematic role. This would allow them to generate hypotheses about possible upcoming event structures and the thematic role of upcoming entitites. We have a lot of evidence that adult parsing integrates multiple levels of information incrementally without waiting until the processing at lower levels is complete. This enables the parser to create structural and semantic anticipations about the next possible continuations (for reviews see, Elman, Hare, & McRae, 2005; Treiman, Clifton Jr., Meyer, & Wurm, 2003; van Gompel, 2013; Kuperberg & Jaeger, 2016). The present results are consistent with such a comprehension model for a developing parser as well. Children assign an immediate analysis for each linguistic cue they encounter even if this cue is a short morpheme and they eagerly hypothesize about the properties of the utterance yet to unfold.

To be able to succeed in such as task, children need to have adultlike incremental processing mechanisms. Apart from this, they also need broader semantic abstractions about thematic roles and how these roles are encoded in their language. This is not in line with the late abstraction accounts that predict a prolonged and lexically-based linguistic abstraction process in children. According to this account, children lack prewired semantic or syntactic predispositions allowing them to expect regularities in the linguistic input so their abstraction process begins around age three but the necessary processing strategies (e.g., preemption, analogy making) allowing them to reach fuller abstractions and hypotheses about how language works develops only after age five or six (e.g., Ambridge & Lieven, 2011; Goldberg, 2006; Savage, Lieven, Theakston, & Tomasello, 2003; Tomasello, 1992; Tomasello, 2003). The present findings are also incompatible with the late neural syntactic maturation account that predicts a prolonged acquisition of complex syntactic functions due to the late-developing brain regions that are assumed to host these syntactic functions (Friederici, 2011, 2012). Our study provides crosslinguistic support for the perspectives assuming broad semantic expectations and early syntax-semantic mapping (Fisher, 2002; Gertner, Fisher, & Eisengart, 2006; Yuan & Fisher, 2009; Kline & Demuth, 2014; Naigles, 1990; Arunachalam & Waxman, 2010; Waxman & Markow, 1995).

5. Conclusion

This paper has provided the earliest evidence in German children for an incremental and predictive interpretation of case markers that is independent of the verb and word order cues. This indicates that the previous findings showing protracted patterns of acquisition of case markers in German along with non-adultlike processing patterns cannot be due to children's limited abstract semantic and syntactic knowledge, immature processing mechanism, or to their late-maturing syntax. Our methodology using a natural task, a meaningful visual context, and typical utterances enabled us to detect adultlike incremental and predictive processing patterns as early as age four in German children (i.e., at a similar age to Turkish children). These findings reveal that children at age four are already adept at adapting their parsing strategies to their language model. This would be far more difficult without a system with broad semantic representations that is ready to look for abstractions and statistical regularities in the input.

Credit author statement

Duygu Özge: Conceptualization, Methodology, Data Collection, Formal Analysis, Writing-original draft preparation, Reading and Editing, Funding Acquisition.

Jaklin Kornfilt: Adaptation of the German stimuli, Writing-original draft preparation, Reading and Editing.

Katja Maquate: Data Collection, Writing-original draft preparation.

Aylin C. Küntay: Supervision, Reading and Editing.

Jesse Snedeker: Conceptualization, Supervision, Methodology, Resources, Supervision, Reading and Editing.

Acknowledgements

This research was supported by a grant from European Commission Research Executive Agency Marie Curie International Outgoing Fellowship (FP7-PEOPLE-2011-IOF 301637) to D. Özge. We thank Pia

[5] A related question is whether the verb-medial structures could be biasing the parser for a subject-initial configuration. However, we think this is not very likely as the subject-before-object configuration is more frequent both in verb-medial and verb-final orderings crosslinguistically. In fact, 97% of the world's language locate the subject before object regardless of the location of the verb (Dryer, 2011). Thus, children actually have a bias to treat the first entity as the subject in verb-initial and verb-final languages (Slobin & Bever, 1982).

Knoeferle for hosting the study at University of Bielefeld and for her collaboration in this study, Ebru Evcen for data trimming and the Growth Curve Analysis, and Ilgar Veryeri for the visual material.

