Running head: Thematic prediction in children

Please Note: This is pre-proof copy

Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension

Duygu Özge[1,2,3], Aylin Küntay[2] & Jesse Snedeker[3]

Middle East Technical University,[1] Koç University,[3] & Harvard University[3]

Address for correspondence:

Duygu Özge[1]

Harvard University

33 Kirkland Street

02138 Cambridge MA

Phone: 0090 505 787 5035

duyguo@metu.edu.tr

Word count: 20049 (includes abstract, introduction, methods, results, discussion; excludes cover page, figures, and references)

[1] Present address: Middle East Technical University, 06800 Ankara

ABSTRACT

During language comprehension we must rapidly determine the thematic roles of arguments (who did what to whom) in order to semantically integrate the players into a single event and predict upcoming structure. While some languages signal these relations mostly with reliable word order, others rely more on case markers. The present study explores whether Turkish-speaking children use case marking predictively during online language comprehension. Specifically, we use the visual-world paradigm to test whether 4-year-olds (and adults) can use a contrast between nominative and accusative case on the first noun to predict the referent of the second noun in verb-medial and verb-final spoken sentences. In verb-medial sentences, both children and adults used case to predict the upcoming noun, but children did so only after hearing the verb. In verb-final structures, however, both children and adults made predictive looks to the correct referent prior to the second noun (and the verb). Thus, Turkish-speaking preschoolers interpret case marking incrementally, independent of the verb, and use it to anticipate the upcoming argument. These findings are inconsistent with the hypothesis that the online interpretation of case marking depends on a late maturing neural circuit. The predictive use of case at four provides strong evidence that children's comprehension relies on broad, abstract mappings between syntax and semantics, which allow children to determine the event role of a case marked argument prior to identifying the verb.

An essential part of language comprehension is determining how the meanings of words come together to make the meaning of the utterance. One critical component of this is identifying the thematic roles of the arguments in a clause (in short, understanding *who did what to whom*). While languages have diverse ways of encoding these thematic roles, two patterns have been widely observed (Jespersen, 1922; Sapir, 1921; Weil, 1844/1887). First, in about 35% of the world's languages, the subject typically appears before the verb while the object typically occurs after it (Dryer, 2013a). Most of these languages have little or no case marking on common nouns (Dryer, 2013b; Iggesen, 2013), and many, like English, have strict word order constraints (Steele, 1978; Newmeyer, 2001). In such a language, the identity of the verb and the order of the nouns reveal the role that each noun plays in the event. For example, an English speaker who hears a sentence that begins as in (1) can surmise that *Gamera* is the agent of the action and that the noun phrase that follows *attacks* will be the patient.

(1) Gamera attacks the…

Critically, this inference depends not just on the order of the noun phrases, but also on the verb in the utterance; in (2) we interpret *Gamera* as the experiencer of the emotion (and expect the stimulus to follow), while in (3) we interpret *Gamera* as the stimulus and expect an experiencer.

(2) Gamera fears the…

(3) Gamera frightens the…

In contrast, about 42% of the world's languages have a canonical order in which the verb follows its object (Dryer, 2013a). Many of these languages, like Turkish, are characterized by flexible word order (Steele, 1978; Newmeyer, 2001) and the systematic use of case markers to signal thematic roles (Greenberg, 1963; Dryer, 2013b; Iggesen, 2013). In languages like this, the verb often appears late and the position of a noun is a poor predictor of its thematic role. In such a language, the morphological case of the noun can provide immediate information about its thematic role and support inferences about upcoming arguments. For example, a Turkish speaker who hears a sentence that begins as in (4) is most likely to infer that *the rabbit* is the patient of some action (as yet unknown) and that another entity is the agent of this action (maybe a fox or hunter). This inference depends on the case marker: if the initial noun is unmarked (as in (5)), then it is highly likely to be interpreted as a nominative agent.

(4) Tavşan-ı birazdan…

rabbit-Acc soon

(5) Tavşan-∅ birazdan…

rabbit-Nom soon

In short, the cues that are available first and most reliably in a language will depend, among other things, on the canonical word order, the degree of word order flexibility, and the presence of morphological case. Consequently, there are systematic crosslinguistic differences in the optimal strategy for incremental thematic-role assignment. Unsurprisingly, adults are highly adept at using whatever sources of information their language provides. English speakers show a general bias to treat sentence initial nouns as agents and make quick use of the information in the verb to predict upcoming arguments (Boland & Tanenhaus, 1991; Trueswell, Tanenhaus & Garnsey, 1994; McRae, Ferretti & Amyote, 1997; Altmann & Kamide, 1999). Speakers of languages with morphological case and freer word order rapidly use case marking to determine the role of a noun, regardless of its position, and to predict the arguments that will follow it (Kamide, Altmann & Haywood, 2003; Kamide, Scheepers & Altmann, 2003).

We know far less about how these abilities develop, because the study of online language comprehension in young children is still relatively new. Are children, like adults, able to make incremental use of whatever information their language provides? Or are some cues intrinsically easier to acquire than others? The results to date are somewhat surprising. Previous studies in English have consistently found that young children use phonetic information to make lexical predictions (Swingley, Pinto, & Fernald, 1999; Fernald, Swingley, & Pinto, 2001), they also use verb information to resolve temporarily ambiguous thematic roles (Trueswell, Sekerina, Hill & Logrip, 1999; Snedeker & Trueswell, 2004) and to predict upcoming arguments (Nation, Marshall, & Altmann, 2003; Borovsky, Elman & Fernald, 2012; see Mani & Hueting, 2012 for data showing that German children can predict upcoming argument that is a thematic fit to the verb). But the findings for case-marked languages have been mixed.

Two studies, both using the visual-world paradigm, suggest that children *may* be able to make incremental use of case markers. First, Choi and Trueswell (2010) report a pattern of processing errors, which suggests that Korean-speaking five-year-olds use case to guide interpretation. However, because the case marker was not actually manipulated in this study, this inference is indirect. Second, Huang and colleagues provide evidence that five-year old Mandarin-speaking children use morphosyntactic information for thematic-role assignment prior to the main verb (Huang, Zheng, Meng & Snedeker, 2013). However, the morphemes that they tested (*ba* & *bei*) are typically analyzed as prepositions or co-verbs (Kit, 1998) rather than case markers. In addition, this study measured thematic-role interpretation by using a pronoun whose reference was disambiguated by its role. Consequently, the critical measure occurred after both arguments had been heard and thus the study does not address whether children use morphosyntax to *predict* upcoming arguments.

Curiously, recent studies of young German speakers strongly suggest that children have great difficulty using case markers during online processing. In a series of ERP and fMRI studies, Friederici and colleagues manipulated the case marking in German transitive sentences and found that children were unable to use this information for thematic analysis until six to seven years of age (Knoll, Obleser, Schipke, Friederici & Brauer, 2012; Schipke, Friederici & Oberecker, 2011; Schipke, Knoll, Friederici & Oberecker, 2012). These findings are both surprising and important. If this pattern generalized to other case-marked languages (and other constructions), then we would be forced to conclude that a good portion of the world's children are using an incremental thematic processing strategy that is poorly suited to the language that they are hearing.

But there are good reasons to doubt this. As we discuss below, German is in many ways atypical for a case-marked language: it is a mixed-headed language which is head-final in embedded clauses and head-initial in main clauses; the order of arguments is fairly predictable; accusative marked objects generally appear following the verb and argument drop is very rare, so object-initial sentences are not very common; case marking is marked both on the determiner or the noun; and the encoding of case involve ambiguities because the gender and number information is also embedded in it (Comrie, 2009; Bader & Haussler, 2009). Could Turkish be a better place to look for the early use of case? It is a verb-final language with highly flexible word order and argument drop, which makes object-initial orders quite frequent. Case information is marked via nominal suffixes and there is no additional grammatical information such as gender or number blended with the case morpheme. These features might be making the Turkish case system more reliable than other case marking languages like German. However, there are also justified reasons why one might be skeptical about this expectation as Turkish case system also involves critical ambiguities (for more information about Turkish word order and case, and its ambiguities, see Appendix 1).

The present study tests whether Turkish-speaking 4-year-olds can use nominative and accusative case to predict upcoming arguments in verb-medial and verb-final sentences. Our goal is to address two research questions. First, can young children use case marking to make thematic predictions, or do they wait to hear each argument, or perhaps rely solely on word order during comprehension? Second, if children are able to use case marking for thematic prediction, does this ability depend on accessing the verb or reliance on canonical word order? Or can they make these predictions even in verb-final sentences and in object-initial sentences? In the remainder of the introduction, we motivate these questions by discussing: 1) theories of how children acquire thematic mappings and their relation to online interpretation, and 2) prior studies on the use of case marking during language comprehension.

Thematic role assignment in theories of language development

All theories of language acquisition must explain how children learn the syntactic devices that their language uses to encode semantic roles. The explanations that are offered vary widely. For example, some domain-specific nativist theories propose that children have innate abstract syntactic categories, innate thematic roles, and innate mappings between the two (Pinker, 1984; Baker, 1988). On such a theory, the primary problems of acquisition are learning the meanings of the words and the morphosyntactic devices that are used to mark the abstract syntactic categories. Other theorists do away with the innate syntactic and semantic categories and mappings and instead posit that children learn the relevant syntactic cues by noting the correlations between meaning and form (Bates & MacWhinney, 1987). Two theoretical questions are particularly relevant to the present studies.

1. Are some syntactic cues, such as linear order, privileged during development?

There are a couple reasons why order might be a more readily available cue to children than case marking. Detecting order does not require prior linguistic experience. In fact, the ability to represent the linear order of elements is a prerequisite for learning in *any* domain where information is distributed over time. In addition, some theorists have argued that human beings have an innate bias to assume that the agent will come before the patient (McNeil, 1975; Osgood & Bock, 1977; Osgood & Tanz, 1977; Byrne & Davidson, 1985): After all, the vast majority of languages with a dominant word order follow this pattern (97% according to Dryer, 2011) and the pattern has a plausible nonlinguistic precursor in the need for a cause to precede its effect. In contrast, case markers are purely linguistic forms with no obvious foothold in prelinguistic cognition. The specific phonological forms used for this function vary across languages, and thus they must be acquired through experience.

If order is an easier cue to acquire, then we would expect: that children would show delays in acquiring consistent thematic mapping rules in languages that rely primarily on case marking; that children acquiring languages in which both case and order predict thematic roles would rely on order more than adults; and that children might be slower in making predictions on the basis of case-makers, particularly when they violate the cross-linguistic preference for agents to come before patients.

Nevertheless, there are good reasons, empirical and theoretical, to be skeptical of the claim that order is developmentally privileged. If it were true, then there would be substantial pressure on languages, over historical time, to adopt word order as the primary device for thematic mapping.[2] Yet, according to the World Atlas of Language Structures (WALS), languages with flexible word order and case marking are quite common (Haspelmath, Dryer, Gil, & Comrie, 2011). Second, while the concept of order might be a perceptual primitive, the application of this concept requires that children know which units they need to track the position of (e.g., noun phrases rather than words, syllables or phonemes). Third, during the pre-linguistic period infants soak up the distributional structure of their native language, building up perceptual representations of function morphemes and the local contexts in which they occur (see e.g., Gomez & Lakusta, 2004; Hallé, Durand, & Boysson-Bardies, 2008; Hohle & Weissenborn, 2003; Lany & Gomez, 2008; Marquis & Shi, 2012; Mintz, 2006; Shafer, Shucard, Shucard, & Gerken, 1998; Shi, Werker & Cutler, 2006; Shi & Lapage, 2008). These analyses could provide a list of candidate morphemes for infants to consider as they try to crack the code for mapping thematic roles to syntactic positions. Finally, case markers conform to what is arguably the most general principle of the syntax/semantics interface: that every chunk of the conceptual representation be expressed by some chunk of

[2] We do not claim that the developmental pressure is the only pressure on language change; there must be other competing pressures from other domains that justify cross-linguistic variety in word ordering. Here, we simply say one-for-all strategy (e.g., Agent-first) would have put half of the world's people at a processing disadvantage.

the phonological structure. Using order to express the conceptual roles of arguments violates this assumption. If children base their approach to thematic interpretation on their earlier experience of word learning, then linking *order* to thematic roles may not be the first hypothesis that occurs to them.

2. How broad are children's initial generalizations?

The nature of early generalizations has been the focus of considerable research over the past two decades largely due to the prominence of Tomasello's verb-island hypothesis (1992). Among the recent versions of this perspective, the most relevant to our study are the usage based (constructivist) account (Tomasello, 2003) and the graded representations hypothesis (Abbott-Smith, Lieven and Tomasello, 2008), which we will collectively refer to as late-abstraction accounts in the rest of the paper.

According to these accounts, children around 22-24 months of age are already productive in combining pivot schemas such that they arrive at categorizations of some sort. However, these are more like unanalyzed chunks devoid of any sort of syntactic marking of different roles. Children then move on to a new stage where they begin to use syntactic markings to mark participant roles only within the constructions that are familiar to them, so there are still no generalized semantic or syntactic roles in this stage. Crucially, late abstraction accounts expect verb-based interpretation of syntactic/morphosyntactic marking even after children begin making more general abstractions. Age 3 is marked as a critical milestone where abstractions begin to emerge (Tomasello, 2003). However, the abstraction process is not an abrupt point in development. It begins around 3 years of age (slightly earlier around 2;6 or slightly later around 3;6-4;0 depending on the language and depending on the individual child) and continues gradually until children compile enough samples of constructions, onto which analogies are made.

Structural relationships and groupings of the constructional islands at this stage are not equal to abstract semantic representations. Based on these files of inventories, children slowly begin to abstract out the relationships between these constructions. This process leads to the so-called *abstract constructions*. However, these constructional representations are far from reflecting the abstraction of syntactic and semantic roles either, they are 'constructional representations that are less than fully abstract' (Boyd & Goldberg, 2012; p. 459). This abstract-constructions stage is not complete until late in acquisition; and reaching abstract semantic and syntactic roles is 'triggered in the end by the need to master constructions involving cross-clausal control as children approach school age' (Tomasello, 2003; p.169). This is attributed to the late development of the necessary processing strategies that are most critical for reaching broader abstractions (e.g., preemption, generalizations on the basis of semantic verb classes, and analogy-making).

In line with this, Boyd & Goldberg (2012) show that 5-year-old children are less successful than 7-year-olds and adults at understanding the linking rules behind a novel construction, given the same amount of input. The authors link this to limited pattern-detection and abstraction capacities of 5-year-old children compared to older children and adults. Thus, according to this account, children still struggle with prolonged item-based behavior and limited analogy-making abilities at age five.

In short, proponents of late abstraction accounts noted that, if the shift toward broader generalizations is prolonged and gradual process, then we should expect children's language processing to reflect the lexically-based categories long after broader generalizations begin to emerge in novel-verb inference tasks (Boyd & Goldberg, 2012; Savage, Lieven, Theakston &

Tomasello, 2003). If this account is correct then children learning verb-final languages would face substantial difficulties during online comprehension. Under the late abstraction accounts, children's representation of case markers should be specific to individual verbs or individual constructions until they have full abstractions about the semantic and syntactic roles. Without broad thematic categories, they should be unable to determine the role of an accusative or nominative noun phrase until after they encounter the relevant verb or they find out about the particular construction. This delay in integrating the nouns into the event could make it more difficult for children to remember the nouns and their case markers and might result in a processing breakdown when the child reaches the verb and has to integrate all the arguments at once. A strong support for this comes from German studies showing that children as old as five or six still fail to interpret case and rely on the verb until age four and on the word order information until age 6 or 7 (Knoll, Obleser, Schipke, Friederici & Brauer, 2012; Boyd & Goldberg, 2012).

The alternative hypothesis is that children build broad mappings between syntactic features and thematic roles from the earliest relevant stages of language acquisition (Gertner, Fisher & Eisengart, 2006; Fisher, 2002). Early abstraction is predicted by theories where children are innately equipped with language-specific representations of the relevant breadth (e.g., Pinker, 1984; Gleitman, 1990). However, it is also expected on any theory in which the learner is biased to represent events in terms of their broad semantic roles and to look for systematic patterns across utterances, regardless of the verb and construction that is used. We will refer to such set of accounts as early-abstraction theories. Over the past decade or two, the evidence has mounted that children under the age of two have broad mappings between syntax and semantics that allow them to interpret sentences with novel verbs and produce sentences in new constructions (Gertner, Fisher, & Eisengart, 2006; Yuan & Fisher, 2008;

Yuan, Fisher & Snedeker, 2012; Kline & Demuth, 2014; Naigles, 1990; Arunachalam & Waxman, 2010). In fact, priming studies suggest that language production and comprehension in preschoolers is more dependent on abstract structure, than it is in adults (Thothathiri & Snedeker, 2008; Rowland 2012). Evidence that preschoolers are able to interpret case independently of the verb would contribute to this emerging picture.

Experimental work on the use of case marking during interpretation

Several prior studies have found that children have difficulty interpreting thematic roles in non-canonical structures. Dittmar, Abbot-Smith & Tomasello (2008) investigated the comprehension of the accusative and nominative case markers in German-speaking monolingual children (at the ages of 2;7, 4;10, and 7;3) with two offline tasks. In both studies, children were tested on their interpretation of spoken utterances with canonical and non-canonical NP1-Verb-NP2 orders. Until about age 4, they were highly verb dependent: they correctly interpreted the structures with the familiar verbs but failed to use either word order or case marking cues in structures with nonsense verbs (for similar findings from other languages, see Chan, Lieven & Tomasello, 2009; c.f., Dabrowska & Tomasello, 2008). In contrast, the four- and five-year-olds used word order, both with known and novel verbs, but largely ignored the case marking cues. Only at the age of seven did the children assign the sentence-initial accusative marked NP to its correct thematic role, showing that they were able to use case marking to override word order.

Subsequent studies with German children have found parallel patterns using ERP and fMRI paradigms (Schipke, Friederici, Oberecker, 2011; Schipke, Knoll, Friederici, Oberecker, 2012; Knoll, Obleser, Schipke, Friederici & Brauer, 2012). For example, in an ERP violation paradigm, German-speaking children between the ages of 4 and 6 showed a pattern of effects, which suggested that their thematic predictions were completely dependent on word order. Specifically, upon hearing a sentence fragment with an initial nominative NP, they experienced a processing disruption (a P600) when they encountered a second nominative argument, much like adults (Frisch & Schlesewsky, 2005), suggesting that they were expecting an accusative object. However, when the same children heard a sentence with an initial accusative NP they also showed an increase in positivity to a subsequent nominative, suggesting that despite the case marker on the first noun they were still expecting to find an object.[3] The authors have posited that this inability to interpret case is caused by the slow maturation of the dorsal pathway between the temporal cortex and Broca's area. Research shows that children –until 7 years of age– lack fully myelinated dorsal pathways connecting the Broca's area to the posterior temporal cortex (for a review, Friederici, 2011). This region is proposed to serve the computation of hierarchical syntactic operations responsible for the processing of complex sentences with non-adjacent dependencies or dislocated arguments (e.g., object-initial sentences, wh-questions, relative clauses) (Brauer, Anwander, & Friederici, 2011). We will refer to this perspective as late neural syntactic maturation hypothesis.