Appendix 1. Test items

1
SOV: Der Hase wird im nächsten Moment den Kohl aufspüren, nicht wahr?
'The rabbit will shortly (in the next moment) eat the cabbage, right?'
OSV: Den Hasen wird im nächsten Moment der Fuchs aufspüren, nicht wahr?
'The fox will shortly eat the rabbit, right?'
2
SOV: Der Kater wird im nächsten Moment den Käse begucken, nicht wahr?
'The cat will shortly look at the cheese, right?
OSV: Den Kater wird im nächsten Moment der Hund begucken, nicht wahr?
'The dog will shortly look at the cat, right?'
3
SOV: Der Hahn wird im nächsten Moment den Weizen schlucken, nicht wahr?
'The rooster will shortly swallow the wheat, right?'
OSV: Den Hahn wird im nächsten Moment der Wolf schlucken, nicht wahr?
'The wolf will shortly swallow the rooster, right?'
4
SOV: Der Junge wird im nächsten Moment den Lollie lecken, nicht wahr?
'The boy will shortly lick the lollipop, right?'
OSV: Den Jungen wird im nächsten Moment der Esel treten, nicht wahr?
'The donkey will shortly kick the boy, right?'
5
SOV: Der Bär wird im nächsten Moment den Honig finden, nicht wahr?
'The bear will shortly find the honey, right?'
OSV: Den Bär wird im nächsten Moment der Jäger finden, nicht wahr?
'The hunter will shortly find the bear, right?'
6
SOV: Der Affe wird im nächsten Moment den Pfirsich schnappen, nicht wahr?
'The monkey will shortly grab the peach, right?'
OSV: Den Affen wird im nächsten Moment der Löwe schnappen, nicht wahr?
'The lion will shortly catch the monkey, right?'
7
SOV: Der Dieb wird im nächsten Moment den Geldbeutel suchen, nicht wahr?
'The thief will shortly seek the wallet, right?'
OSV: Den Dieb wird im nächsten Moment der Polizist suchen, nicht wahr?
'The policeman will shortly seek the thief, right?'
8
SOV: Der Schimpanse wird im nächsten Moment den Apfel fressen, nicht wahr?
'The chimpanzee will shortly eat the apple, right?'
OSV: Den Schimpansen wird im nächsten Moment der Wolf fressen, nicht wahr?
'The wolf will shortly eat the chimpanzee, right?
9
SOV: Der Jungen wird im nächsten Moment den Ball treten, nicht wahr?
'The boy will shortly kick the ball, right?'
OSV: Den Jungen wird im nächsten Moment der Hund treten, nicht wahr?
'The dog will shortly kick the boy, right?'

Appendix 2. Graphs plotting how gaze patterns on each referent on the visual context changed through the course of the utterance in each condition for both groups for both experiments

Figs. 4 and 5 depict that participants looked at each referent as it was mentioned and there were overall more looks at the topic referent compared to other possibilities on the screen until the presentation of NP2. This is not surprising because this has been the only referent that has been overtly mentioned until the second referent is revealed towards the end of the utterance and it is the referent the utterance is mainly about (i.e., the sentence is about what the first referent (i.e., *the rabbit*) did or what happened to it).

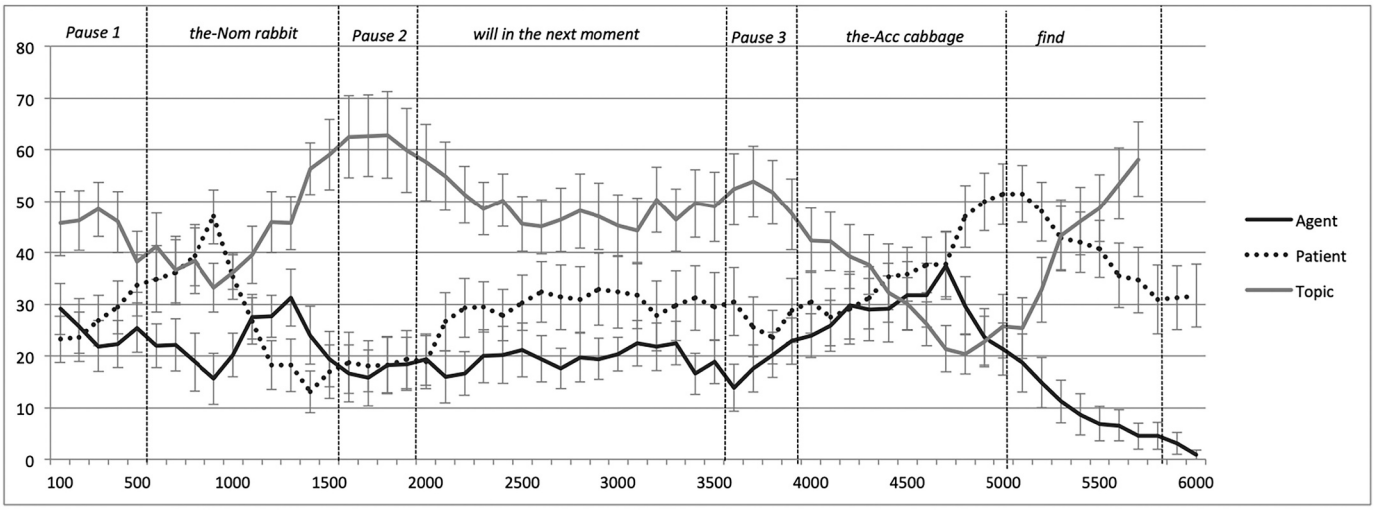

Fig. 4. Proportion of fixations to each of the objects (the topic, the plausible agent, and the plausible patient) on the visual scene in the nominative condition.

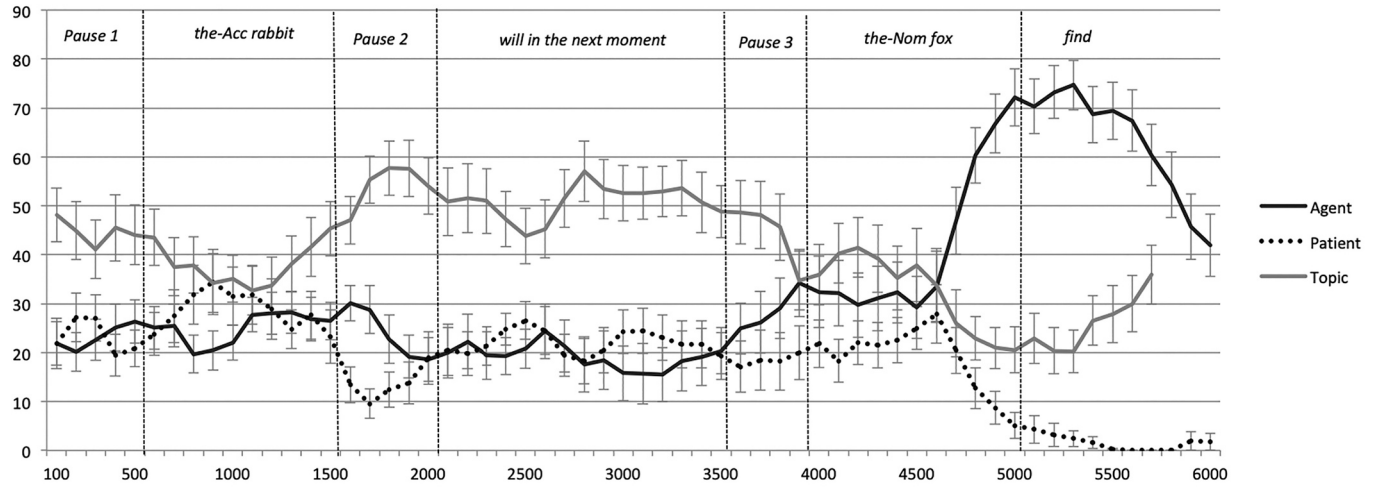

Fig. 5. Proportion of fixations to each of the objects (the topic, the plausible agent, and the plausible patient) on the visual scene in the accusative condition.

We were particularly interested in whether the case marking influenced the looks to the two referents yet to be mentioned (i.e., *the fox* – plausible agent and *the cabbage* – plausible patient) during the predictive time windows. If the sentence-initial case information were interpreted incrementally, we would see greater looks to the plausible patient (*the cabbage*) right after the nominative marked NP1 and greater looks to the plausible patient (*the fox*) right after the accusative marked NP1. This is what we observe in the figures above. As we see in Fig. 4, the looks to the plausible patient are greater during the fourth time window that consists of the auxiliary and the time adverbial (*will in the next moment*), which reveals that the participants launched predictive looks to the patient entity upon hearing the nominative NP1. Similarly, it is clear in Fig. 5 that the looks to the plausible agent are greater during the third time window, which corresponds to the pause right after NP1, which shows that the participants launched predictive looks to the agent entity upon hearing the accusative NP1. In Fig. 5, we also see that this pattern is not maintained during the fourth time window (the region that hosts the auxiliary and the time adverbial) during which the participants shifted their gaze back and forth between the plausible agent and plausible patient until the end of this region. Form that moment on, they once more have a constant focus on the target (Agent) referent during the prosodic pause (Pause 3) until the onset of NP2 and until the end of the utterance. Thus, these figures show that participants prefer to look at the topic referent (i.e., being the referent under discussion) until late in utterance for both conditions and we see clear predictive effects in both conditions prior to NP2 although there are points within the predictive time windows where the pattern becomes murkier.