We further know from Friederici & Frisch's (2000) study that German-speaking adults show left lateralized biphasic negativity-P600 in argument-structure violations. This negativity is attributed to the incongruency between the case marker on the sentence-initial object noun and the clause-final verb. Importantly, this pattern is taken as a reflection of

[3] While Schipke and colleagues reach similar theoretical conclusions, they interpret this effect as an N400 to the violation in the double accusative sentences rather than as a P600 to accusative-nominative sentences. The statistical analysis does not strongly support either interpretation. We favor the one given above on the basis of parsimony--it provides a single explanation for children's adult-like and non adult-like patterns.

event structure prediction on the basis of case marking information. In other words, the dorsal fibers connecting Broca's area to posterior temporal cortex are also taken to be crucial in predictive interpretation of the morphosyntactic cues. Given immature dorsal fibers until age 6 or 7, children should lack adult-like biphasic negativity in object-initial sentences and any structural prediction on the basis of case marking cues. This is exactly the pattern observed in Schipke, Knoll, Friederici, and Oberecker's (2012) study that measured ERP patterns in object-initial versus subject-initial sentences in German. They report that children only begin to show adult-like negativity after the object-initial sentences at age 6. Even at this age, they differ from adults at the onset of the second noun in that they fail to integrate the second argument into the available structure; hence fail to assign correct thematic roles to the arguments offline.

If we assume that the neural system subserving complex syntactic processing matures at around the same age for all children irrespective of the typological features of the language they acquire, then this hypothesis predicts that children acquiring languages other than German should also have difficulty interpreting case markers and using morphosyntax in predictive manner, but they should instead rely heavily on word order.[4]

This prediction is counter-intuitive given what we know about the acquisition of case. Probably the best-studied example of an early emerging case system is Turkish. Turkish-speaking children use case marking productively in their spontaneous speech from

[4] Although the basis of the reasoning in the late abstraction accounts and late neural syntactic maturation accounts differ theoretically, these perspectives expect children either to lack the necessary cognitive mechanism (e.g., analogy-making abilities) or the mature neural architecture (e.g., dorsal fiber bundles) enabling them to acquire and process complex structures until late in acquisition.

around the age of two (Aksu-Koç & Slobin, 1985; Ketrez & Aksu-Koç, 2009; Ural, Yüret, Ketrez, Koçbaş & Küntay, 2009). Accusative case, as a marker of the direct object, is the earliest case morphology to appear in children's speech (Ketrez, 2004). In act-out tasks, two-year-olds are able to use nominative and accusative case to reliably interpret sentences with both canonical and non-canonical word order (Slobin & Bever, 1982; Batman-Ratyosyan & Stromswold, 1999). In addition, Göksun, Küntay & Naigles (2008) demonstrated that two- to five-year-old Turkish-speaking children produce more causative enactments of intransitive verbs in two-argument structures when there is a sentence-initial accusative-marked object. This suggests that children use case marking as a cue to transitivity. Critically, however, all of these studies have two properties in common. First, they are offline: they measure the product of comprehension and production but, unlike the Friederici studies, they do not tell us about how these products are constructed during language processing. Second, the sentences in all of these studies contained a familiar verb, thus it is not clear whether children can interpret case marking in isolation or only in conjunction with the verb or within the whole construction.

To the best of our knowledge there is only one published study that explores online language processing in Turkish-speaking children (Özge, Marinis & Zeyrek, 2015). This experiment used a self-paced listening paradigm to assess processing of relative clauses in adults and 5-8 year old children. All groups showed slower reaction times to the genitive marked embedded subject in the object relative clause as compared to the accusative marked embedded object in the subject relative clause. While this could reflect incremental thematic interpretation, it could also reflect differences in the frequency, distribution and ambiguity of the two markers. Crucially, Özge et al.'s (2015) findings are based on children between the ages of 5 and 8, so we still do not know whether younger Turkish-speaking children are able to use case morphemes to incrementally assign thematic roles and predict the upcoming arguments during the course of online processing.

Goals of the current studies

The present study seeks to fill this gap. We use the visual world eye-tracking paradigm to test Turkish-speaking children (4;0-5;0 years old) and adults on sentences where the first NP has either accusative or nominative case. By doing this, we can address two questions. First, do children use case marking to anticipate the role of the upcoming argument during real-time language comprehension or do they primarily rely on word order (interpreting the first noun as an agent regardless of case)? Second, do children need to hear the verb in the sentence before they can interpret the case marker or are they able to use case to assign an abstract role independent of the verb? Being able to predict upcoming information independent of the verb is crucial because the verb would provide information about the argument valency and selectional restrictions, thereby making prediction via mere lexical associations easier. In Experiment 1, we explored whether children and adults could use the case on the first argument to predict the thematic role of the second argument in verb-medial structures (i.e., SVO vs. OVS). In these cases, prediction can occur after verb information is available. In Experiment 2, we used verb-final structures (i.e., SOV vs. OSV) in which any anticipation of the second argument must, by necessity, occur prior to the verb.

Experiment 1

Our experiments were modeled on a study by Kamide, Scheepers, and Altmann (2003), which demonstrated that German-speaking adults can use the case of the first argument (nominative or accusative) to predict the second argument in a transitive sentence with either an SVO or OVS order. In this study, participants were shown a visual context with three related referents (e.g., a hare, a cabbage, and a fox) and heard a pre-recorded sentence (e.g., *'The hare-Nom eats soon the cabbage-Acc'* versus *'The hare-Acc eats soon the fox-Nom'*). During the adverb region, but not during the verb region, the adults looked more often at the plausible patient (i.e., the cabbage) in the nominative (i.e., Nom-V-Acc) condition and more at the plausible agent (i.e., the fox) in the accusative (i.e., Acc-V-Nom) condition. This effect demonstrates that German-speaking adults are able to use case to predict an upcoming argument, though it still leaves open the possibility that they need information from the verb to do so.

In another study, Kamide, Altmann, and Haywood (2003) presented Japanese adults with a visual scene with a waitress, a customer, a hamburger, and a baby-high chair, accompanied by spoken utterances in dative and accusative condition (e.g., *waitress-Nom customer-Dat* vs. *waitress-Nom customer-Acc*). There were more looks to the plausible theme (*hamburger)* in the dative condition compared to the accusative condition, suggesting that adult speakers are able to interpret case markers to predict the thematic role of the upcoming argument. Different from Kamide et al.'s (2003) German study, Japanese adults were able to predict the upcoming referent before they heard the verb in this study. However, the contrast in this study was between the double-object constructions and the transitive constructions. The authors discuss that the interpretation of the dative case along with the lexico-semantic information provided by more than one noun (waitress + customer) might have led to the expectation of an extra object. Thus, although these studies strongly suggest that adults integrate case markers in a predictive manner, there is still a possibility that case marking and the meaning of the only available noun on its own may not be enough for thematic prediction even for adults.

In the present study, we used verb-medial structures similar to Kamide and colleagues' German study to test whether children are able to use the case marking on the first noun to anticipate the second argument. Different from Kamide and colleagues' (2003) Japanese study, our experiment requires participants to rely only on the identity of the first noun and its case marker to be able to predict the upcoming argument. Our linking assumption is that the parser is an eager parser (Earley, 1970). It would incrementally interpret the case marker on the first noun (together with the world knowledge triggered by the visual context), create the most plausible/probable hypothesis about the thematic role of this given noun, which would automatically lend itself to the expectations about the upcoming event structure and its possible arguments (for a review, see Levy, 2008). If children are able to use case markers to make thematic predictions, then we should see more looks to the agent in the accusative condition compared to the nominative condition. It is important to note that the scene does not prevent the participants from activating all possible events that involve the three referents on the visual scene. For instance, the first nominative marked referent (e.g., *rabbit*) might act on the *carrot* (e.g., *The rabbit will soon eat the carrot*) but it could also act on the *fox* (e.g., *The rabbit will soon flee from the fox*). The same also applies to the accusative condition. On the basis of the findings of previous studies, we simply assume that anticipating a patient would be the most likely choice in the nominative condition and anticipating an agent would be the most likely choice in the accusative condition (for a similar discussion, see Kamide, Altmann, & Haywood, 2003). However, we might have little differences in the looks to either of the referents or frequent shifts between them.

Another assumption here is that we expect predictive looks to the upcoming plausible argument also in Turkish despite the fact that it frequently allows argument drop. We believe that the processor would still create a conceptual representation of all arguments regardless of whether or not they are overtly available in the utterance. Previous filler-gap processing literature is also in line with our assumption that missing or dislocated arguments are conceptually reactivated as soon as the parser realizes that there is a slot for this argument (Frazier, 1987; Frazier et al., 2015; Omaki et al., 2015). This is also supported by the pragmatic constraints on argument drop in Turkish dictating that the referent should be recoverable from context in order to be dropped. Thus, we think that the argument should be conceptually represented regardless of whether or not it is overtly stated.

In this experiment, we expect to find the biggest predictive effects immediately before the target noun (in anticipation of the referent). In verb-medial structures these looks could reflect either abstract thematic prediction on the basis of the case marker or the integration of case and verb-specific role information. Experiment 2 addresses this more directly.

Method

Participants

Our participants were 37 typically developing children (4;0-5;0 years, M = 4;5) and 39 undergraduates at Koç University. Turkish was the first language of all participants. Our participants were reported to have vision that was normal or corrected to normal through the use of contact lenses. The adult participants and the parents of the children reported that they had no known neurological or behavioral problems.

Stimuli

The critical test items consisted of verb-medial mono-transitive simple sentences with two overt arguments, as in (6) and (7). The first noun was in nominative case in the SVO order (i.e., N_{Nom}-V-N_{Acc}) and in accusative case in the OVS order (i.e., N_{Acc}-V- N_{Nom}). The first noun was preceded by an adjective, making it natural to place a prosodic boundary after this noun phrase. As a result that case marker occurred at the end of prosodic phrase making it more salient, and more clearly articulated, than it would otherwise be.

(6) Nominative condition (SVO: N_{Nom}-V-N_{Acc}*):*

Minik	tavşan-∅	birazdan	yiy-ecek	şurada-ki	havuc-u.
little	rabbit-Nom	soon	eat-Fut	there-Rel	carrot-Acc

"The little rabbit will soon eat the carrot over there."

(7) Accusative condition (OVS: N_{Acc}-V- N_{Nom}*):*

Minik	tavşan-ı	birazdan	yiy-ecek	şurada-ki	tilki-∅.
little	rabbit-Acc	soon	eat-Fut	there-Rel	fox-Nom

"The fox over there will soon eat the little rabbit."

We included the time adverbial *birazdan* (soon) between the first noun and the verb and the modifier *şurada-ki* (there-Rel: the one over there) between the verb and the second noun to provide enough time windows both before and after the verb, in which predictive effects of case might emerge.

There were 20 critical items (10 for each condition), (Appendix 2). We used the following agentive verbs in the critical sentences: *kick, swallow, kiss, hold, bite, lick, eat,* and *find*. Each verb was used once, except for *eat* and *find*, which were used twice. There were 10 filler items containing intransitive sentences with the subject composed of a complex noun phrase (NP) marked in the genitive case, as in (8), (Appendix 3).

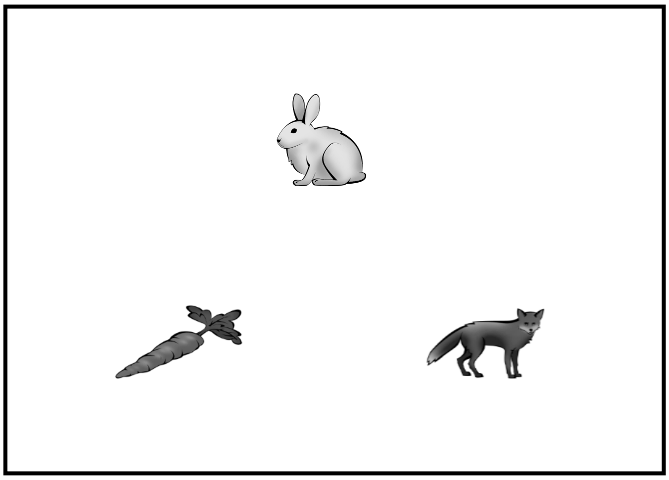
(8) Ağaç-ın yan-ın-da-ki kurbağa birazdan zıpla-yacak.

tree-Gen near-Poss-Loc-Rel.3sg frog soon jump-Fut

"The frog next to the tree will soon jump."

Each critical sentence was paired with a visual scene with three referents (Figure 1): 1) the topic (i.e., the expressed noun; e.g., *rabbit*), the referent of the first noun in the sentence and thus the entity that the sentence is about; 2) the plausible agent (e.g., *fox*), an entity who could reasonably be expected to perform some action on the topic and the referent of the second noun in the OVS sentences; and 3) the plausible patient (e.g., *carrot*), an entity which could be affected by an action performed by the topic but is unlikely to act on the topic and the referent of the second noun in the SVO sentences. A visual scene with three objects also accompanied each filler item. The pictures and animations were prepared by a professional artist. All pictures were in color and at the resolution of 640 x 480 pixels. The location of each object on the screen was pseudorandomized such that the topic, the plausible agent, and the plausible patient appeared equal often in the three positions on the screen (upper right, upper left and lower middle).

Figure 1: Gray-scale version of a sample visual display for the following sentences



Experiment 1:

a- Minik tavşan	birazdan	yi-yecek	şura-da-ki	havuç-u.
little rabbit-Nom	soon	eat-Fut	there-Loc-Rel	carrot-Acc

“The little rabbit will soon eat the carrot over there.”

b- Minik tavşan-ı	birazdan	yi-yecek	şura-da-ki	tilki .
little rabbit-Acc	soon	eat-Fut	there-Loc-Rel	fox-Nom

“The fox over there will soon eat the little rabbit.”

Experiment 2:

a- Minik tavşan	birazdan	şura-da-ki	havuç-u	yi-yecek.
little rabbit-Nom	soon	there-Loc-Rel	carrot-Acc	eat-Fut

"The little rabbit will soon eat the carrot over there."

b- Minik tavşan-ı	birazdan	şura-da-ki	tilki	yi-yecek.
little rabbit-Acc	soon	there-Loc-Rel	fox-Nom	eat-Fut

"The fox over there will soon eat the little rabbit."

To see to what extent the topic entity (rabbit) is related to the plausible agent (fox) and plausible patient (carrot) on the visual scene, we conducted a norming study with adult participants. Fifty-five Turkish speakers were presented with a pair of items composed of the topic and the plausible agent (e.g., rabbit-fox) or the topic and the plausible patient (e.g., rabbit-carrot) and were asked to rate how closely related the two items were) using a 7-point rating scale (1 = not related at all, 7 = very related). Relatedness was defined as the likelihood that the two items would appear together. The data for this study is publicly available in Özge, Küntay, & Snedeker (2018). Mann Whitney-U test showed that there were five items with a patient (object) bias (i.e., marked as more related with the plausible patient than the plausible agent it appeared with), namely the chick in relation to the wolf (agent) and the corn (patient), the monkey in relation to the lion (agent) and the banana (patient), the squirrel in relation to the fox (agent) and the nut (patient), the bear in relation to the hunter (agent) and the honey (patient) and one item with a agent (subject) bias (the girl in relation to the woman and the doll). This resulted in an overall patient preference ($U = 381.5$, $p < .01$). Our eye-tracking experiment aims to test whether children can interpret the accusative and the nominative case in Turkish to predict the upcoming argument. If there was a bias to mark the first referent as agent regardless of the case marking of the first noun, then we would expect more predictive looks to the patient entity. In such a case, these five items with the patient preference would further strengthen the first-entity-as-agent bias. If on the other hand we find an effect of case, we would expect greater agent looks in the accusative condition compared to the nominative condition. This would mean the overall expectation to relate the topic entity to the plausible patient entity can be overridden with the interpretation of case markers.

A female native speaker of Turkish recorded the test sentences with a focus accent on the verb. There was no focus accent on the first noun to avoid the implication that the sentence-initial referent was to be topicalized or contrasted with another referent. We edited the sound files to control for the duration of the pauses between phrases. Each utterance was preceded by a 200-milliseconds of silence and a1500-ms of silence at the end of the utterance before the animation began. There was a natural prosodic break of around 300 ms after the first noun. To ensure that its length is the same across all items, we replaced it with a 300-ms of silence. No other time window was edited to ensure naturalness.

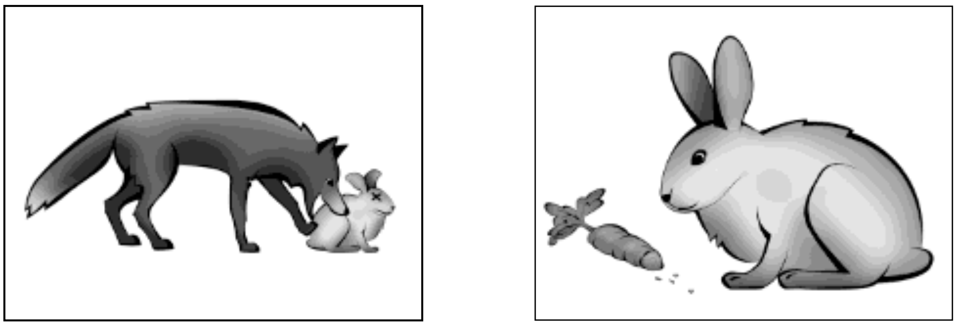
The case marking of the first noun was manipulated within subjects. Two counterbalanced lists were constructed so that each one contained 5 items in each condition and each item appeared in both conditions across the two lists. Each list was constructed so that two critical items from the same condition never occurred back to back. The same fillers were used across the lists.

Procedure

Each participant was seated in front of the screen of the Tobii T60 eye-tracker and calibrated using Tobii Studio. The presentation of each trial was gaze contingent to ensure that participants were looking at the screen and being tracked throughout the study. If the participant got up from their seat during the experiment, the calibration process was repeated.

At the beginning of the trial the three objects appeared on the screen. In the version of the experiment for the children, each object was named as it appeared on the screen. In the adult version, there was no naming of the referents. The spoken utterance, which was played over external speakers, began 200 ms later (due to the silence at the beginning of the sound file). There was 1500 ms of silence at the end of the sentence, then the objects disappeared, and a simple animation began. Following the animated videos, participants saw the final state of the depicted event as a still picture. On half of the trials, the animation depicted the event described in the sentence, and on half of the trials it depicted some other event. For the critical trials, the incorrect event was always the event that would have been correct in the other condition. For example, given a target sentence such as '*The rabbit will soon eat the carrot'*, the correct animation showed the rabbit eating the carrot while the incorrect animation showed the rabbit being eaten by the fox (Figure 2). They were told to listen to the sentence while looking at the scene on the screen, watch the animation and then tell the experimenter whether the event in the animation matched the sentence that they had heard. The participants' responses were coded by the experimenter during the session. Both OVS and SVO sentences were followed by incorrect events on half of the trials.

Figure 2: Gray-scale version of sample still images shown after the animated video



Data analysis

Below, we will first present the offline performance of our participants in the end-sentence animations. This will allow us to see whether they listened to the utterances carefully for comprehension.

We will then graph their eye-gaze patterns to understand how the pattern of their looks changed through the course of the utterances they heard. For the purpose of these graphs, we synchronized the files at the onset of the speech stream and divided fixations into 100-ms time-windows that continued until the end of the sound file (1500 ms after sentence offset). Our eye-tracker samples the position of participants' gaze 60 times per second, which corresponds to around once in 17 ms (or around 3.5 looks per second and around 5.88 looks per 100 ms). We processed the gaze data using a program written in Python to count the number of looks to each referent (i.e., how many times a participant shifted her gaze from one location to another on the screen) in each 100-ms time-window. We then divided each sentence into five time windows: the first noun, the adverb, the verb, the modifier and the second noun. Each window began at the onset of the relevant word and ended at the onset of the next word. Thus the first noun window included the 300 ms pause that followed this noun. These time windows were not offset by 200 ms, because we wanted to ensure that any predictive looks in the modifier time window were truly predictive and not due to rapid effects of phonological information.

As our dependent measure, we will use a score we call *agent preference*: a binary variable indicating whether the participant looked more at the potential agent or the potential patient during a given time window (for a similar method, see Huang, Leech and Rowe, 2017). This measure best captures the underlying distribution of our data: While our eye-tracker samples gaze 60 times a second, people typically make only a couple of saccades in a second. Consequently, in a short time window, most participants will only fixate one of the objects, and thus any measure of fixation proportion within that window is essentially binary. To calculate agent preference we took the number of samples (for a given trial and a given window) in which the participant looked at the plausible agent and subtracted the number of samples in which they looked at the plausible patient. If this number was positive, agent preference was 1. If it was negative, then agent preference was 0. If participants looked at neither object during the time window (or both equally), then the trial was excluded from the analysis. Our primary interest will be in the looking pattern after the first noun and prior to the second noun. If participants are able to use case to predict the upcoming referent, then we should see greater agent preference during the predictive region (i.e., after the first noun and before the second noun) in the accusative condition than in the nominative condition.

We will then analyze the data using mixed-effects logistic regressions including the fixed effects for case (first noun accusative or nominative) and age (child or adult), and the interaction of these variables. We will conduct two omnibus analyses (i.e., analyses including multiple degrees of freedom) including first all five time-windows and then eliminating the final time-window revealing the identity of the second noun. Our aim in those analyses will be to see if the agent preference score would change through the course of the utterance, if the pattern would differ with respect to case marking, and if children would behave similarly to adults. Finally, we will have separate models for each time window to see if the agent preference significantly differs especially in the predictive time windows.

The data for both eye-tracking studies is publicly available in Özge, Küntay, & Snedeker (2018).

Results

Table 1 presents the percentage of correct responses to the end-sentence animations in each condition by children and adults. Both groups showed above 95% success rate in the judgment of the animations in relation to the utterance for both conditions.

Table 1: Percentage (%) of correct responses in each condition for children and adults

	Adults	Children
Nominative	99.7	95.92
Accusative	99.7	95.11

The eye-gaze patterns for the two groups of how agent-preference changed in each time window through the course of utterance are shown in Figures 3 and 4. From these figures, we see that adults show greater agent preference in the nominative condition compared to the accusative condition in the first time window during the first noun region. This is maintained until the end of the second time window (the adverbial region), where the pattern shifts with greater agent preference in the accusative condition compared to the nominative condition until the end of the sentence. In children, we see greater agent preference in the nominative condition towards the end of the first noun region until the middle of the adverbial region, where the difference between the two case markers is not so different. Children then begin to show greater agent preference in the accusative condition compared to the nominative condition towards the end of the verb region and this pattern peaks right after the verb region during the modifier region. We were mainly curious about whether the participants would show greater agent preference in the accusative compared to the nominative case between the adverbial and modifier region. From these figures, we see that adults show the expected pattern before the verb region while children show it later than adults (i.e., after the verb) in the modifier region. To further check if participants looked at the expressed referent and how their looks to each of the three referents changed in line with the utterance they heard, we also plotted the percentage of looking time to each of the objects (the topic, the plausible agent, and the plausible patient) during each of these time windows. These graphs are presented in Appendix 3 for the sake of simplicity and coherence in streamlining our data analysis here.

Figure 3: Gaze patterns of agent preference in each time window for adults in Experiment 1

Error bars indicate the standard error of the mean. Values above zero indicate preference to look at potential agent, and values below zero indicate a preference to look at the potential patient.

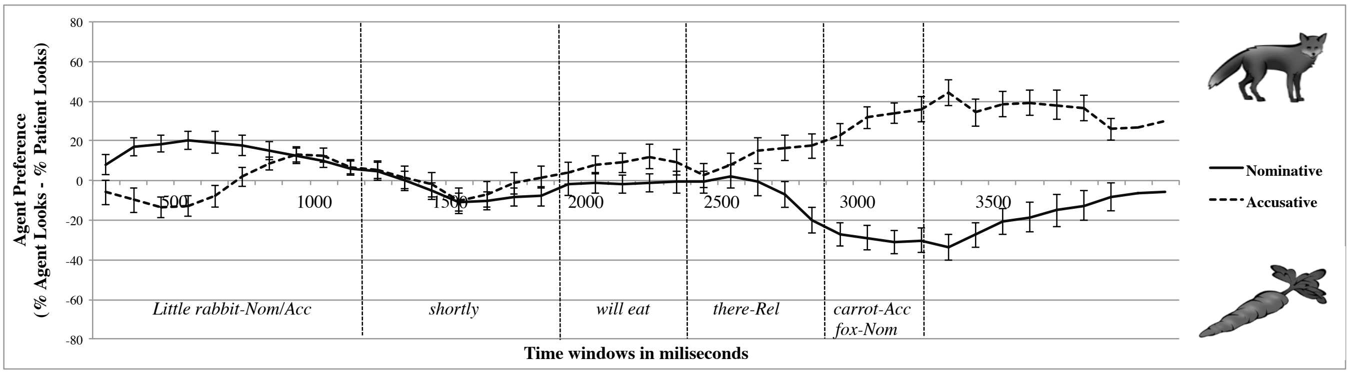
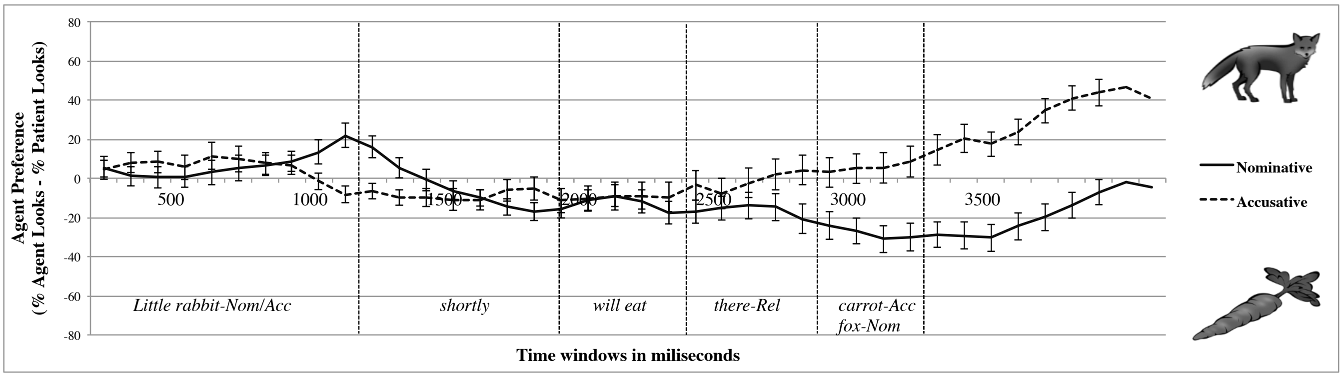


Figure 4: Gaze patterns of agent preference in each time window for children in Experiment 1

Error bars indicate the standard error of the mean. Values above zero indicate preference to look at potential agent, and values below zero indicate a preference to look at the potential patient.



For our statistical analyses, we used the agent preference score as our dependent variable and analyzed the data using mixed-effects logistic regressions which included fixed effects for case (first noun accusative or nominative) and age (child or adult), and the interaction of these variables.[5] For these binary variables, an effects-coding scheme was used, with the first listed level of each variable coded as 1 and the second as -1. Thus the main effects in these analyses can be interpreted as if they were ANOVA's. All reported analyses included random intercepts for both participants and items. We also conducted parallel analyses with all the potentially applicable random effects. When these analyses converged, they did not provide a significantly better fit to the data ($p > .05$) indicating that the additional random predictors were not justified.

Figure 5: Agent preference in each time window for the adults in Experiment 1

Error bars indicate the standard error of the mean. Values above .5 indicate preference to look at potential agent, and values below .5 indicate a preference to look at the potential patient.

---

[5] Following Barr (2008), we preferred to use mixed-effects logistic regressions for our analysis. However, to see if our results were compatible with the analysis used in Kamide et al. (2003) and Carminatti & Knoeferle (2013), we conducted two complementary analyses for the critical time window using the ANOVAS for Experiment 1 (Appendix 4). These analyses were compatible with the analysis we present here.

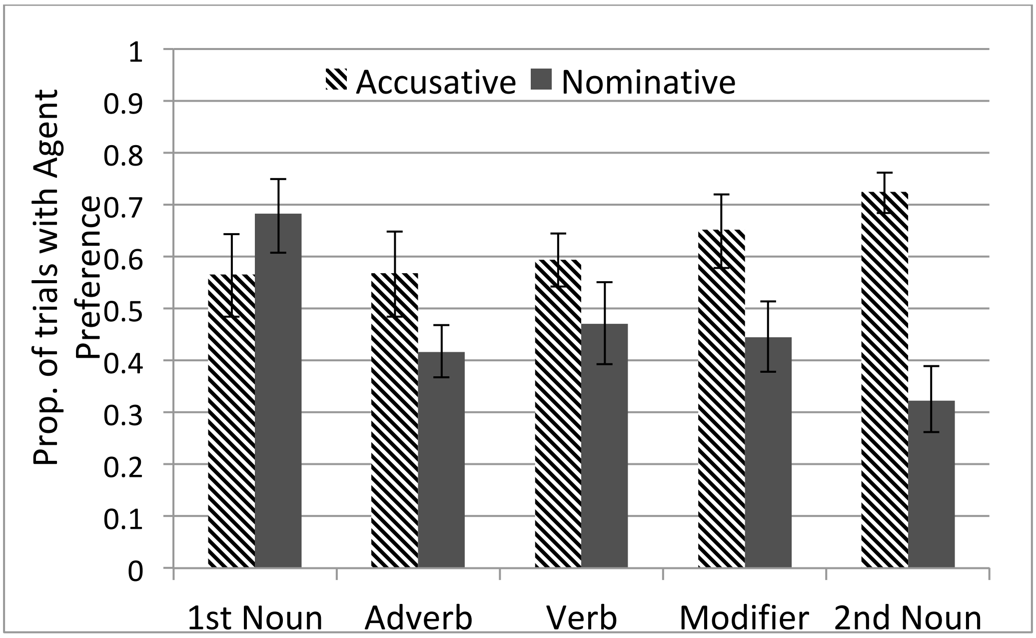
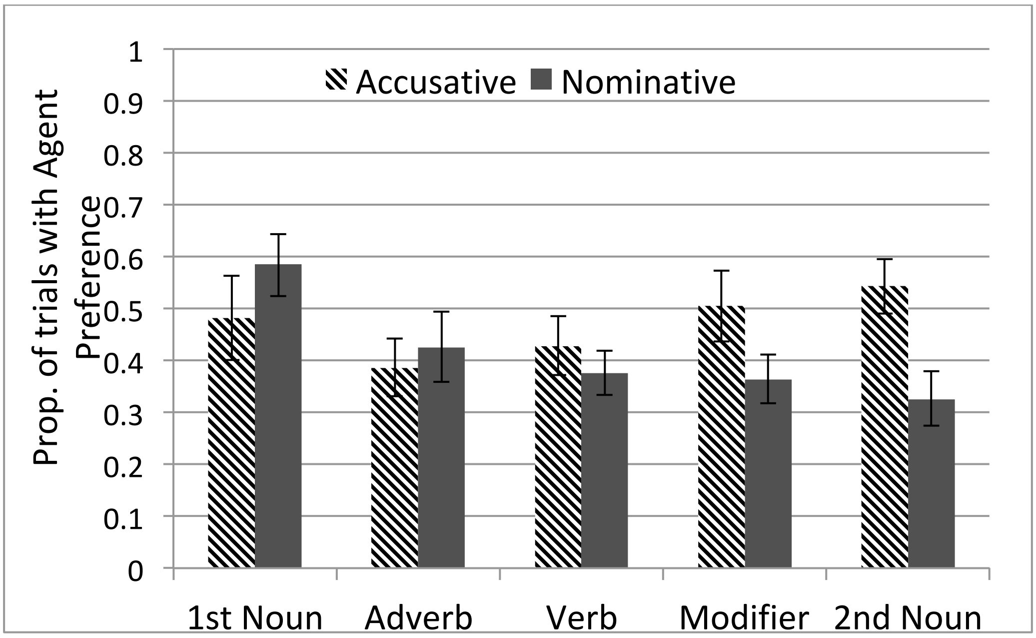


Figure 6: Agent Preference in each time window for the children in Experiment 1

Error bars indicate the standard error of the mean. Values above .5 indicate preference to look at potential agent, and values below .5 indicate a preference to look at the potential patient.



Figures 5 and 6 show the proportion of trials with an agent preference in each group in each of the five time windows (see above). To explore whether these scores changed over the course of the trial, and the role of case in any of these changes, we conducted an omnibus analysis of all five time windows, with fixed effects for case, age group and time window. The time window was coded with four variables (using an effects coding scheme with the first noun window as the reference). The predictors and their interaction terms were added to the model in the order listed, and the contribution of each one was assessed with a log-likelihood test (see "Full Sentence" in Table 2).[6]

[6] In cases where the new main effect or interaction corresponds to a single fixed effect in the model, this test provides essentially the same information to the z statistic for that effect. In those cases, the p-values for the two tests were very similar. However, the log-likelihood test also provides a single summary statistic for main effects or interactions that are spread across multiple fixed effects in the model (which occurs whenever a variable has more than two levels). Thus for consistency, we report log-likelihood tests for all fixed effects in the omnibus analyses.

This analysis revealed the following patterns. First, there is an effect of case in the expected direction: participants look more at the potential agent in the accusative condition than in the nominative condition. Second, there is an effect of age: children looked at the potential patient more often than adults, suggesting that they have a stronger bias for the first noun to be an agent. Third, there is an interaction of case and age: over the sentence as a whole, the adults were more likely to look at the correct referent of the final noun phrase than the children. Fourth, there is a main effect of time window, which reflects a decrease in agent preference as the sentence progresses. Finally, and most critically, there is an interaction between case and the time window, indicating that the differences between the two sentence types emerge as the utterances unfold. However, because this analysis included the entire sentence, this interaction is presumably driven in part by looks that occur after participants recognize the final noun, and thus it does not provide clear evidence for a predictive effect of case.

Table 2: Omnibus analyses of agent preference in Experiment 1

Effect Added	Full Sentence	First Noun and Predictive Windows
Case	$\chi^2(1) = 27.45$	$\chi^2(1) = 4.30$

	p < .0001**	p = .04*
Group	$\chi^2(1) = 13.16$ p = .0003**	$\chi^2(1) = 10.42$ p = .001**
Case * Group	$\chi^2(1) = 7.16$ p = .007**	$\chi^2(1) = 3.03$ p = .08†
Time Window	$\chi^2(4) = 18.98$ p = .0008**	$\chi^2(3) = 18.91$ p = .0003**
Case * Time Window	$\chi^2(4) = 52.96$ p < .0001**	$\chi^2(3) = 22.15$ p < .0001**
Group * Time Window	$\chi^2(4) = 1.85$ p = .76	$\chi^2(3) = 1.82$ p = .61
Case * Group * Time Window	$\chi^2(4) = 3.93$ p = .42	$\chi^2(3) = 2.25$ p = .52

The shaded cells are the effects that would indicate the predictive use of case.

† .05 < p < .1 * .01 < p < .05 ** p < .01

To explore the predictive effects more directly, a second omnibus analysis was conducted that excluded the final noun (see "First Noun and Predictive Windows", Table 2). The effect of age persisted in this analysis, with children showing a greater preference for the potential patient. In contrast, the interaction between age and case, which had been reliable when the second noun was included, was reduced to a trend. Critically, the main effect of case and the case by time window interaction persisted in this analysis, indicating that the case marker shaped interpretation *prior* to the second noun and that this effect shifted, in the predicted direction, over time.

To pin down the timing of the case effects and further understand the interactions in the omnibus analyses, we constructed separate models for each time window with case and age group as fixed effects (Table 3). We then followed up these analyses by looking at the each age group separately (Table 4). Taken together, these analyses show how the use of case changes as the sentence unfolds.

At the first noun, we found a reliable effect of case, but one which was in *the opposite direction* to the effect that we expected to find later in the sentence (Table 3). Participants who heard "rabbit" with nominative case were <u>more likely</u> to look at the fox, in the period during and immediately after the noun (see Figures 5 & 6). The effect was similar in the two groups (β = -.22 and -24), but it reached conventional significance only for the adults (Table 3). Given the timing and direction of this effect, it could reflect the use of case as cue to the identity of first argument, rather than as a basis for predicting the second noun. In other words, participants may be using the noun stem and the case marker in parallel to identify the relevant referent. The case marker on the first noun might be leading our participants to focus on the referents that could take on the thematic role that is consistent with this first case marker: the fox and the rabbit have features we expect for a nominative noun and the carrot and rabbit have the features expected for an accusative noun.

Table 3: Time window analyses for agent preference in Experiment 1

Time Window	Effect		
	Case	Group	Case * Group
First Noun + Pause	β = -.23 z = -2.69 p =.01**	β = -.16 z = -1.76 p =.08†	β = .006 z = .07 p =.95
Adverb	β = .09 z = .96 p =.34	β = -.14 z = -1.44 p =.15	β = -.16 z = -1.76 p =.08†
Verb	β = .18 z = 2.07 p =.04*	β = -.27 z = -2.81 p =.005**	β = -.05 z = -.60 p =.55
Modifier	β = .34 z = 3.97 p < .0001**	β = -.22 z = -2.21 p =.03*	β = -.06 z = -.69 p =.49
Second Noun	β = .62 z = 7.15 p < .0001**	β = -.20 z = -2.17 p =.03*	β = -.18 z = -2.12 p =.03*

The shaded cells are effects that would indicate the predictive use of case.

† $.05 < p < .1$ * $.01 < p < .05$ ** $p < .01$

Table 4: Analyses for the children and for the adults in Experiment 1

Time	Effect

Window	Child Intercept	Child Case	Adult Intercept	Adult Case
First Noun + Pause	$I = .13$ $z = .57$ $p = .57$	$\beta = -.22$ $z = -1.81$ $p = .07$†	$I = .46$ $z = 1.77$ $p = .08$†	$\beta = -.24$ $z = -2.01$ $p = .05$*
Adverb	$I = -.38$ $z = -2.11$ $p = .03$*	$\beta = -.08$ $z = -.63$ $p = .53$	$I = -.08$ $z = -.39$ $p = .69$	$\beta = .28$ $z = 2.00$ $p = .05$*
Verb	$I = -.41$ $z = -2.65$ $p = .008$**	$\beta = .14$ $z = 1.09$ $p = .28$	$I = .15$ $z = .95$ $p = .34$	$\beta = .23$ $z = 1.85$ $p = .06$†
Modifier	$I = -.28$ $z = -1.73$ $p = .08$†	$\beta = .29$ $z = 2.35$ $p = .02$*	$I = .19$ $z = .80$ $p = .42$	$\beta = .41$ $z = 3.33$ $p = .0009$**
Second Noun	$I = -.26$ $z = -1.97$ $p = .05$†	$\beta = .44$ $z = 3.64$ $p = .0003$**	$I = .14$ $z = .89$ $p = .38$	$\beta = .80$ $z = 6.56$ $p < .0001$**

The dependent variable is agent preference.