Appendix 3. Growth curve analyses

To see whether looks to the agent changed over the course of predictive time windows (i.e., the period between NP1 and NP2), we conducted growth curve analyses (hereafter GCA; Mirman et al., 2008; Mirman, Landrigan, & Britt, 2017) on participants' looks to the agent in accusative and nominative condition. GCA allows us to model the rise and the fall in the gaze patterns by fitting the curves to the proportion of looks over time and assessing statistical significance of the bends in these curves (Mirman et al., 2008).

For the analyses, we aggregated looks to the agent across participants and items and calculated empirical logit transformation of fixation probabilities to the agent for correction[6] (Barr, 2007). The empirical logit transformed proportions to the agent served as our dependent variable. We conducted separate growth curve analyses for the two predictive windows and time window including NP2: NP1 window (i.e., NP1 in either accusative or nominative condition + Pause 2) (TW1 and TW2), Adverbial window (i.e., *will in the next moment* + Pause 3) (TW3 and TW4), and NP2 window (i.e., NP2 and Pause 4) (TW5). For the analyses, we fitted a model to predict empirical logits from fixed effects of Case Marking (Accusative vs. Nominative), *Time* and their interaction. 2-level factor Case Marking was deviation coded before model fitting (Accusative, −0.5; Nominative 0.5). *Time* (*Time1, Time2, Time3*) was a continuous variable modeled with third-order orthogonal (i.e., uncorrelated therefore more appropriate for multiple regression) polynomial (cubic). We constructed our models with a third-order polynomial since the curve only had two bends and the model with a fourth-order polynomial gave worse fits of the data (Mirman et al., 2008). This third-order polynomial allows us to capture how the terms (*Time1, Time2, Time3*) independently affect the form of the curve: *Time1* is a straight line indicating a single change in focus from neutral start, *Time2* is a U-shaped curve indicating two changes in focus from initial flatness to the first decrease in looks to the agent and the reversal at the bottom, *Time3* indicates three changes in focus from initial flatness to the first increase in looks to the agent followed by a decrease and another increase (i.e., looks to the agent, to the patient, back to the agent) (Mirman et al., 2008; Mirman et al., 2017). The model included random intercepts for participants and items. We report the regression coefficient, the standard error and t-value. Parameter estimate degrees of freedom and corresponding *p*-values were estimated using Satterthwaite's method.

1. *Analysis of NP1 window (TW1 & TW2)*

We found a significant main effect of Time 2 ($\beta = -3.02$, SE = 0.62, $t = -4.8$, $p < .001$), Time 3 ($\beta = -3.62$, SE = 0.6, $t = -6.03$, $p < .001$) and a significant interaction between Case Marking and Time 3 ($\beta = 2.86$, SE = 1.2, $t = 2.37$, $p < .05$) (Fig. 6).

This interaction was due to the following: In the accusative condition, looks to the agent initially decreased and this was followed by a significant greater/faster increase and a greater/faster secondary decrease than the looks in the nominative condition. In other words, patterns in the nominative condition were flatter.