The shaded cells are effects that would indicate the predictive use of case.

† $.05 < p < .1$ * $.01 < p < .05$ ** $p < .01$

The results of the verb and adverb windows are suggestive but difficult to interpret. The adverb window is the first window where we thought we might find predictive effects of case. In the analysis of both groups, there is no hint of a case effect, but there is a trend toward an interaction between case and age (Table 3). Taken alone, the children are at chance, while the adults show a small but statistically significant effect of case, this time in the direction that would be expected if participants were using the case marker to predict the second noun (Table 4). In the verb window, the effect of case is robust when the two age groups are collapsed, and there is no evidence of an interaction (Table 3). However, when the age groups are analyzed separately, the effect fails to reach the conventional level of significance in either one (Table 4).

In contrast, the effects in the final predictive region, the modifier window, are clear and robust. The main effect of case is reliable when the groups are combined (Table 3), and the contrast is significant for each group in isolation (Table 4). Both adults and children show a greater agent preference for sentences where the first noun had accusative case, suggesting that they are predicting what the upcoming argument will be.

Finally, after the second noun begins and the referent is phonologically disambiguated, the difference between the two sentences becomes even larger. This final window is the only place where there is a clear interaction between age and case. While this could indicate that lexical interpretation is more sensitive to developmental change than thematic prediction, it could also reflect a difference in our ability to detect an interaction in each time window. The main effect of the case condition is much greater in the final window than in the earlier ones, making any modulation of this effect across age groups easier to detect.

Despite our initial relatedness study where participants showed greater lexical association between the expressed entity and the plausible patient entity, we found an effect of case in our eye-tracking study. Even if there was a tendency in children to treat the first entity as agent initially right after the first noun, we later found an effect of case after the verb region where they showed greater agent preference in the accusative condition compared to the nominative condition. This shows that the lexical associations observed in our relatedness study and first-argument-as-agent bias were overridden thanks to the case marking cues.

Also, one may raise the question as to whether the argument anticipations in the post-verbal region might be influenced by the word order frequencies of Turkish (Table 10 in Appendix 1). Could it be the case that our participants had greater agent preference in the accusative condition due to the fact that OVS (13.96%) is more frequent than SVO (2.25%) (i.e., anticipating a subject argument after an OV structure more than anticipating an object argument after an SV structure)?[7] In principle, we accept that frequency plays a role in parsing. However, we do not think that frequency is at play in this particular case because of the following reasons. First, we think the parser conceptually activates all arguments in all word order configurations regardless of whether or not the language or the utterance processed involves argument drop. Dropping of arguments is a pragmatically motivated choice on the part of the speaker such that only the arguments that are already derivable from context are omitted from the actual utterance. The listener has to link this referential expression (be it a pronoun or argument-drop) to its antecedent to be able to interpret the utterance successfully. Having an access to the same contextual constraints as the speaker, the listener can detect the identity of the omitted argument. This, however, does not mean that these omitted arguments are not conceptually activated (or anticipated). On the contrary, we know from discourse anaphora studies that the arguments that are more conceptually accessible are referred with shorter referential expressions or they are dropped altogether in argument drop languages (for a review see, Arnold, 1998). Therefore, we think the

[7] We thank one of our anonymous reviewers for bringing up this question.

participants would conceptually (hence visually) search for an antecedent of an anticipated argument regardless of whether there is a chance the argument could be grammatically omitted or replaced by a pronoun. This is why we think the post-verbal region is also relevant in prediction of arguments. Second, even if the dropped arguments had not been anticipated, frequency cannot explain the pattern here because OVS might be more frequent than SVO but OV (subject-drop) is much more frequent than both OVS and SVO. As shown in Table 10 in Appendix 1, 40.18% of the utterances involve subject-drop (OV), therefore if the listeners had been anticipating arguments on the basis of their frequency, they should have assumed that there would be no subject argument in the upcoming structure because OV is much more frequent than OVS, but this was not the case.

Taken together, these findings support three conclusions. First, in verb medial constructions, children, but not adults, have a bias to interpret the first noun as the agent and predict that a patient will follow. This was evidenced by the reliable negative intercepts for the children in the adverb and verb regions (Table 4), which indicate that there is a reliable patient preference and by the main effect of age group in the omnibus analyses (Table 2). Second, despite this bias children, like adults, can use the case marking on a noun to predict an upcoming argument. Third, in adults, the effect of case was visible in the adverb time window, prior to the onset of the verb. This suggests that adults are able to assign a thematic role to the first argument and predict the second argument, independent of the particular verb used in the utterance. Importantly, the predictive effects we observed in our adult participants were earlier than those in German adults in Kamide et. al. (2003).

Experiment 1, however, leaves one critical question unanswered: can children use case to predict arguments independent of the verb? In children, we found no reliable effects of case until the region that followed the verb. This could indicate that young children, unlike adults, need to use the verb and the specific event roles that it encodes, to interpret case markers. If this is the case, the effects we observed might have arisen just on the basis of an activation of event schema that relied on lexico-semantic associations between the first noun and the verb (Bar, 2009). In other words, thematic fit between the fist noun and the meaning of the verb might be activating children's previous experiences of the event type denoted in the sentence leading to the predictions of the event structure (Feretti et al., 2001; MacRea et al., 2005). This same problem also applies to previous studies (Borovsky, Elman & Fernald, 2012; Mani & Huetting, 2012; Kamide et al., 2003). In Borovsky et al. (2012) study, for instance, young children viewed a visual scene with four referents (e.g., a cat, a bone, a box of treasure, and a ship) accompanied by a spoken utterance and they were able to use the semantic properties of the verb (e.g., *hide* vs. *chase*) in line with the semantic properties of the first noun (e.g., *the dog* vs. *the pirate*) to predict the second argument. More specifically, children had more looks to the plausible referent that could be hid by the dog (i.e., *bone*) when they heard a sentence like *The dog will hide* and they had more looks to the plausible referent that could be chased by the dog (i.e., *cat*) when they heard a sentence like *The dog will chase*. Here too, the semantic relatedness between *dog* and *hide* (*dog* and *chase*) might have automatically led predictive looks to the correct upcoming referent. If this were the case, this could be achieved without any use of morphosyntactic information.

Alternatively, young children, like adults, could have the ability to use syntactic and morphosyntactic information to assign broad thematic roles independent of the verb and possible semantic associations between lexical items, but that ability could be more difficult to detect for a variety of reasons. For example, they might be slower to process information resulting in a delayed pattern of effects. Perhaps they would be able to use case prior to the verb, if we increased the length of time between these two morphemes. Alternatively, children's processing could be as rapid as adult processing but noisier, and as a result small effects (like those in the adverb window) might disappear in the noise, while larger ones remain detectable. If shifts in eye-gaze are driven by an attempt to synchronize incoming perceptual information, then we should expect that looks to the upcoming argument would increase when participants hear that a new noun phrase is beginning. This would lead to predictive effects being largest in the modifier region, regardless of where the verb is.

In Experiment 2, we explore this possibility by conducting a parallel experiment using verb-final sentences. In these sentences, we can assess thematic prediction prior to the verb by measuring gaze in the modifier window. If children are simply noisier than adults, we should see predictive looks in the modifier region even though the verb has not occurred. If children require verb information to make predictions, then the effect in the modifier window should disappear.

Experiment 2

In Experiment 2, we made two changes to the stimuli. First, we moved the verb to the end of the utterance. As a result, the predictive region consisted of the adverb and the modifier. Second, the utterances had a natural prosodic break after the adverb. This resulted in a predictive region that was roughly equivalent in length to the predictive region in the first experiment.

Method

Participants

Forty children between the ages of 4;0-5;0 (M = 4;54) and 21 university undergraduates, all with Turkish as their first language, participated in this study. Participants were reported to have vision that was normal or was corrected to normal through the use of contact lenses. The parents reported that the children had no known neurological or behavior problems and the adults reported that they did not either. None of the participants had taken part in Experiment 1.

Stimuli

The sentences in experiment 2 were the same as those in experiment 1 except that the words appeared in a different order. Specifically, the verb was placed at the end of the utterance, resulting in SOV and OSV orders where the case marking on the first noun was manipulated between the nominative and the accusative. Thus, we had sentences in two conditions, namely the nominative (i.e., N_{Nom}-N_{Acc}-V) condition and the accusative (i.e, N_{Acc}-N_{Nom}-V) condition, as in (9) and (10).

(9) Nominative condition (SOV: N_{Nom}-N_{Acc}-V*):*

Minik tavşan-∅	birazdan	şurada-ki	havuc-u	yiy-ecek.
little rabbit-Nom	soon	there-Rel	carrot-Acc	eat-Fut

"The little rabbit will soon eat the carrot over there."

(10)   Accusative condition (OSV: N_{Acc}-N_{Nom}-V*):*

Minik tavşan-ı	birazdan	şurada-ki	tilki-∅	yiy-ecek.
little rabbit-Acc	soon	there-Rel	fox-Nom	eat-Fut

“The fox over there will soon eat the little rabbit.”

The sentences were accompanied by the same pictures and animations as in Experiment 1. The same female native speaker recorded the sentences using a natural prosody with a focus accent on the verb and no topicalizing or contrastive focus on the first noun. We edited the sound files to control for the duration of the pauses. Each sound file was preceded by 200-ms silence, there was a pause of 300 ms after the first noun and a pause of 1500 ms at the end of the utterance before the animation started.

Procedure

The procedure was the same as in Experiment 1.

Data analysis

The same data analysis procedure was followed as in Experiment 1.

Results and Discussion

Table 5 presents the percentage of correct responses to the end-sentence animations in each condition by children and adults. We see above 93% success rate in the offline evaluation of the animations in relation to the utterance heard for both conditions.

Table 5: Percentage (%) of correct responses in each condition for children and adults

	Adults	Children
Nominative	100	93.41
Accusative	97.72	91.58

Figures 7 and 8 show the eye-gaze patterns of how agent-preference changed in each time window through the course of the utterance for both children and adults in Experiment 2. What we see in these figures is that both adults and children begin with greater agent preference in the nominative condition compared to the accusative condition in the first time window, showing no difference with respect to case in the second time window during the adverbial region. Towards the end of this time window, children begin to show a shift with greater agent preference in the accusative condition compared to the nominative condition, which becomes clearer during the modifier region (before the second noun and before the verb). Adults show the same pattern in the modifier region.

Figure 7: Gaze patterns of agent preference in each time window for adults in Experiment 2

Error bars indicate the standard error of the mean. Values above zero indicate preference to look at potential agent, and values below zero indicate a preference to look at the potential patient.

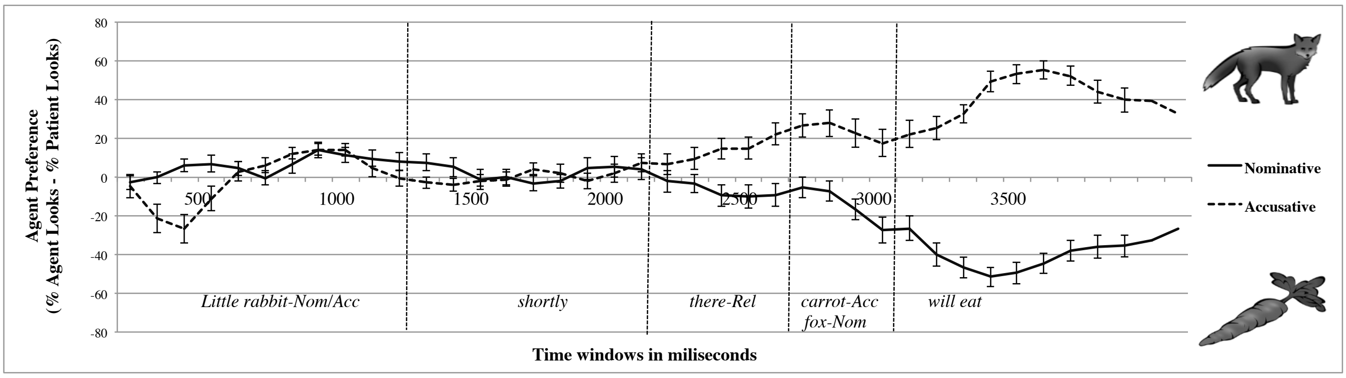
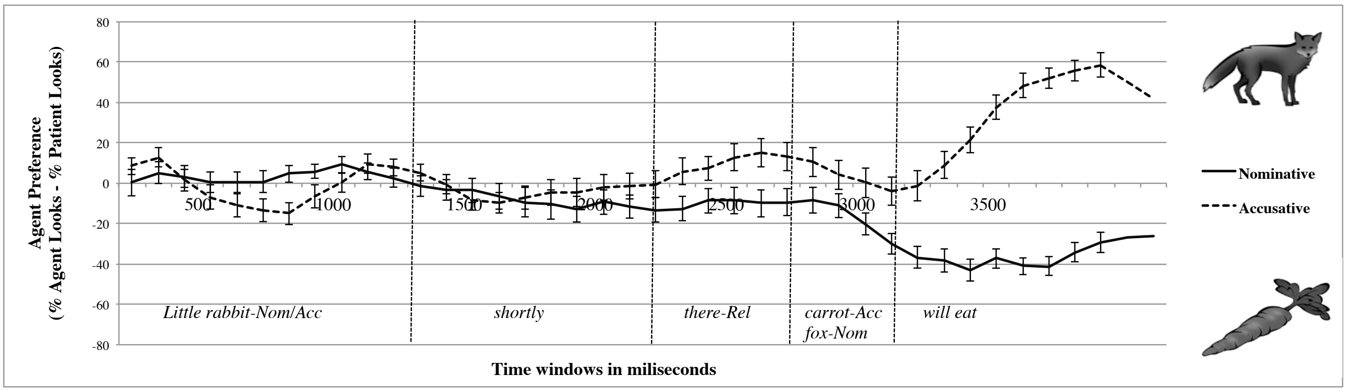


Figure 8: Gaze patterns of agent preference in each time window for children in Experiment 2

Error bars indicate the standard error of the mean. Values above zero indicate preference to look at potential agent, and values below zero indicate a preference to look at the potential patient.



As in Experiment 1, we calculated agent preference and looked at the eye-movements in time windows corresponding to the five critical words. Again, these time windows were not offset by 200ms and we grouped the pauses with the words that they followed (the first noun and the adverb). The same data analysis strategy was employed: we used mixed-effects logistic regressions, with fixed effects for case (first noun accusative or nominative) and age (child or adult), and the interaction of these variables.[8] We included random intercepts for both participants and items in all analyses and conducted parallel analyses with other

[8] As in Experiment 1, we also conducted alternative analyses with ANOVA for the critical time window as we did in Experiment 2. The findings were again compatible with our analysis (Appendix 5).

potential random effects. When these analyses converged, they did not provide a significantly better fit to the data ($p > .05$) indicating that the additional random predictors were not justified.

Figure 9: Agent Preference in each time window for the adults in Experiment 2

Error bars indicate the standard error of the mean. Values above .5 indicate preference to look at potential agent, and values below .5 indicate a preference to look at the potential patient.

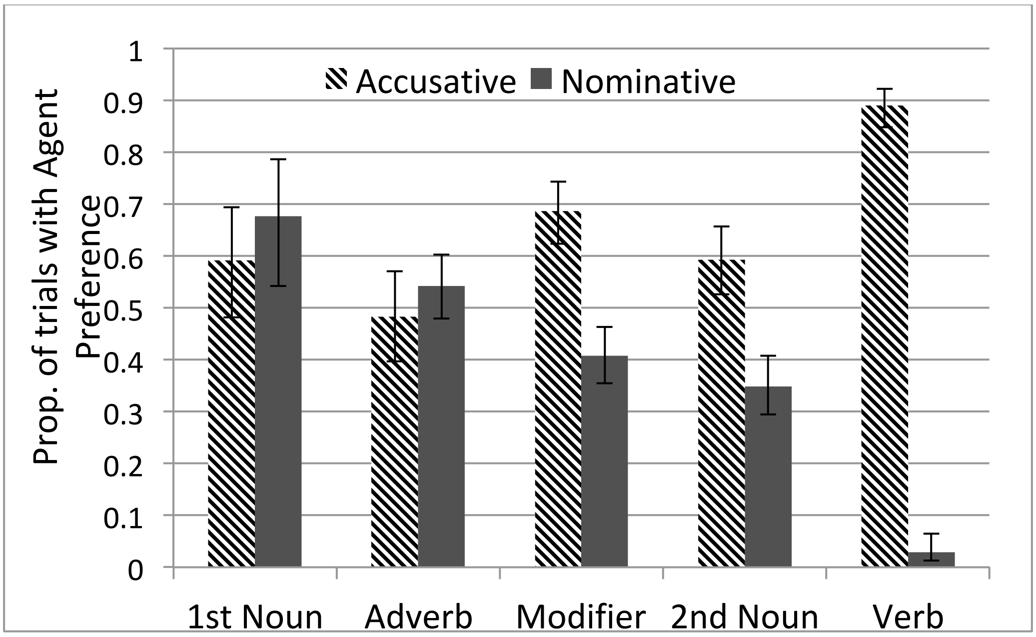
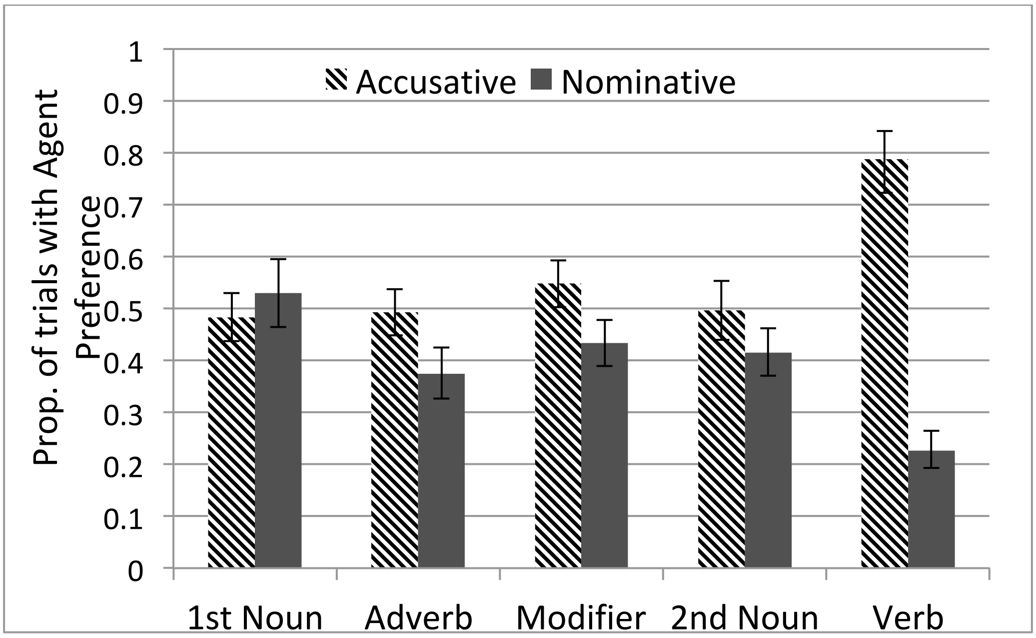


Figure 10: Agent Preference in each time window for the children in Experiment 2

Error bars indicate the standard error of the mean. Values above .5 indicate preference to look at potential agent, and values below .5 indicate a preference to look at the potential patient.



Figures 9 and 10 show the proportion of trials with an agent preference in each group across the five time windows. We conducted an omnibus analysis of all five time windows, with fixed effects for case, age group and time window to see whether these scores changed over the course of the trial, and how case influenced these changes. We added the predictors and their interactions to the model in the order given, and the contribution of each one was assessed with a log-likelihood test (see "Full Sentence" in Table 6).

Table 6: Omnibus analyses of agent preference in Experiment 2

Time Window	Full Sentence	First Noun and Predictive Windows

Case	$\chi^2(1) = 91.02$ p < .0001**	$\chi^2(1) = 5.18$ p = .02*
Group	$\chi^2(1) = 3.31$ p = .07†	$\chi^2(1) = 5.15$ p = .02*
Case * Group	$\chi^2(1) = 3.82$ p = .05†	$\chi^2(1) < .01$ p = .94
Time Window	$\chi^2(4) = 6.42$ p = .17	$\chi^2(2) = 4.10$ p = .13
Case * Time Window	$\chi^2(4) = 137.75$ p < .0001**	$\chi^2(2) = 10.68$ p = .005**
Group * Time Window	$\chi^2(4) = 1.90$ p = .75	$\chi^2(2) = .32$ p = .85
Case * Group * Time Window	$\chi^2(4) = 12.45$ p = .01*	$\chi^2(2) = 2.49$ p = .29

The shaded cells are the effects that would indicate the predictive use of case.

† .05 < p < .1   * .01 < p < .05   ** p < .01

This analysis revealed the following patterns. First, there is an effect of case condition in the expected direction (more looks at the potential agent in the accusative condition compared to the nominative condition). Second, there is a marginal effect of age: children looked more at the potential patient than adults. Third, there is an interaction between case and the time window, confirming that the difference between the two sentence types emerges as the utterances unfold. Finally, there is an interaction between case, age and time window reflecting a greater differentiation of the two constructions as the sentence progresses in the adults (relative to the children). Critically, this analysis is based on the entire sentence, and thus the pattern of effects reflects the identity of the second noun, in addition to any predictive processes.

To focus on the predictive effects, we conducted another omnibus analysis excluding the final noun and verb (see "First Noun and Predictive Windows", Table 6). In this analysis, there was no hint of the three-way interaction between age, case and time window, which had been present in the analysis of the full sentence. The marginal age by case interaction also disappeared. Thus there are no clear differences between children and adults in the degree to which they use case to predict upcoming arguments, suggesting that the interactions in the Full Sentence analysis are driven by differences that emerge after the second noun begins.

Three effects did persist in the analysis of the predictive regions. First, there was a reliable effect of age: prior to the second noun, children looked at the potential patient more often than adults, consistent with a bias for interpreting the first noun as the agent. Second, there was a robust effect of case, in the expected direction, indicating that participants were predicting the upcoming noun. Third, there was a significant case by time window interaction indicating that this effect changed as the sentence unfolded. To explore this further, we conducted separate analyses for each time window with case and age group as fixed effects (Table 7). We then analyzed each age group separately (Table 8).

Table 7: Time window analyses for agent preference in Experiment 2

Time Window	Effect		
	Case	Group	Case * Group
First Noun + Pause	β = -.09 z = -.96 p =.34	β = -.18 z = -1.83 p =.07†	β < .01 z = .03 p =.98
Adverb + Pause	β = .07 z = .67 p =.51	β = -.16 z = 1.63 p =.10	β = .16 z = 1.67 p =.10†
Modifier	β = .41 z = 4.02 p < .0001**	β = -.11 z = -.97 p =.33	β = -.18 z = -1.74 p =.08†
Second Noun	β = .32 z = 3.28 p = .001**	β = -.02 z = -.20 p =.84	β = -.15 z = -1.60 p =.11
Verb	β = 1.56 z = 12.50 p < .0001**	β = -.05 z = -.36 p =.72	β = -.36 z = -2.86 p =.004*

The shaded cells are effects that would indicate the predictive use of case.

† .05 < p < .1   * .01 < p < .05   ** p < .01

Table 8: Analyses for the children and for the adults in Experiment 2

Time	Effect

Window	Child Intercept	Child Case	Adult Intercept	Adult Case
First Noun + Pause	I = .02 z = .11 p =.91	β = -.09 z = -.76 p =.45	I = .47 z = 1.14 p =.26	β = -.12 z = -.74 p =.46
Adverb + Pause	I = -.25 z = -1.98 p =.05*	β = .23 z = 2.05 p =.04*	I = .07 z = .44 p =.66	β = -.10 z = -.61 p =.54
Modifier	I = -.03 z = -.19 p =.85	β = .24 z = 2.14 p =.03*	I = .19 z = 1.06 p =.29	β = .57 z = 3.42 p =.0006**
Second Noun	I = -.16 z = -1.47 p =.14	β = .16 z = 1.47 p =.14	I = -.12 z = -.70 p =.49	β = .49 z = 3.03 p = .002**
Verb	I = -.09 z = -.57 p =.57	β = 1.17 z = 9.47 p < .0001**	I < .01 z = .02 p =.99	β = 2.05 z = 8.93 p < .0001**

The dependent variable is agent preference.

The shaded cells are effects that would indicate the predictive use of case.

† .05 < p < .1   * .01 < p < .05   ** p < .01

At the first noun, there were no effects of case in any of these analyses, just a marginal effect of group (with children looking more to the patient than adults, consistent with a greater agent-first bias). The pattern at the adverb is ambiguous: in the analysis of both groups there is no effect of case, and only a marginal interaction between case and group. However, the children show an effect of case, in the predicted direction, which reaches significance, while the adults do not.

Our critical predictions, however, were about the modifier region. If children, and adults, can use case predictively *prior to encountering the verb*, then we should expect a robust effect of case in the combined analysis (Table 7) and reliable effects of case for both adults and children (Table 8). This is precisely what we find. In addition, there is also marginal interaction between case and age group (Table 7), reflecting the fact that the effect may be larger in adults than children (β = .57 and .24, respectively)

After the second noun begins, and the referent is disambiguated, the difference between the two sentences persists. This results in a case effect in the pooled analysis and in the adults, but oddly not in the children. The absence of the effect in children could indicate that they are slower in processing the noun or that they are more likely to have shifted away from the inferred referent (since the preference for the children begins back at the adverb). However, in the absence of a robust interaction between group and case, there is no strong reason to posit a difference between the age groups. In the verb region, at the end of the sentence, the difference between the two conditions is robust in both groups (β = 1.17 and 2.05 for children and adults, respectively). In this time window, there is also a robust interaction between age group and case condition, indicating that children are less categorical in their tendency to look at the correct referent.

In sum, the results clearly demonstrate that both children and adults are able to use the case marking on the first noun to predict the second noun prior to the verb. Our studies are the first to clearly demonstrate that children and adults create thematic expectations about the upcoming referents solely on the basis of the identity of the first noun and its case marker, combined with the real-world plausibility information provided by the visual context – without any further information provided by the identity of the verb or of a second referent (c.f., Borovsky, et al. 2012; Mani & Huetting, 2012; Kamide, Scheepers, & Altmann, 2003; Kamide, Altman, & Haywood, 2003). The present pattern requires that our participants assign a thematic role to the first noun, without knowing the verb that will follow, and then use this information to infer the likely thematic role of the second noun. In addition, these findings, like those of Experiment 1, indicate that children have a bias to initially interpret the first noun as the agent, resulting in more looks to the possible patient. This is evident in the group effects in the omnibus analyses (Table 6) indicating that children look at the patient more than adults) and in the reliable negative intercept for the children in the adverb region (Table 8) indicating that children have a preference for the patient over the agent.

General Discussion

Our findings show that children acquiring Turkish can use case markers to incrementally assign thematic roles to nouns and predict upcoming arguments in sentences, even when they have yet to encounter the verb. In Experiment 1, we tested four-year-olds and adults on verb-medial structures and found that both groups used the contrast between the nominative and the accusative case on the first argument to anticipate the second noun. In adults, these predictive looks emerged before the verb, but in children they only appeared after the verb (at the modifier), raising the possibility that children might need verb information to make thematic predictions. Alternatively, children might be able to make predictions independent of the verb, but they may be slower in parsing incoming information at multiple layers (e.g., integrating the available noun and the adverbial in the syntactic hierarchy, processing the morphosyntax on the available lexical items, integrating semantic and discourse information with the visual display which would then lend itself to structural and semantic predictions about the upcoming utterance) or their predictions may become stronger and more stable as they get closer to the expected position of the predicted argument, resulting in larger effects immediately before the noun. In Experiment 2, we disentangled these possibilities by using verb-final sentences and found that four-year-olds (like adults) were able to use the case distinction to predict the upcoming argument prior to the verb (with reliable effects in the adverb and modifier regions).

To the best of our knowledge, this is the first clear evidence that young children can use case markers for thematic role assignment independent of verb meaning. This is a central skill for many of the world's children. 42% of languages have a preference for verb-final word order (Dryer, 2013a). Most of these languages have productive case marking on nouns (79%, Dryer, 2013b) and fairly flexible word order (Steele, 1978; Newmeyer, 2001). Another 14% of languages are reported to have no dominant word order, but again most of these languages (72%) have productive case marking (Dryer, 2013ab). Children learning these kinds of languages would be ill-served by a processing strategy that relied solely on verb-specific information or word order.

In the remainder of the discussion, we will address: 1) how these findings fit into our current understanding of language comprehension; 2) the role of psycholinguistics in understanding typological variation; 3) how these findings constrain theories of language acquisition; 4) how to reconcile our findings with the prior work on case comprehension in German speaking children.

More evidence that language comprehension is flexible and opportunistic

The interpretation of our findings becomes clearer when we place them in the context of contemporary theories of moment-to-moment language comprehension. While these proposals vary in many respects, several decades of intensive experimental research has resulted in convergence on several critical points (for reviews see, Elman, Hare & McRae 2005; Treiman, Clifton, Meyer et al. 2003; van Gompel, 2013; Kuperberg & Jaeger, 2016). First, understanding language involves constructing representations of an utterance at multiple levels, which are partially ordered with respect to the speech stream (from sounds to words to ideas). Second, construction of these levels is incremental: processing at a higher level does not wait until processing at lower levels is finished, instead information quickly propagates upward. For example, we begin activating the meanings of words as the first phoneme comes in. As a result, local ambiguity is rampant: the sounds we hear activate multiple words, a given word form can have many different meanings, and a string of words is often compatible with several syntactic structures or semantic analyses. Third, at each level, processing is interactive: we resolve this ambiguity by drawing on information from other levels of representation, some higher in the processing stream and some lower, to determine which analyses to pursue. This is a powerful tool because language is characterized by systematic correspondences across levels (words link sound to meaning, syntactic structures link linear strings to semantic propositions). Interactive processing allows us to exploit those correspondences, using knowledge at each level to reduce uncertainty at the others. Fourth, as a consequence of incremental, interactive processing we engage in prediction of upcoming information at multiple levels. When we activate higher level structures on the basis of incomplete information at lower levels, this leads to the expectation that the missing elements in those structures will soon appear. This information propagates downward, resulting in prediction at lower levels.

Over the past fifteen years, there has been a surge of research exploring how this system develops, prompted in part by the development of robust methods for studying language processing in young children. This work has demonstrated that the core properties of the adult comprehension system are present early in development (for review see Snedeker & Huang, 2015). Children also process language in incremental and interactive fashion, activating the meanings of candidate words as they unfold, constructing syntactic analyses of partial clauses, and making predictions about upcoming constituents.

The present results fit firmly into this model of language comprehension. First, as they hear the initial noun, listeners identify its case and activate the thematic role(s) associated with that case. This was demonstrated most clearly in the Experiment 1: children and adults who heard rabbit in nominative case were more likely to look at the other potential agent (the fox) than those who heard it in accusative case. Presumably these looks are not predictions about the second noun--they would be flat out wrong and surprisingly early. Instead they appear to be part of the initial processing of the first noun phrase; case information is being accessed during the time when the root noun is still being processed and both pieces of information are being used in parallel to disambiguate reference. In the nominative condition this leads participants to look to items that are plausible agents, in the accusative it leads them to look at plausible patients.

Second, by assigning a thematic role to the first noun, listeners are constructing one or more hypotheses about the event structure of this sentence. For example, in assigning the rabbit to the role of patient (narrowly defined) we are positing that it is the affected entity in a change of state event. If we assign it to the role of patient, we are considering the possibility that it was moved. These bits of event structure generate expectations about upcoming arguments (there should be a cause of the change of state or motion), which lead to predictions about the plausible referent of this upcoming argument, generating the eye movements that we observed right before the second noun. Thus our child, and adult, participants have developed a processing strategy that is well suited to the specific predictive cues provided by their language. But this language specific strategy makes use of the same basic building blocks (incremental, interactive interpretation resulting in prediction) that characterize comprehension across different populations, different tasks, and different levels of representation. This model of language comprehension raises additional questions about typological variation and language acquisition, which are central to the present paper and are discussed in the sections below.

Comprehension as a constraint on typology.

Languages vary along dimensions, such as word order and argument marking, that logically should affect how and when predictions are made. Do different languages provide equivalent levels of predictive constraints or are some languages more constraining and thus simpler to process? Many suggest that languages have been shaped by selective pressures operating over historical time (see e.g., Deutscher, 2005; Christiansen & Chater, 2008; Gibson, Piantadosi, Brink, Bergen, Lim & Saxe, 2013; Tomasello, 1999/2011). The challenge of accurate online language comprehension could be a particularly potent selector, shaping how languages package syntactic information and distribute it across an utterance. Specifically, the inability of humans to accurately remember long lists of unintegrated items could select for languages that allow conceptual integration to begin early in the sentence. In some languages, this need could be filled by predictable word order and the early arrival of the verb (Pozzan & Trueswell, 2015). For example, English speakers may provisionally assume that the first noun is an agent (see e.g., Bates, MacWhinney, Caselli, Devescovi, Natale & Venza, 1984; Bever, 1970; Meyer, Mack & Thompson, 2012) and then use the verb to retrieve an event template that allows them to predict (and interpret) the arguments that follow (see e.g., Altmann & Kamide, 1999; Arai & Keller, 2013; Trueswell & Kim, 1998). In other languages, case markers may satisfy this need. For example, Turkish speakers may use accusative case on the first noun to begin building an event template that allows that noun to be integrated (as a patient), posits the existence of an agent, and generates the prediction that a noun phrase filling this role may appear. Our results are clearly compatible with this hypothesis, as are prior findings demonstrating the predictive use of case in Japanese-speaking and German-speaking adults (Kamide, Altmann & Haywood, 2003; Kamide, Scheepers & Altmann, 2003).

Table 9: The relation of word order and case marking across languages

	Verb Initial	Verb Medial	Verb Final	No Dominant Order
Languages With Case	35	89	326	103
Languages Without	47	187	87	40
Total	82	276	413	143
Percentage Case Marked	43%	32%	79%	72%

According to the data from Dryer (2013a; 2013b), languages without case are only those classified as having neither case affixes or adpositional clitics. All others are categorized as languages with case. The verb initial languages are mostly VSO (78%), and the verb medial and verb final languages are almost entirely SVO and SOV respectively (97%, 99.5%).

If online comprehension constrains typological variation in this way, then we would expect case marking to be more common in languages where the verb follows its complements. As Table 9 illustrates this is precisely the pattern that we see. Verb-final languages are more likely to have productive case markers and more likely to have rich case marking. But there is considerably more work to be done before we will know what role comprehension plays in constraining typological variation. The sentence processing studies merely demonstrate that speakers can use information from case marking and information from the verb to make predictions. We do not yet know whether these two information sources are equivalent in strength or the degree to which they are in complementary distribution in the actual utterances that children hear. Nor do we know how the pressures of online comprehension shape reiterated language learning (but see Fedzechkina, Jaeger, & Newport (2012) for evidence that the ambiguity created by word order variation affects the retention of case marking in artificial language learning task).

Implications for theories of language acquisition

Above we sketched a picture of language comprehension in which listeners make use of the strongest cues in their language to make predictions about the rest of the utterance. On this model, one critical part of language acquisition is to discover the cues that are relevant in your language and gain the skill to employ them rapidly and incrementally. Prior studies have found that young English speakers make rapid use of word order and verb information (Nation, et al., 2003; Borovsky, et al., 2012; Trueswell, et al., 1999; Snedeker & Trueswell, 2004). But this data pattern, by itself, leaves open two possibilities: 1) young children may rapidly acquire language specific parsing strategies, or 2) young children may rely on universal strategies (use word order and verb meanings) that are better suited for some languages than others. Our findings provide clear evidence that, by four years of age at least, the first interpretation is the correct one. Young Turkish speakers do not simply interpret the first noun as the agent, instead they use case to flexibly assign thematic roles and make predictions about upcoming arguments. As we noted in the introduction, these findings are relevant to two central issues in language development.

Does word order have a privileged status in language acquisition?

Order is a necessary property of spoken language because multiple morphemes cannot be produced simultaneously. In narrative, languages typically use the order of sentences to mark the order of events, suggesting that there might be a predisposition to assume a temporal isomorphism between language and the world (Deutscher, 2005). A bias toward temporal isomorphism should favor an agent-first word order. Agents initiate the action and exist prior to and independent of the event (Dowty, 1991). Patients, in contrast, typically move or change as a result of the agent's actions and sometimes only come into existence as a part of the event (*bake cake* or *tell a lie*). If infants expect temporal isomorphism, and this expectation is strong enough, then they should find it easier to learn word order cues than case marking cues. Our results clearly demonstrate that by 4 years of age children have a robust mastery of case marking. Obviously these results cannot tell us when this ability is acquired or whether younger children rely more heavily on word order. However our findings do show that by the time children are in preschool their knowledge of case is sufficiently robust that they use it spontaneously and rapidly to make predictions as a sentence unfolds.

There is one feature of our data which suggests that children might be more reliant on order than adults. In Experiment 1, children were more likely to look at the potential patient than adults, particularly after the onset of the verb. This bias could reflect a greater reliance on word order (a residual tendency to assume that the first argument is an agent). But it could also reflect differences in perceptual or conceptual biases (maybe kids are just more interested in the objects than the adults are), or the children's assumption that some of these events (e.g., eating or biting and finding) are likely to have only one animate argument.

Critically, this difference between the groups was not present in Experiment 2 in which the crucial regions came before the verb. Interestingly, we saw no evidence in our data that the adults had a general preference for agent or subject-initial sentences, though this pattern has been found in a variety of languages and tasks (e.g., Bever, 1970; Frazier, 1987; de Vincenzi, 1991; Schiefers, Friederici, & Kühn, 1995; Schlesewsky, Fanselow, Kliegl, & Krems, 2000; Frisch & Schlesewsky, 2005; Demiral, Schlesewsky, & Sclesewsky, 2008).