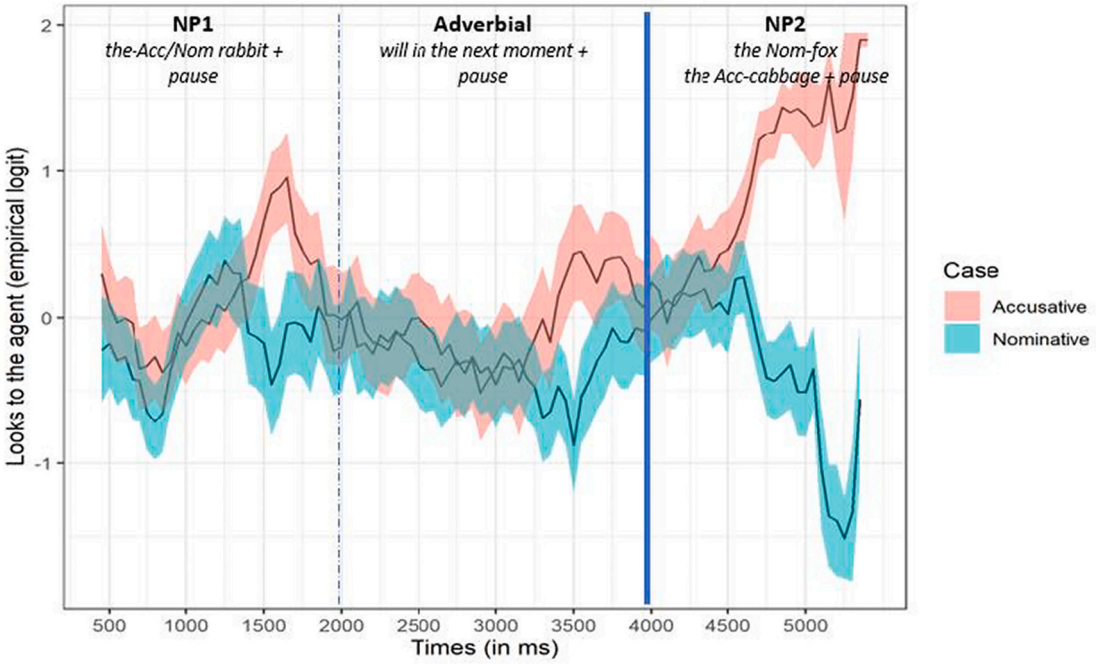

Fig. 6. Looks to the agent from the NP1 onset to the NP2 offset by Case Marking. Standard errors of the mean are represented by ribbons.

2. *Analysis of adverbial window (TW3 & TW4)*

We found a significant main effect of Case ($\beta = -0.5$, SE = 0.09, $t = -5.34$, $p < .001$), *Time2* ($\beta = 4.67$, SE = 0.7, $t = 6.63$, $p < .001$) and a significant interaction between Case and *Time2* ($\beta = -3.09$ SE = 1.42, $t = -2.17$, $p < .05$) as well as Case and *Time3* ($\beta = 4.28$, SE = 1.51, $t = 2.82$, $p < .01$) (Fig. 6). The main effect of Case was caused by the fact that the looks to the agent in the accusative condition were significantly greater than those in the Nominative condition. Positive coefficients of *Time2* indicated that there was an overall decrease in the looks to the agent followed by an increase in both conditions (i.e., U- shaped curvature). However, the interaction between Case and *Time2* showed that there was a greater curvature in looks (i.e., steeper and faster looks to the agent) in the accusative condition compared to the nominative condition. Likewise, in the nominative condition, looks to the agent were less stable and fluctuated significantly more (i.e., there was an initial slight increase which was followed by a decrease and a secondary increase).

[6] When eye-tracking data is aggregated across participants and items, sample-by-sample binomial responses are lost. Therefore, empirical logit, which takes proportional data as and returns a linear-corrected DV as output is calculated (Dink & Ferguson, 2015).

3. Analysis of NP2 window (TW5 & TW6)

There were a significant main effect of Case Marking ($\beta = -1.86$, SE = 0.1, $t = -18.17$, $p < .001$) and a significant interaction between Case and *Time1* ($\beta = -12.01$, SE = 1.41, $t = -8.47$, $p < .001$) and Case and *Time2* ($\beta = -8.44$, SE = 1.32, $t = -6.37$, $p < .001$) (Fig. 6). The main effect of Case is due to the greater looks to the agent in the accusative condition than in nominative condition. These interactions were due to the following: In the accusative condition, looks to the agent continued to increase linearly whereas in the nominative condition the increasing looks to the agent started to decrease gradually (i.e., reversed U-shaped pattern). This is an expected pattern because the target NP (i.e., patient in the nominative condition) was revealed in this time window.

References

Altmann, G. T., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science, 33*(4), 583–609. https://doi.org/10.1111/j.1551-6709.2009.01022.x

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Ambridge, B., & Lieven, E. V. M. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge University Press.

Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition, 114*(3), 442–446. https://doi.org/10.1016/j.cognition.2009.10.015

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bader, M., & Häussler, J. (2010). Word order in German: A corpus study. *Lingua, 120*(3), 717–762. https://doi.org/10.1016/j.lingua.2009.05.007

Barr, D. J. (2007, September 12). Walkthrough of an "empirical logit" analysis in R. http://talklab.psy.gla.ac.uk/tvw/elogit-wt.html.

Biran, M., & Ruigendijk, E. (2015). Do case and gender information assist sentence comprehension and repetition for German-and Hebrew-speaking children? *Lingua, 164*, 215–238. https://doi.org/10.1016/j.lingua.2015.06.012

Bornkessel, I., Schlesewsky, M., & Friederici, A. D. (2002). Grammar overrides frequency: Evidence from the online processing of flexible word order. *Cognition, 85*(2), B21–B30. https://doi.org/10.1016/S0010-0277(02)00076-8

Borovsky, A., Elman, J., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112*(4), 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Cho, S.-J., Brown-Schmidt, S., & Lee, W.-Y. (2018). Autoregressive generalized linear mixed effect models with crossed random effects: An application to intensive binary time series eye-tracking data. *Psychometrika, 83*(3), 751–771. https://doi.org/10.1007/s11336-018-9604-2

Chomsky, N. (1965). *Aspects of the theory of syntax*. The MIT Press.

Chomsky, N. (1981). Lectures on government and binding. *Foris* (Publications).

Degen, J., Kursat, L., & Leigh, D. D. (2021). Seeing is believing: Testing an explicit linking assumption for visual world eye-tracking in psycholinguistics. *Proceedings of the Annual Meeting of the Cognitive Science Society, 43*(43), 1500–1506.

Dink, J. W., & Ferguson, B. (2015). EyetrackingR: An R library for eye-tracking data analysis. https://www.eyetracking-r.com/.

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German children's comprehension of word order and case marking in causative sentences. *Child Development, 79*(4), 1152–1167. https://doi.org/10.1111/j.1467-8624.2008.01181.x

Dressler, W. U. (2019). Natural morphology. In *Oxford research encyclopedia of linguistics*. https://doi.org/10.1093/acrefore/9780199384655.013.576

Dryer, M. S. (2011). The evidence for word order correlations. *Linguistic Typology, 15*, 335–380. https://doi.org/10.1515/lity.2011.024

Elman, J. L., Hare, M., & McRae, K. (2005). Cues, constraints, and competition in sentence processing. In M. Tomasello, & D. I. Slobin (Eds.), *Beyond nature-nurture: Essays in honor of Elizabeth Bates* (pp. 111–138). Lawrence Erlbaum Associates Publishers.

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development, 72*(4), 1003–1015. https://doi.org/10.1111/1467-8624.00331

Ferretti, T. R., Kutas, M., & McRae, K. (2007). Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 182–196. https://doi.org/10.1037/0278-7393.33.1.182

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language, 44*(4), 516–547. https://doi.org/10.1006/jmla.2000.2728

Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science, 5*(1), 55–64. https://doi.org/10.1111/1467-7687.00209

Friederici, A. D. (2011). The brain basis of language processing: From processing to function. *Physiological Reviews, 91*(4), 1357–1392. https://doi.org/10.1152/physrev.00006.2011

Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences, 16*(5), 262–268. https://doi.org/10.1016/j.tics.2012.04.001

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science, 17*(8), 684–691. https://doi.org/10.1111/j.1467-9280.2006.01767.x

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition: A Journal of Developmental Linguistics, 1*(1), 3–55. https://doi.org/10.1207/s15327817la0101_2

Göksun, T., Küntay, A. C., & Naigles, L. R. (2008). Turkish children use morphosyntactic bootstrapping in interpreting verb meaning. *Journal of Child Language, 35*(2), 291–323. https://doi.org/10.1017/S0305000907008471

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199268511.001.0001

van Gompel, R. P. G. (2013). *Sentence processing: Current issues in the psychology of language*. Psychology Press.

Haider, H. (2010). *The Syntax of German*. Cambridge University Press. https://doi.org/10.1017/CBO9780511845314

Hemforth, B., & Konieczny, L. (2000). Cognitive parsing in German: An introduction. In B. Hemforth, & L. Konieczny (Eds.), *German sentence processing* (pp. 1–23). Springer.