Another interesting issue is the fact that we observed more looks to the plausible patient in the nominative condition despite the fact that the visual scene in our study made possible both types of event construal where the first noun is the patient (*The rabbit will shortly be eaten by the fox.*) or the experiencer (i.e., *The rabbit will shortly flee from the fox*). This was also the case in previous studies (e.g., Kamide, Scheepers, & Altmann, 2003; Altmann & Kamide, 1999). This pattern might be due to a probabilistic tendency ranking an agent higher than an experiencer (for a related discussion, see Kamide, Altmann, \& Haywood, 2003). There might also be a cognitive tendency to construct the perspective of the sentence from the perspective of the most accessible entity (i.e., agent) (MacWhinney, 1977; Bock & Warren, 1985), and when the case marker on the first noun corresponded to this expectation, the participants might have constructed the event from the perspective of the agent acting on another entity. Dowty (1991) similarly suggests there might be a cognitive tendency to take a given argument that morphosyntactically fits in the subject role as bearing the strongest agent entailments (e.g., intention, causation, change of state) before entertaining a less strong thematic possibility (i.e., experiencer, patient). Dowty links this to human survival, suggesting that an expectation that the agent-like entity will act in the most agent-like manner might be life saving (Dowty, 1991; p.601).

Do children have grammars that are less abstract than adults?

For the last 25 years, there has been an ongoing debate about whether the granularity of grammatical generalizations changes across development. Late abstraction accounts propose that children initially begin with narrow generalizations based on the distributional patterns and meanings of individual words and then gradually form broader, more abstract constructions by generalizing across these exemplars (see e.g., Ambridge & Lieven, 2011; Goldberg, 2006; Tomasello, 1991). On this theory, children's initial mappings between structural positions and argument roles are specific to individual verbs. Broad semantic roles (e.g., patient or agent) and syntactic relations (e.g., subject and object) are a gradual developmental achievement. In contrast, generativists propose that languages have abstract semantic and syntactic relations precisely because children are predisposed to analyze their world and their language with categories of this breadth (Fisher, Hall, Rakowitz & Gleitman, 1994; Pinker, 1984). If the late abstraction accounts are right, and children's initial grammars are based on verb-based generalizations (lexical islands) then they should be unable to assign a role to a noun phrase prior to hearing the verb. How long this inability should last is unclear--presumably the pace of generalization would depend on what the initial semantic hypothesis space is, how the categories of agent and patient are constructed from this space and what data is available to the child to guide learning. These questions are, to date, largely unaddressed. However, late abstraction accounts have claimed that children as old as four (Savage, Lieven, Theakson & Tomasello, 2003) or five (Boyd & Goldberg, 2012) may rely more on lexically based generalizations or word order. Thus, on the late abstraction account, we might expect that preschoolers would struggle to use case during language comprehension. In fact, Dittmar and colleagues (2008) propose that gradual generalization of constructions may account for difficulty that German speaking five-year-olds have in using case to interpret sentences with novel verbs.

Our second experiment provides clear evidence that, by 4 years of age, children can interpret case markers independent of the verb, independent of the construction and independent of the word order configuration; and in fact they do so spontaneously and rapidly. Note that nothing in our experimental task requires that children generalize (or even predict). Young children spontaneously access representations that are broad enough to allow them to integrate nouns into an event structure without precise knowledge of the kind of event under discussion. Given the age group we are testing, we remain agnostic about whether thematic roles and grammatical abstractions are innate. However, the pattern we found in Turkish speaking children is clearly inconsistent with the picture painted by late abstraction accounts suggesting that the *abstract constructions* stage – which begins as of age 3 and which is far from reflecting full thematic and grammatical abstraction – lasts until age 5 or 6 when children develop the necessary processing strategies to reach broader abstractions (preemption, generalizations on the basis of semantic verb classes, and analogy-making) (Tomasello, 2003; p.169). As we highlighted before, early abstraction is not only expected by theories assuming innate grammatical representations but it is also expected on any theory assuming a bias in human beings to represent events in terms of their broad semantic roles and a strong statistical learning skills detecting patterns across utterances. The present pattern is fully consistent with such a view and a growing body of work suggesting that grammatical abstraction is early (affecting novel verb comprehension and production by age 2: Fisher, 2002; Gertner, Fisher, & Eisengart, 2006; Yuan & Fisher, 2008; Kline & Demuth, 2014; Naigles, 1990; Arunachalam & Waxman, 2010; Waxman & Dana, 1995) and pervasive (shaping the comprehension and production of sentences with known verbs by age 3: Thothathiri & Snedeker, 2008; Rowland et al., 2012).

Why do children fail to use case in some studies but not others?

In the introduction, we noted that prior studies have found that German speaking children have difficulty understanding OVS sentences with case marked initial nouns until the age of 6 or 7: young children perform poorly when verb-specific information is not available (Dittmar et al., 2008), fail to detect case marking violations in ERP paradigms (Schipke et al., 2011), and do not show the increase in the activation in the left inferior frontal gyrus that is associated with comprehension of these structures in more competent children and adults (Knoll et al., 2012). In contrast, we find that four-year-old Turkish speakers readily predict upcoming arguments on the basis of case markers, making more predictive looks to the agent when the first noun is accusative than when it is nominative. Our findings are inconsistent with the strongest interpretation of the German data. If the comprehension of case marking and non-canonical word order depends on a late maturing neural connection (Brauer, et al., 2011; Friederici, 2012), then we should expect that these abilities will be absent in children under 6 regardless of which language they are learning. Nevertheless, the German data provide clear evidence that under some circumstances children do struggle to interpret case marking. Thus the challenge is to understand both the successes and the failures. We see two broad possibilities for the divergent findings: the difference could be in the experiments and what they measure, or it could be in the languages (or both).

Experimental Differences: Our experiment had several features that might make it more sensitive to children's understanding of case. First, in contrast with Dittmar (but like Knoll and Schipke) our task did not require children to act out the sentence or choose between two very similar pictures. They simply had to listen to the sentences. Second, while our experiment was designed to measure fleeting thematic predictions, the German studies have focused on offline interpretation and neural indices of processing difficulty. Perhaps children use case on the fly, to make predictions, but have difficulty holding on to those predictions long enough to guide explicit judgments or resolve processing difficulties.

Third, and perhaps most critically, all of our transitive sentences described highly plausible events and had two arguments that differed systematically in the degree to which they would make a good agent or patient. In other words, our sentences were much like the transitive utterances that children typically hear. In contrast, the German studies used reversible transitives: sentences that would be equally plausible (or implausible) under either role assignment (e.g., the bunny hits the cat). This is a necessary feature of offline studies (if the sentences are not reversible, children can solve the problem without using case). But it is unnecessary in studies of online processing where we can look at the prediction or real time integration of semantic information. The use of atypical reversible transitives could conceivably create cumulative doubt across the study as the child repeatedly arrives at analyses that do not seem particularly plausible. This problem is compounded in studies that use violation paradigms (e.g., Schipke et al., 2011) because half of the critical sentences will contain case marking errors. A rational learner should eventually stop making predictions under these circumstances. If we assume that younger children have weaker priors, then we might expect them to give up their predictive strategies more quickly in the face of such unsatisfying outcomes.

Crosslinguistic Differences: There are two properties of the Turkish case marking system that might facilitate acquisition and lead children to lean heavily on case marking for thematic interpretation. First, in Turkish case is marked with syllabic nominal suffixes which vary a bit depending on the phonological context, but do not carry additional grammatical information (e.g., number or gender). Nominative case is the unmarked case, and accusative is consistently marked for definite nouns. In contrast, German case is marked on both the determiner and the noun, but in a manner that depends on a complex interaction between gender and number. Critically, the distinction between accusative and nominative case is only visible for masculine singular nouns. Thus German children must unpack case from these other cues during acquisition and will often encounter sentences where the case of the first noun cannot be determined.

Second, Turkish children frequently encounter object initial sentences, and thus they would be ill served by a strong bias to always assume that the first noun is an agent. Although Turkish is considered to be an SOV language, the word order is highly variable and argument dropping is common (Kornfilt, 1997).

In contrast, in German, as in English, subject dropping is very rare. While OVS sentences are grammatical, they are dispreferred in most contexts and their acceptability depends on information structure, pronominalization, animacy, definiteness, verbal semantics, and prosody (Gorrell, 2000; Weskott, Hornig, Fanselow & Kliegl, 2011; Primus, 1993). In both text and adult-directed speech, over 90% of sentence initial noun phrases are nominative subjects (Gorrell, 2000). In a corpus study analyzing the relative ordering of subject and object, Bader and Haussler (2009) concludes that 'within the middlefield, an object occurs in front of the subject mainly in order to adhere to lexical-semantic constraints. In the rare case [6%] that an object precedes a subject for discourse-related reasons, this is due to the constraint requiring a focused constituent in preverbal position. For putting the objects in the prefield, and thereby in front of the subject [it happens 38% of the cases], lexical-semantic constraints and discourse-related constraints play both a significant role. The main discourse-related constraint is the constraint requiring topics to occur in clause-initial position.' (p.757). This pattern extends to child-directed speech: Ditmar and colleagues (2008) found that the subject preceded the object in 68% of the transitive sentences of the Szagun (2004) corpus. Presumably, this number would be much higher if nominative intransitive subjects were included as well. In short, children acquiring German may have good reasons to get into the habit of assigning the agent role to the first noun that they encounter: for lexical noun phrases case will often be ambiguous and most of the time this initial hypothesis will be correct. In naturalistic contexts, the cases where the rule would fail may be clearly marked by prosody or discourse cues (see Dittmar et al., 2008 for discussion).

If German-speaking children have an initial parsing strategy based on word order, then case marking may mostly be used as a secondary cue for revising incorrect role assignments. On this construal, the Dittmar and Friederici data would suggest that German speaking children have difficulty with revision until approximately 7 years of age, that revision is indexed by activity in left inferior frontal gyrus, and that development of revision strategies may be subserved by a late maturing connection between temporal and frontal language regions. This reconceptualization is strongly consistent with proposals by

Trueswell and colleagues about the development of language comprehension. They argue that young children quickly assign thematic roles on the basis of cues that are early and statistically robust (e.g., Trueswell, et al. 1999; Snedeker & Trueswell, 2004; Choi & Trueswell, 2010). When subsequent information appears that contradicts this initial analysis, 5-year-olds fail to revise their interpretation (Trueswell, et al. 1999; Choi & Trueswell, 2010). By 8 years of age, children are able to override these biases (Weighall, 2008). Trueswell and his colleagues have proposed that children's errors reflect immature executive function. Specifically they argue that revision requires the use of cognitive control which is subserved in part by the left inferior frontal gyrus (January, Trueswell & Thompson-Schill, 2009; Novick et al., 2005; Novick, Trueswell, & Thompson-Schill, 2010). Thus while Friederici's and Trueswell's explanations are very different, their data patterns and the neural changes that they propose are quite similar.

Final remarks

In this paper, we found the Turkish-speaking preschoolers and adults are able to rapidly use case markers to predict upcoming arguments, even when the verb has yet to appear. These results demonstrate that the difficulty that German-speaking children have with non-cannonical word orders cannot be attributed to a universal developmental limitation driven by the late maturation of core syntactic processes. Thankfully, children learning head final languages are able to develop parsing strategies that make use of the syntactic cues that appear early in their utterances. This capacity to use case taps into broad grammatical abstractions that allow children to assign general thematic roles to arguments before encountering the verb. Thus our findings are consistent with theories of acquisition that are grounded in early semantic and syntactic abstraction (Fisher, 2002; Pinker 1984) and theories of processing which assume that comprehenders make incremental predictions based on the cues that are most reliable in their language (MacWhinney, Pleh, & Bates, 1985; Bates & MacWhinney, 1987; MacDonald et al. 1994; Tanenhaus, et al. 1995; Snedeker & Trueswell, 2004; Trueswell & Gleitman, 2007).

References

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). Graded representations in the acquisition of English and German transitive constructions. *Cognitive Development*, *23*(1), 48-66.

Akhtar, N., & Tomasello, M. (1997). Young children's productivity with word order and verb morphology. *Developmental Psychology, 33(6),* 952-965. doi: 10.1037/0012-1649.33.6.952

Aksu-Koç, A. A., & Slobin, D. I. (1985). The acquisition of Turkish. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition: Vol. 1, the data* (pp. 839-880). Hillsdale, NJ: Erlbaum.

Altan, A. (2006). Acquisition of a null subject language. In *Unpublished paper presented at the 13th International Conference on Turkish Linguistics* at Uppsala University, Sweden.

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

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

Arai, M., & Keller F. (2013). The use of verb-specific informa- tion for prediction in sentence processing. Language and Cognitive Processes, 28, 525–560. doi:10.1080/ 01690965.2012.658072

Arnold, J. (1998). Reference form and discourse patterns. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

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

Bader, M., & Häussler, J. (2010). Word order in German: A corpus study. *Lingua*, *120*(3), 717-762.

Baker, M. (1988). *Incorporation: a theory of grammatical function changing.* Chicago: University of Chicago Press.

Baker, C. L. (1979). Syntactic theory and the projection problem. *Linguistic Enquiry, 10(4)*, 533-581. Stable URL: http://www.jstor.org/stable/4178133

Barr, D. J. (2008). Analyzing 'visual world'eyetracking data using multilevel logistic regression. *Journal of memory and language*, *59*(4), 457-474.

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition.* New Jersey: Lawrence Erlbaum Inc. Publishers.

Bates, E., MacWhinney, B., Caselli, C., Devescovi, F. N., and Venza, V. (1984). A cross-linguistic study of development of sentence interpretation strategies. *Child Development, 55(2),* 341-354. doi: 10.2307/1129947

Batman-Ratyosyan, N. & Stromswold, K. (1999). What Turkish acquisition tells us about underlying word order and scrambling. *Proceedings of the 23rd Annual University of Pennsylvania Linguistics Conference,* 6, 37-52.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the Development of Language* (pp. 279-362). New York: Wiley.

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, *21*(1), 47–67. Borsley, R., & Börjars, K. (2009). *Non-transformational Syntax*. Oxford: Blackwells.

Boland, J. E., & Tanenhaus, M. K. (1991). *The role of lexical representations in sentence processing.* In G. Simpson (Ed.), Understanding word and sentence (pp. 331-366). New York: North-Holland.

Borovsky, A., Fernald, A., & Elman, J. (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. doi:10.1016/j.jecp.2012.01.005

Boyd J. K., & Goldberg AE. (2012). Younger children fail to fully generalize a novel argument structure construction when exposed to the same input as older learners. *Journal of Child Language, 39(3),* 457-481. DOI:10.1017/S0305000911000079

Bar, M. (2009). The pro-active brain: Memory for predictions. Philosophical Transactions of the Royal Society B, 364, 1235–1243. doi:10.1098/ rstb.2008.03103

Brauer, J. Anwander, A., Perani, D. and Friederici A.D. (2013). Dorsal and ventral pathways in language development. *Brain and Language, 127(2),* 289-295. doi:10.1016/j.bandl.2013.03.001

Byrne, B., & Davidson, E. (1985). On putting the horse before the cart: exploring conceptual bases of word order via acquisition of a miniature artificial language. *Journal of Memory and Language 24(4),* 377-389. doi:10.1016/0749-596X(85)90035-X

Candan, A., Küntay, A. C., Yeh, Y. C., Cheung, H., Wagner, L., & Naigles, L. R. (2012). Language and age effects in children's processing of word order. *Cognitive Development*, *27*(3), 205-221.

Carminati, M. N., & Knoeferle, P. (2013). Effects of speaker emotional facial expression and listener age on incremental sentence processing. *PloS one*, *8*(9), e72559.

Cartwright, T. A. & Brent, M. R. (1997). Syntactic categorization in early language acquisition: Formalizing the role of disctributional analysis. *Cognition, 63(2),* 121-170. doi:10.1016/S0010-0277(96)00793-7

Chan, A., Lieven, E. & Tomasello, M. (2009). Children's understanding of the agent – patient relations in the transitive construction: cross-linguistic comparisons between Cantonese, German and English. *Cognitive Linguistics 20(2),* 267-300. doi: 10.1515/COGL.2009.015

Choi, Y., & Trueswell, J. (2010). Children's (in)ability to recover from garden paths in a verb-final language: evidence for developing control in sentence processing. *Journal of Experimental Child Psychology*, *106*(1), 41-61. doi:10.1016/j.jecp.2010.01.003

Christiansen, M. H. & Charter, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*, p. 489-558. doi: 10.1017/S0140525X08004998

Clair, M. C., Monaghan, P. & Christiansen, M. H. (2010). Learning grammatical categories from distributional cues: Flexible frames for language acquisition. *Cognition, 116(3),* 341-360. doi:10.1016/j.cognition.2010.05.012

Comrie, E. (2009). *The world's major languages.* Oxon, UK: Routledge.

Dabrowska, E. & Tomasello, M. (2008). Rapid learning of an abstract language-specific category: Polish children's acquisition of the instrumental construction. *Journal of Child Language, 35(3),* 533-558. doi: 10.1017/S0305000908008660

Deutscher, G. (2005). *The Unfolding of Language*. New York: Holt Paperbacks.

De Vincenzi, M. (1991). *Syntactic parsing strategies in Italian: The minimal chain principle.* Dordrecht: Kluwer Academic.

Demiral, S. B. I., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2008). On the universality of language comprehension strategies: Evidence from Turkish. *Cognition*, *106*(1), 484-500. doi: 10.1016/j.cognition.2007.01.008

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. doi: 10.1111/j.1467-8624.2008.01181.x

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67(3), 547-619.

Dryer, M. S., 2008b. Relationship between the order of object and verb and the order of adjective and noun. In: Haspelmath, M., Dryer, M., Gil, D., Comrie, B. (Eds.), The World Atlas of Language Structures Online. Max Planck Digital Library, Munich, chapter 91. Available online at http://wals.info/feature/91.

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

Dryer, M. S. (2013a). Order of adjective and noun. In Matthew S. Dryer M. S. and Haspelmath, M. (eds.), The World Atlas of Language Structures Online. Max Planck Institute for Evolutionary Anthropology, Leipzig.

Dryer, M. S. (2013b). Order of adposition and noun phrase. In Dryer M. S. and Haspelmath, M. (eds.), The World Atlas of Language Structures Online. Max Planck Institute for Evolutionary Anthropology, Leipzig.

Elman, J., Hare, M. & McRae, K. (2005). Cues, Constraints, and Competition in Sentence Processing. In: *Beyond nature-nurture: Essays in honor of Elizabeth Bates.*

Fedzechkina, M., Jaeger, T.F., and Newport, E. L. (2012). Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences, 109(44)*, 17897-17902.

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. doi: 10.1111/1467-8624.00331

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 601–681). Hillsdale, NJ: Erlbaum.

Frazier, M., Ackerman, L., Baumann, P., Potter, D., & Yoshida, M. (2015). Wh-filler-gap dependency formation guides reflexive antecedent search. *Frontiers in psychology*, *6*: 1504. doi: 10.3389/fpsyg.2015.01504

Friederici, A. D. (2011). Brain basis of language processing: From processing to function. *Physiological Reviews, 91(4),* 1357-1392. doi: 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.

Friederici, A. D., & Frisch, S. (2000). Verb argument structure processing: The role of verb-specific and argument-specific information. *Journal of memory and language*, *43*(3), 476-507.

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. doi: 10.1111/1467-7687.00209

Fisher, C., Hall, D. G., Rakowitz, S., & Gleitman, L. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua*, *92*, 333-375.

Frisch, S. & Schlesewsky, M. (2005). The resolution of case conflicts from a neurophysiological perspective. *Cognitive Brain Research, 25(2),* 484-498. doi:10.1016/j.cogbrainres.2005.07.010

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. doi: 10.1111/j.1467-9280.2006.01767.x

Gibson E, Piantadosi ST, Brink K, Bergen L, Lim E, Saxe R. (2013). A noisy-channel account of crosslinguistic word order variation. *Psychological Science, 24(7),* 1079-1088. 2013. doi: http://dx.doi.org/10.1177/0956797612463705

Gleitman, L. (1990). The structural sources of verb meanings. *Language acquisition*, *1*(1), 3-55.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press on Demand.

Goldberg, A. E. (2011). Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics, 22(1),* 131-154. doi:10.1515/cogl.2011.006

Gorrel, P. (2000). The subject-before-object preference in German clauses. In Hemforth, B. and Konieczny, L. (eds.). *German sentence processing, Studies in Theoretical Psycholinguistics 24,* Dordrecht: Kluwer Academic Publishers, 25-63.

Gómez, R. L., and Lakusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science, 7,* 567-80. doi: 10.1111/j.1467-7687.2004.00381.x

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. doi: 10.1017/S0305000907008471

Greenberg, J. H. (1963). Some universals of grammar with particular reference to thr order of meaningful elemants. In J. H. Greenberg (ed.). *Universals of Language, (*pp. 73-113). London: MIT Press.

Hahne, A., Eckstein K., & Friederici, A. (2004). Brain Signatures of Syntactic and Semantic Processes during Children's Language Development. *Journal of Cognitive Neuroscience, 16(7)*, 1302-1318. doi:10.1162/0898929041920504

Hallé, P., Durand, C., & de Boysson-Bardies, B. (2008). *Do 11-month-old French infants process articles? Language and Speech, 51,* 23-44. doi: 10.1177/00238309080510010301

Haspelmath, M., Dryer, M., Gil, D., Comrie, B. (Eds.), 2005. *The World Atlas of Language Structures.* Oxford University Press, Oxford.

Höhle, B., and Weissenborn, J. (2003). German learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science, 6(2),* 122-127. doi: 10.1111/1467-7687.00261

Huang, Y., Zheng, X., Meng, X., & Snedeker, J. (2013). Assignment of grammatical roles in the online processing of Mandarin passive sentences. *Journal of Memory and Language, 69(4),* 589-606. doi:10.1016/j.jml.2013.08.002

Huang, Y. T., Leech, K., & Rowe, M. L. (2017). Exploring socioeconomic differences in syntactic development through the lens of real-time processing. *Cognition*, *159*, 61-75.

Iggesen, O. A. (2013). *Number of Cases.* Max Planck Institute for Evolutionary Anthropology, Leipzig.

January, D., J. C. Trueswell, and S. L. Thompson-Schill. 2009. Co-localization of Stroop and syntactic ambiguity resolution in Broca's Area: Implications for the neural basis of sentence processing. *Journal of Cognitive Neuroscience, 21(12),* 2434-2444. doi:10.1162/jocn.2008.21179

Jespersen, O. (1922). *Language: Its nature and development.* London: Allen and Unwin.

Kamide, Y., Scheepers, C., & Altmann, G. T. (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.

Kamide, Y., Altmann, G., & 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. doi:10.1016/S0749-596X(03)00023-8

Ketrez, & Aksu-Koc. (2009). Early nominal morphology in Turkish: Emergence of case and number. In U. Stephany & M. D. Voeikova (Eds.), *Development of Nominal Inflec-*

tion in First Language Acquisition: A Cross-linguistic Perspective (pp. 15-48). Berlin: Mouton De Gruyter.

Ketrez, F. N. (2004). Children's accusative case and indefinite objects. Dilbilim Araştırmaları, 15, 63-74.

Kit, C. (1998). Ba and Bei as milti-valence prepositions in Chinese –An LFG perspective. In B. K. T'sou (ed.) *Studia Linguistica Sinica.* City University of Hong Kong: Language Information Sciences Research Centre, 497-522.

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

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. doi:10.1016/j.neuroimage.2012.05.014

Kuperberg, G. R. and Jaeger, F. T. (2016). What do we mean by prediction in language comprehension?

Kornfilt, J. (1997). *Turkish*. London: Routledge.

Küntay, A. C. & Slobin, D. I. (2002). Putting interaction back into child language: examples from Turkish. *Psychology of Language and Communication 6,* 5-14.

Lany, J., and Gómez, R.L. (2008). Twelve-month-old infants benefit from prior experience in statistical learning. *Psychological Science, 19(12),* 1247-1252. doi: 10.1111/j.1467-9280.2008.02233.x

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. doi: 10.1037/a0029284

Marquis, A., & Shi, R. (2012). Initial morphological learning in preverbal infants. *Cognition, 122(1),* 61-66. doi:10.1016/j.cognition.2011.07.004

McNeill, D. (1975). Semiotic extension. In R. L. Solso (Ed.), *Information processing and cognition: the Loyola symposium.* Hillsdale, N.J.: Lawrence Erlbaum.

McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specic concepts. *Language and Cognitive Processes,12(2/3),*137-176. doi: 10.1080/016909697386835

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological review*, *101*(4), 676.

MacWhinney, B. (1977). Starting points. *Language*, *53*(1), 152–168.

Meyer, A.M., Mack, J.E., and Thompson, C.K. (2012). Tracking passive sentence comprehension in agrammatic aphasia. *Journal of Neurolinguistics, 25(1)*, 31-43. doin: 10.1016/j.jneuroling.2011.08.001

Mintz, T. H. (2006). Finding the verbs: distributional cues to categories available to young learners. In Hirsh-Pasek K., Golinkoff, R. M., (eds.), *Action Meets Word: How Children Learn Verbs.* New York: Oxford University Press, 31-63.

Naigles L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language. 17(2),* 357-374. doi: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. doi:10.1016/j.jecp.2003.09.001

Newmeyer, F.J., (2001). Grammatical functions, thematic roles, and phrase structure: their underliying disunity. In: Davies, D.W., Dublinsky, S. (Eds.), *Objects and Other Subjects.* Kluwer Academic Publishers, Dordrecht, 53-75.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass, 4(10),* 906-924.

Oflazer, K., Say, B., Hakkani-Tür, D. Z., & Tür, G. (2003). Building a Turkish treebank. *Treebanks*, 261-277.

Omaki A., Lau E. F., Davidson White I., Dakan M. L., Apple A., Phillips C. (2015). Hyper-active gap filling. Front. Psychol. 6:384. 10.3389/fpsyg.2015.00384

Osgood, C. E., and J. K. Bock (1977). Salience and sentencing: some production principles. Sentence Production - Developments in Research and Theory, (ed.) S. Rosenberg, (pp. 89-140). Hillsdale, NJ: Erlbaum.

Osgood, C. E., & Tanz, C. (1977). Will the real direct object in bitransitive sentences please stand up? In A. Juilland (Ed.), *Linguistic studies offered to Joseph Greenberg on the occasion of his sixtieth Birthday.* Saratoga, CA: Anna Libri.

Özge, D., Marinis, T., and Zeyrek, D. (2015). Incremental processing in head-final child language: on-line comprehension of relative clauses in Turkish-speaking children and adults. *Language, Cognition and Neuroscience (formerly titled Language and Cognitive Processes), 27*. doi: 10.1080/23273798.2014.995108

Özge, D., Küntay, A., Snedeker, J. (2018). Data for: Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension, *Mendeley Data, v1.* http://dx.doi.org/10.17632/yxn6cc8st4.1

Özge, U., & Bozsahin, C. (2010). Intonation in the grammar of Turkish. *Lingua*, *120*(1), 132-175.

Pozzan, L. and Trueswell, J.C. (2015). Revise and resubmit: How real-time parsing limitations influence grammar acquisition. *Cognitive Psychology, 80,* 73-108. doi: 10.1016/j.cogpsych.2015.03.004

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

Rowland, C., Chang, F., Ambridge, B., Pine, J. M. & Lieven, E. V. (2012) The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition, 125(1),* 49-63. doi:10.1016/j.cognition.2012.06.008

Sapir, E. (1921). *Language: An introduction to the study of speech.* New York: Harcourt, Brace.

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. doi: 10.1111/1467-7687.00312

Shafer, V. L., Shucard, J. L., Gerken, L. A., Shucard, D. W. (1998). 'The' and the brain: an electrophysiological examinations of infants' responses to familiar and unfamiliar function morphemes. *Journal of Speech, Language and HearingResearch, 41,* 874-886.

Schipke CS, Friederici AD, Oberecker R. 2011. Brain responses to case-marking violations in German preschool children. *Neuroreport, 22(16),* 850-854. doi: 10.1097/WNR.0b013e32834c1578

Schipke CS, Knoll LJ, Friederici AD, Oberecker R. 2012. Preschool children's interpretation of object-initial sentences: Neural correlates of their behavioral performance. Developmental Science, 15(6), 762-774. doi: 10.1111/j.1467-7687.2012.01167.x

Schlesewsky, M., Fanselow, G., Kliegl, R., & Krems, J. (2000). The subject preference in the processing of locally ambiguous wh-questions in German. In B. Hemforth & L. Konieczny (Eds.), *German sentence processing*, 65-93. Dordrecht: Kluwer.

Schriefers, H., Friederici, A. D., & Kühn, K. (1995). The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language, 34(4),* 499-520. doi:10.1006/jmla.1995.1023

Shi, R., Werker, J.F., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy, 10(2),* 187-198. doi: 10.1207/s15327078in1002_5

Shi, R. & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants. Developmental Science, 11(3), 407-413. doi: 10.1111/j.1467-7687.2008.00685.x

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

Snedeker, J. & Huang, Y. T. (2015). Sentence Processing. In Bavin, E. L. and Naigles, L. R. (eds.). *The Cambridge Handbook of Child Language. Second Edition.*Cambridge: Cambridge University Press.

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), 238-99. doi:10.1016/j.cogpsych.2004.03.001

Steele, S. (1978). *Word order variation: A typological study.* Stanford: Stanford University Press.

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

Szagun, G. (2004). Learning by ear: on the acquisition of case and gender marking by German-speaking children with normal hearing and with cochlear implants. *Journal of Child Language, 31(1),* 1-30. doi: 10.1017/S0305000903005889

Taylan, E. E. (1984). *The function of word order in Turkish grammar* (Vol. 106). LA: University of California Press.

Thothathiri, M., & Snedeker, J. (2008). Syntactic priming during language comprehension in three- and four-year-old children. *Journal of Memory and Language*, *58*(2), 188-213. doi:10.1016/j.jml.2007.06.012

Tomasello, M. (1992). *First verbs: A case study in early grammatical development.* Cambridge University Press.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

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

Tomasello, M. (2003). Constructing a language: A usage-based account of language acquisition. *Cambridge, MA*.

Treiman, R., Clifton, C., Meyer, A. & Wurn, A. (2003). Language comprehension and production. In Healy, A.; Proctor, R. (eds), *Handbook of Psychology: Experimental Psychology,* Vol. 4 (pp. 527-547). Hoboken, NJ, US: John Wiley & Sons, Inc.

Trueswell, J. C., & Gleitman, L. R. (2007) Learning to parse and its implications for language acquisition, in G. Gaskell (ed.) Oxford Handbook of Psycholinguistics. Oxford: Oxford Univ. Press.

Trueswell, J., 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. doi:10.1016/S0010-0277(99)00032-3

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role informa- tion in syntactic ambiguity resolution. *Journal of Memory and Language, 33(3),* 285-318. doi:10.1006/jmla.1994.1014

Trueswell, J. C., and Kim, A. E. (1998). How to prune a garden path by nipping it in the bud: fast priming of verb argument structure. *Journal of Memory and Language, 39(1),* 102-123. doi: 10.1006/jmla.1998.2565.

Ural, A. E., Yüret, D., Ketrez, F. N., Koçbaş, D. & Küntay, A. C. (2009). Morphological cues vs. number of nominals in learning verb types in Turkish: The syntactic bootstrapping mechanism revisited. *Language and Cognitive Processes, 24*, 1393-1405.

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

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

Weil, H. (1887). *The Order of Words in the Ancient Languages compared with that of the Modern Languages.* Transl. by Charles W. Super. Boston: Ginn. First published in French, Paris: Joubert [1844].

Weighall, A.R. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Child Psychology,* 99(2), 75-95. doi: 10.1016/j.jecp.2007.10.004

Weskot, T., Hörnig, R., Fanselow, G., and Kliegl R. (2011). Contextual licencing of marked OVS word order in German, *Linguistische Berichte, 225*, 3-18.

Yuan S, & Fisher C. (2008). "Really? She blicked the baby": Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science, 20(5),* 619-626. doi: 10.1111/j.1467-9280.2009.02341.x

Yuan, S., Fisher, C., and Snedeker, J. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child Development, 83(4),* 1382-1399. doi: 10.1111/j.1467-8624.2012.01783.x

Appendix 1: Ambiguities in Turkish word order and case

Turkish is a flexible word order language. Although the SOV is the most neutral order, all six word order variations are possible in line with the packaging of information structure. Generally speaking, the sentence-initial position hosts the topic (i.e., the referent the sentence is about) and the preverbal position receives the default focus accent hosting the new information (i.e., question under discussion) (Taylan, 1984). However, when all the referents are derivable from the context, the focus accent can fall on the verb (Özge, U. & Bozşahin, 2010).

The language heavily involves argument drop. Especially, subject-omission is especially quite frequent, so object-initial orders are natural. If the arguments are derivable from context, all arguments can be dropped, so the only word available for utterance interpretation in those cases would be the verb and the context. In the METU-Sabancı Treebank written adult corpus composed of 7262 grammatical sentences (Oflazer, Say, Hakkani-Tür, 2003), 52% of the NP-initial sentences has the subject and 48% of them has the object as their first argument (Candan, et al., 2012). Similarly, on a count of the first 136580 sentences from the *Milliyet* Corpus, subject-initial orders (SOV, SVO) (35.95%) are more frequent than object-initial orders (OV, OVS, OSV) (16.71%); there is a substantial amount of subject drop (i.e., 40.18% of this amounts to the object initial (OV) ordering increasing the frequency of object initial orders to 56.89%); and verb-final word orders (SOV, OSV, OV) are more frequent (76.64%) than verb-medial ones (SVO, OVS) (16.21%) (Table 10).

Table 10: Frequency of different word orders in Turkish from the *Milliyet* Corpus

Order	Frequency of Occurrence (%)
SOV	33.70%
OSV	2.75%
SVO	2.25%
OVS	13.96%
VOS	.53%
VSO	2.38%
OV	40.18%
VO	4.21%

This pattern is also reflected in Slobin & Bever's (1982) child-directed speech based on the natural utterances of 14 children and their parents, where the verb-final utterances constituted the 53% of the child and 58% of the parent utterances while the verb-medial orders represented the 37% of the child and 38% of the parent utterances. Although this study does not report the percentage of object-initial utterances and subject-drop, it is suggested that children fully reflect adult utterance patterns, so we would expect frequent subject-drop and object-initial sentences in child language as well. This is also supported by Altan's (2006) corpus study based on 36 children aged of 2;0 to 4;08, where children drop subjects around 70% of the time in a pragmatically appropriate manner.

Apart from word order variation, being an agglutinative language, Turkish is also rich in morphosyntax. Nominal suffixes mark case information. Despite the case morphemes are relatively reliable, as they are not conflated with any additional grammatical information, these morphemes still have various types of ambiguities. For instance, the nominative case is an unmarked subject case (e.g., 'tilki/fox' in (14)).

(14)

Tilki	[tavşan	kovala-dı]	bütün gün.
fox	rabbit	chase-Past.3sg	whole day
Subject	Object	Verb	

'The fox chased rabbits whole day.'

Yet, not every unmarked sentence-initial noun has to be the subject in Turkish; it could also be a direct object because the case marking can be optional for generic objects (pseudoincorporation) and for indefinites in some contexts (e.g., 'tilki/fox' in (15)).

(15)

[Tilki	kovala-dı]	tavşan	bütün gün.
fox	chase-Past.3sg	rabbit	whole day
Object	Verb	Subject	

'The rabbit chased foxes whole day.'

The subject marking is not uniform, either. Subject is marked in the nominative case when it is the subject of a simple clause, as in (16). However, if it is the subject of an embedded clause, it could be marked either in the nominative case, as in (17) or in the genitive case, as in (18), depending on the type of the embedded verb.

(16)

Tilki	tavşan-ı	kovala-dı.
fox	rabbit-Acc	chase-Past.3sg

'The fox chased the rabbit.'

(17)

Tilki	tavşan-ı	kovala-dığ-ı	için	yor-ul-du.
fox	rabbit-Acc	chase-Comp-Poss.3sg	for	tired-Pass-Past.3sg

'The fox got tired because of chasing the rabbit/as it chased the rabbit.'

(18)

Tilki-nin	tavşan-ı	kovala-dığ-ı-nı	gör-dü-m.
fox-Gen	rabbit-Acc	chase-Comp-Poss.3sg-Acc	see-Past-1sg

'I saw that the fox chased the rabbit.'

The accusative case as an object marker is a relatively straightforward cue; yet, it is still ambiguous because it is homophonous with the third person singular possessive suffix marked on a consonant ending stem, as in (19) (c.f., 20). Crucially, the third person possessive suffix is very frequent, as it is obligatory in noun-noun compounds (e.g., otobüs bilet-i/ bus ticket-Poss.3sg/ ‘bus ticket’).

(19)

Tavşan-ı	kovala-n-dı.
rabbit-Poss-3sg	chase-Pass-Past.3sg

‘(His/Her) rabbit was chased.’

(20)

Tavşan-ı	kovala-dı.
rabbit-Acc	chase-Past.3sg

‘(Someone/Subject) chased the rabbit.’

Appendix 2: Test items in each condition

Critical Items

Nominative Condition

1- Hızlı tavşan birazdan yiyecek şuradaki havuçu.

“The speedy rabbit will soon eat the carrot over there.”

2- Tatlı civciv birazdan yutacak şuradaki mısırı.

"The sweet chicken will soon swallow the corn over there."

3- Şeker kız birazdan öpecek şuradaki bebeği.

"The cute girl will soon kiss the doll over there."

4- Şirin bebek birazdan yalayacak şuradaki dondurmayı.

"The pretty baby will soon lick the ice-cream over there."

5- Tombul ayı birazdan bulacak şuradaki balı.

"The chubby bear will soon find the honey over there."

6- Ufak sincap birazdan yiyecek şuradaki fındığı.

"The small squirrel will soon eat the hazelnut over there."

7- Küçük çocuk birazdan tekmeleyecek şuradaki topu.

"The little child will soon kick the ball over there."