Hoberg, U. (1981). *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache* (Vol. 10). Max Hueber Verlag.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133–156. https://doi.org/10.1016/S0749-596X(03)00023-8

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research, 32*(1), 37–55. https://doi.org/10.1023/A:1021933015362

Kempen, G., & Harbusch, K. (2005). The relationship between grammaticality ratings and corpus frequencies: A case study into word order variability in the midfield of German clauses. In S. Kepser, & M. Reis (Eds.), *Vol. 85. Linguistic evidence: Empirical, theoretical and computational perspectives* (pp. 329–349). De Gruyter Mouton. https://doi.org/10.1515/9783110197549.329.

Ketrez, F. N., & Aksu-Koç, A. (2009). Early nominal morphology in Turkish: Emergence of case and number. In S. Ursula, & M. D. Voeikova (Eds.), *Vol. 30. Development of nominal inflection in first language acquisition: A cross-linguistic perspective* (pp. 15–48). De Gruyter Mouton. https://doi.org/10.1515/9783110217117.15.

Kline, M., & Demuth, K. (2014). Syntactic generalization with novel intransitive verbs. *Journal of Child Language, 41*(3), 543–574. https://doi.org/10.1017/S0305000913000068

Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye tracking. *Cognitive Science, 30*(3), 481–529. https://doi.org/10.1207/s15516709cog0000_65

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition, 95*(1), 95–127. https://doi.org/10.1016/j.cognition.2004.03.002

Knoll, L. J., Obleser, J., Schipke, C. S., Friederici, A. D., & Brauer, J. (2012). Left prefrontal cortex activation during sentence comprehension covaries with grammatical knowledge in children. *Neuroimage, 62*(1), 207–216. https://doi.org/10.1016/j.neuroimage.2012.05.014

Kröger, J. M., Münster, K., Burigo, M., & Knoeferle, P. (2018). Wiggle, wiggle, wiggle: How visual cues influence thematic role assignment in children and adults. In *Proceedings of the 40th annual meeting of the cognitive science society* (pp. 1969–1974).

Kröger, J. M., Münster, K., & Knoeferle, P. (2017). The influence of prosody and case marking on thematic role assignment in ambiguous action scenes: Adults versus children. In *Proceedings of the 39th annual meeting of the cognitive science society* (pp. 2463–2468).

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 843–847. https://doi.org/10.1037/a0029284

Maquate, K., & Knoeferle, P. (2021). Referential vs. non-referential world-language relations: How do they modulate language comprehension in 4 to 5-year-olds, younger, and older adults? *Frontiers in Psychology, 11*, Article 542091. https://doi.org/10.3389/fpsyg.2020.542091

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition, 33*(7), 1174–1184. https://doi.org/10.3758/BF03193221

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language, 59*(4), 475–494. https://doi.org/10.1016/j.jml.2007.11.006

Mirman, D., Landrigan, J.-F., & Britt, A. E. (2017). Taxonomic and thematic semantic systems. *Psychological Bulletin, 143*(5), 499–520. https://doi.org/10.1037/bul0000092

Münster, K. (2016). Effects of emotional facial expressions and depicted actions on situated language processing across the lifespan [doctoral dissertation, Bielefeld University]. Bielefelder E-Dissertation. https://pub.uni-bielefeld.de/record/2906648.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*(2), 357–374. https://doi.org/10.1017/S0305000900013817

Nation, K., Marshall, C. M., & Altmann, G. T. M. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology, 86*(4), 314–329. https://doi.org/10.1016/j.jecp.2003.09.001

Özge, D., Küntay, A., & Snedeker, J. (2019). Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension. *Cognition, 183*, 152–180. https://doi.org/10.1016/j.cognition.2018.10.026

Özge, D., Marinis, T., & Zeyrek, D. (2015). Incremental processing in head-final child language: Online comprehension of relative clauses in Turkish-speaking children and adults. *Language, Cognition and Neuroscience, 30*(9), 1230–1243. https://doi.org/10.1080/23273798.2014.995108

Özge, Duygu, Kornfilt, Jaklin, Katja, Maquate, Aylin, Küntay, & Snedeker, Jesse (2021). Data for German-speaking children use sentence-initial case marking for predictive language processing at age four. Retrieved from osf.io/mn2ce. *Open Science Foundation, December 19*. osf.io/mn2ce, in this issue.

Pinker, S. (1984). *Language learnability and language development*. Harvard University Press.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science, 6*(5), 557–567. https://doi.org/10.1111/1467-7687.00312

Schipke, C. S., Friederici, A. D., & Oberecker, R. (2011). Brain responses to case-marking violations in German preschool children. *Neuroreport, 22*(16), 850–854. https://doi.org/10.1097/WNR.0b013e32834c1578

Schipke, C. S., Knoll, L. J., Friederici, A. D., & Oberecker, R. (2012). Preschool children's interpretation of object-initial sentences: Neural correlates of their behavioral performance. *Developmental Science, 15*(6), 762–774. https://doi.org/10.1111/j.1467-7687.2012.01167.x

Siegel, D. C. (1974). Topics in English morphology. [Doctoral dissertation, Massachuseits Institute of Technology]. DSpace@MIT. http://hdl.handle.net/1721.1/13022.

Slobin, D. I., & Bever, T. G. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition, 12*(3), 229–265. https://doi.org/10.1016/0010-0277(82)90033-6

Snedeker, J. (2009). Sentence processing. In E. L. Bavin (Ed.), *The Cambridge handbook of child language* (pp. 321–337). Cambridge University Press.

Snedeker, J., & Huang, Y. T. (2015). Sentence processing. In E. L. Bavin, & L. R. Naigles (Eds.), *The Cambridge handbook of child language* (2nd ed., pp. 409–437). Cambridge University Press. https://doi.org/10.1017/CBO9781316095829.019.

Snedeker, J., & Trueswell, J. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*(3), 238–299. https://doi.org/10.1016/j.cogpsych.2004.03.001

Staub, A., & Clifton, C., Jr. (2006). Syntactic prediction in language comprehension: Evidence from either...Or. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(2), 425–436. https://doi.org/10.1037/0278-7393.32.2.425

Stephany, U., & Voeikova, M. D. (2009). Introduction. In S. Ursula, & M. D. Voeikova (Eds.), *Vol. 30. Development of nominal inflection in first language acquisition: A cross-linguistic perspective* (pp. 1–14). De Gruyter Mouton. https://doi.org/10.1515/9783110217117.1.

Strotseva-Feinschmidt, A., Schipke, C. S., Gunter, T. C., Brauer, J., & Friederici, A. D. (2019). Young children's sentence comprehension: Neural correlates of syntax-semantic competition. *Brain and Cognition, 134*, 110–121. https://doi.org/10.1016/j.bandc.2018.09.003

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*(2), 73–108. https://doi.org/10.1016/S0010-0277(99)00021-9

Tomasello, M. (1992). *First verbs: A case study in early grammatical development*. Cambridge University Press. https://doi.org/10.1017/CBO9780511527678

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*(3), 209–253. https://doi.org/10.1016/S0010-0277(99)00069-4

Tomasello, M. (2003). *Constructing a language: A usage-based account of language acquisition*. Harvard University Press.

Treiman, R., Clifton, C., Jr., Meyer, A. S., & Wurm, L. H. (2003). Language comprehension and production. In A. F. Healy, & R. W. Proctor (Eds.), *Vol. 4. Handbook of psychology: Experimental psychology* (pp. 527–547). John Wiley & Sons Inc.. https://doi.org/10.1002/0471264385.wei0419

Trueswell, J. C., Sekerina, I. A., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73*(2), 89–134. https://doi.org/10.1016/S0010-0277(99)00032-3

Vergnaud, J. R. (1977). Formal properties of phonological rules. In R. E. Butts, & J. Hintikka (Eds.), *Basic problems in methodology and linguistics* (pp. 299–318). Springer.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*(3), 257–302. https://doi.org/10.1006/cogp.1995.1016

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition, 99*(2), B63–B72. https://doi.org/10.1016/j.cognition.2005.07.001

Xanthos, A., Laaha, S., Gillis, S., Stephany, U., Aksu-Koç, A., Christofidou, A., … Dressler, W. U. (2011). On the role of morphological richness in the early development of noun and verb inflection. *First Language, 31*(4), 461–479. https://doi.org/10.1177/0142723711409976

Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby": Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science, 20*(5), 619–626. https://doi.org/10.1111/j.1467-9280.2009.02341.x

Zhang, L., & Knoeferle, P. (2012). Visual context effects on thematic role assignment in children versus adults: Evidence from eye tracking in German. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *34. Proceedings of the annual meeting of the cognitive science society* (pp. 2593–2598) (34) https://pub.uni-bielefeld.de/record/2615852.