8- Uslu çocuk birazdan kucaklayacak şuradaki oyuncağı.

"The good child will soon hug the teddy bear over there."

9- Komik maymun birazdan ısıracak şuradaki muzu.

"The funny monkey will soon bite the banana over there."

10- Minik fare birazdan bulacak şuradaki peyniri.

"The small mouse will soon find the cheese over there."

Accusative Condition

1- Hızlı tavşanı birazdan yiyecek şuradaki tilki.

"The fox over there will soon eat the speedy rabbit."

2- Tatlı civcivi birazdan yutacak şuradaki kurt.

"The wolf over there will soon swallow the sweet chicken."

3- Şeker kızı birazdan öpecek şuradaki kadın.

"The woman over there will soon kiss the cute girl."

4- Şirin bebeği birazdan yalayacak şuradaki köpek.

"The dog over there will soon lick the pretty baby."

5- Tombul ayıyı birazdan bulacak şuradaki avcı.

"The hunter over there will soon find the chubby bear."

6- Ufak sincabı birazdan yiyecek şuradaki kurt.

"The wolf over there will soon eat the small squirrel."

7- Küçük çocuğu birazdan tekmeleyecek şuradaki at.

"The horse over there will soon kick the little child."

8- Uslu çocuğu birazdan kucaklayacak şuradaki adam.

"The man over there will soon hug the good child."

9- Komik maymunu birazdan ısıracak şuradaki aslan.

"The lion over there will soon bite the funny monkey."

10- Minik fareyi birazdan bulacak şuradaki kedi.

"The cat over there will soon find the small mouse."

Filler Items

1- Güzel kadının aldığı yumurta birazdan kırılacak.

"The egg that the pretty woman bought will crack soon."

2- Dikkatsiz çocuğun balonu birazdan patlayacak.

"The careless boy's baloon will pop soon."

3- Obur ineğin korktuğu arı birazdan uçacak.

"The bee that the gargantuan cow frightened will fly soon."

4- Geveze zebranın balığı birazdan yüzecek.

"The chatter zebra's fish will swim soon."

5- Genç polisin bindiği gemi birazdan batacak.

"The ship that the young policeman takes will sink soon."

6- Yorgun hemşirenin odasındaki mum birazdan sönecek.

"The candle in the tired nurse's room will go off soon."

7- Şakacı kuzunun kovaladığı eşek birazdan koşacak.

"The donkey that the funny lamb chases will run away soon."

8- Şaşkın penguenin yanındaki buz birazdan eriyecek.

"The ice next to the weird penguin will melt soon."

9- İyi öğretmenin beslediği ördek birazdan dalacak.

"The duck that the good teacher feeds will dive in soon."

10- Büyük ağacın yanındaki kurbağa birazdan zıplayacak.

"The frog next to the big tree will hop soon."

Appendix 3: 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

In order to see whether our participants looked at each object as it was mentioned and how their gaze patterns on the visual scene changed as they heard incoming linguistic input, we plotted percentage of looking time to each of the object (topic, plausible agent, and plausible patient) in each of the time-windows for children and adults for both Experiment 1 and Experiment 2 below.

Figures 11 and 12 show the percentage of looking time to each of the object (topic, plausible agent, and plausible patient) in each of the time-windows for both the children and adults in Experiment 1.

Figure 11: Gaze patterns for adults in Experiment 1

Figure 11a: Percentage of looking time to each of the objects (the topic, the plausible agent, and the plausible patient) on the visual scene in the nominative condition

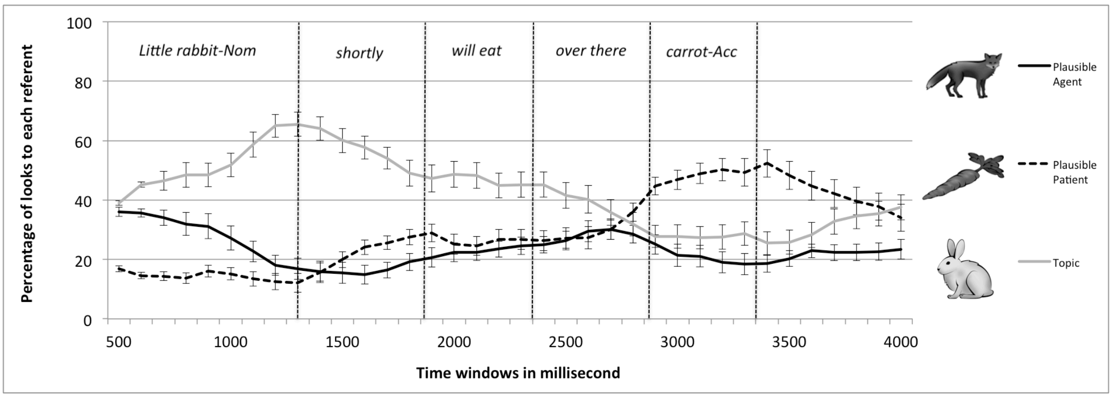


Figure 11b: Percentage of looking time to each of the objects (the topic, the plausible agent, and the plausible patient) on the visual scene in the accusative condition

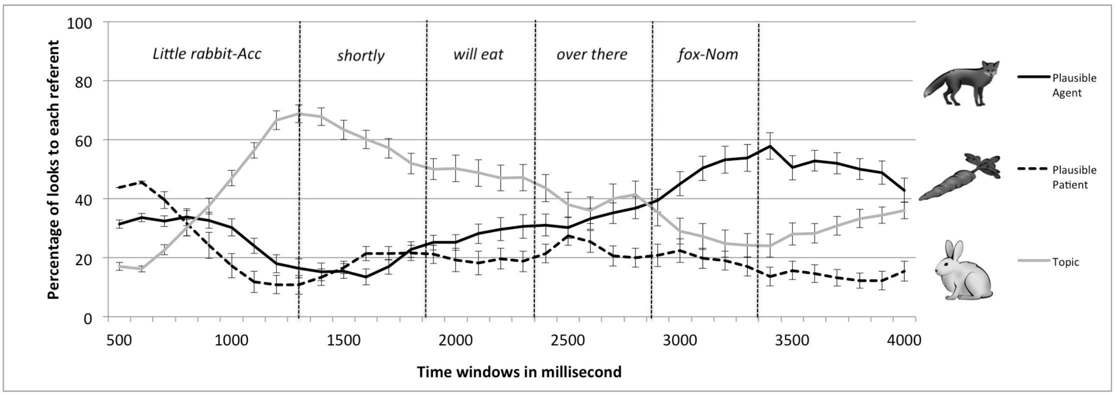


Figure 12: Gaze patterns for children in Experiment 1

Figure 12a: Percentage of looking time to each of the objects (the topic, the plausible agent, and the plausible patient) on the visual scene in the nominative condition

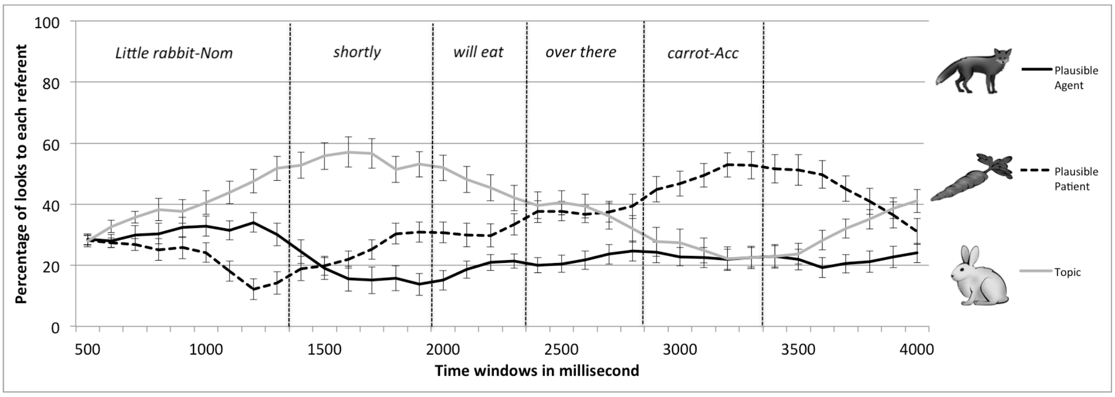
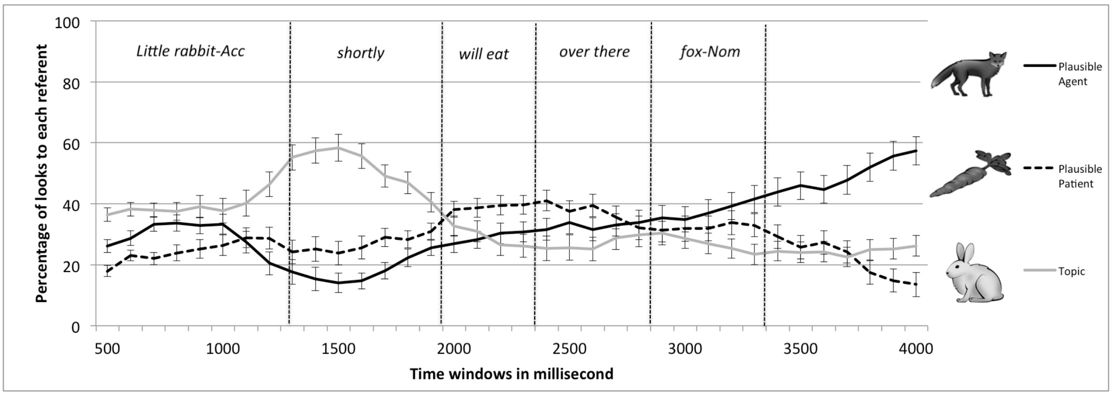


Figure 12b: Percentage of looking time to each of the objects (the topic, the plausible agent, and the plausible patient) on the visual scene in the accusative condition



Participants in both groups looked at each object as it was mentioned, with gaze to the topic peaking in the adverb region and gaze to the referent of the second noun rising toward the end of the trial. Our primary interest was in the looking pattern after the first noun and prior to the second noun. If participants are using case to predict the upcoming referent, then we should see greater looking to the potential agent immediately after encountering the accusative marked patient in the accusative condition than in the nominative condition where they encounter the nominative marked agent, and greater looking to the potential patient in the nominative condition than in the accusative condition. In the nominative condition, adults shifted their looks to the plausible patient (carrot) after the first noun during the adverbial region while shifting their looks to the plausible agent (fox) towards the end of the adverbial region during the verb region, so they were a little slower to shift their gaze to the correct referent in the accusative condition compared to the nominative condition. Children showed a similar pattern to adults in the nominative condition while shifting to the correct referent later than adults (toward the end of the modifier region). Nevertheless, comparing the two conditions, both groups showed greater agent preference in the accusative condition compared to the nominative condition (adults during the verb region, children during the modifier region).

Figures 13 and 14 show the percentage of looking time to each of the object (topic, plausible agent, and plausible patient) in each of the time-windows for both the children and adults in Experiment 2.

Figure 13: Gaze patterns for adults in Experiment 2

Figure 13a: Percentage of looking time to each of the objects (the topic, the plausible agent, and the plausible patient) on the visual scene in the nominative condition

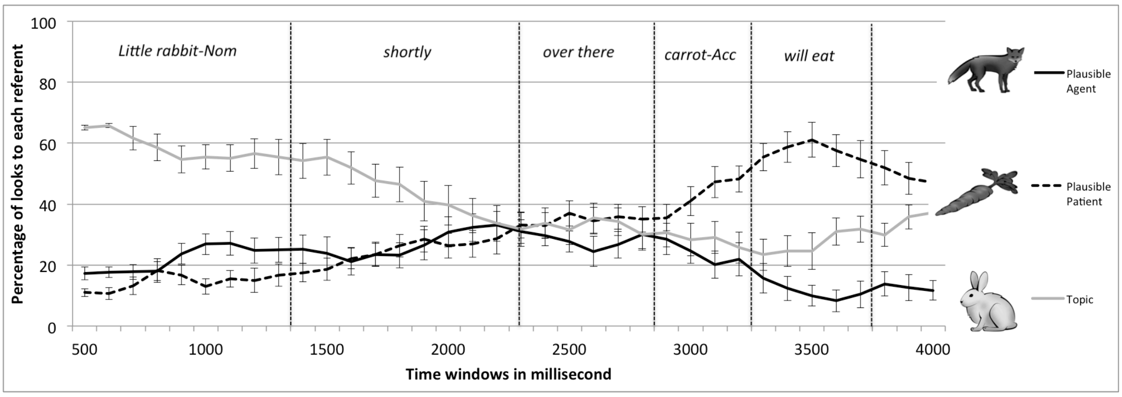


Figure 13b: Percentage of looking time to each of the objects (the topic, the plausible agent, and the plausible patient) on the visual scene in the accusative condition

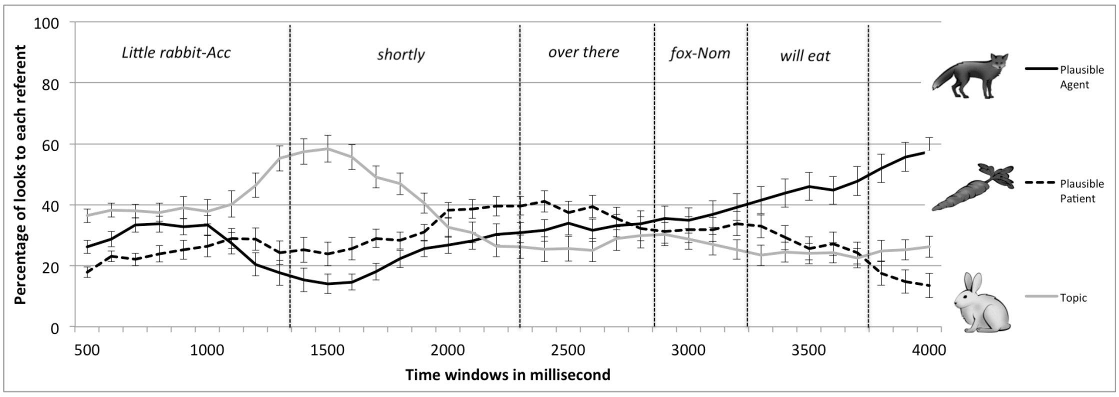


Figure 14: Gaze patterns for children in Experiment 2

Figure 14a: Percentage of looking time to each of the objects (the topic, the plausible agent, and the plausible patient) on the visual scene in the nominative condition

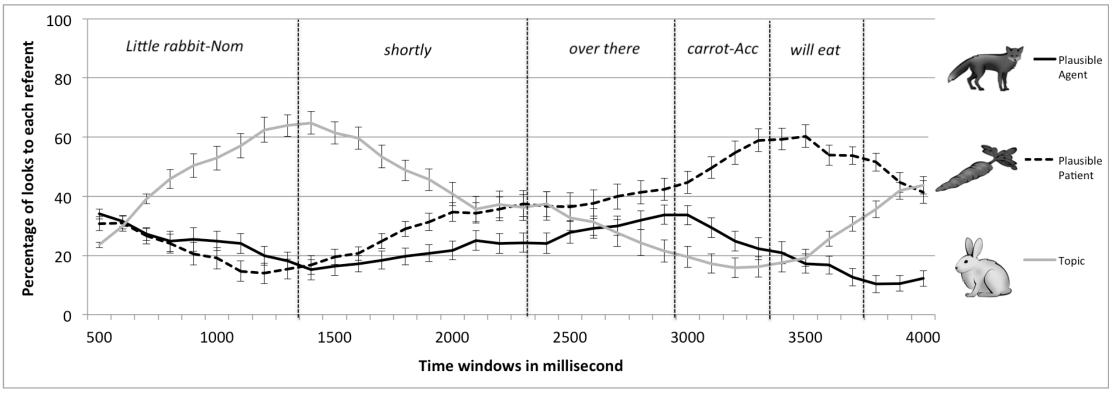


Figure 14b: Percentage of looking time to each of the objects (the topic, the plausible agent, and the plausible patient) on the visual scene in the accusative condition

Similar to Experiment 1, we see that participants look at each object when it is mentioned, gazes to the topic peaks at the end of the first noun and in the adverbial region, and the gaze to the second referent increasing toward the end of the second noun. Adults show more shifts between referents and less clear preference for the correct upcoming referent during the predictive time windows in this experiment (see Figure 13b). We reason that this may due to the hypothetical nature of parsing. The parser assigns the most plausible analysis to an available string and generates a provisional analysis for the upcoming string. The parser hears the first noun (rabbit), at this point it is certain that the sentence is about the rabbit, then it integrates the case marker on the first noun and comes up with a probabilistic interpretation about the role of this first entity and create expectations about the upcoming plausible structure. Yet, since there are various morphosyntactic ambiguities at each point in the utterance, all of the syntactic and semantic analyses (for the available string and for the upcoming string) should remain provisional until further information is encountered. For instance, the utterance with a sentence-initial accusative marked noun could continue in several very plausible ways. Sentence initial noun marked in –i morpheme could be a direct object in simple sentence (11), it could be a direct object of an embedded sentence that does not include any agent entity (12), or it could be a possessed noun in a possessive noun phrase whose possessor (his/her) is dropped (13). The reason why adults show more frequent shifts between referents might be due to greater awareness about these ambiguities both in the nominative and in the accusative condition.

(11)	Minik tavşan-ı	birazdan	şu	tilki	bul-acak.
	small rabbit-Acc	soon	that	fox	find-Fut-3sg

‘That fox will soon find the little rabbit.’

(12) Minik tavşan-ı	birazdan	şu	havuç-un	yanında	gör-eceğ-iz.
small rabbit-Acc	soon	that	carrot-Gen	next to	see-Fut-1pl

‘We will soon see this little rabbit next to that carrot.’

(13) (O-nun)	minik tavşan-ı	birazdan	şu	havuç-u	yi-yecek.
(s/he-Gen)	small rabbit-Poss3sg	soon	that	carrot-Acc	eat-Fut-3sg

‘His/her small rabbit will soon eat that carrot.’

This also explains why we get far more looks to the topic referent compared to the referents yet to be mentioned. The parser is sure of only one thing: the sentence is about the rabbit, all other analyses should be kept provisional so that they can be updated in line with the incoming input until the parser is sure of who is doing what to whom.

Appendix 4: Alternative analyses for Experiment 1

To see whether our analysis would lead to different results if analyzed using a method similar to Kamide and colleagues’ (2003) study, we conducted a repeated measures ANOVA with Case (Nominative, Accusative) and Referent (Plausible-Agent –fox, Plausible-Patient – carrot) as within-subjects factor and Group (Children, Adults) as between-subjects factor for the critical time window (TW7, Modifier Region). As mentioned before, if the participants looked at the appropriate referent upon hearing the case-marked NP1, we would find a significant interaction between Case and Referent. This is exactly what we found. There was no main effect of the Case $F(1, 74) = 2.32$, $p = .132$ or the Referent $F(1, 74) = .001$, $p = .982$, but there was a significant interaction between Case and Referent $F(1, 74) = 10.22$, $p = .002$. Pairwise comparisons with Bonferonni corrections revealed that while there were more looks to the plausible agent (fox) (M=36.43; SE=2.73) than to the plausible patient (carrot) (M=27.58; SE=2.20) in the Accusative condition, and there were more looks to the Plausible Patient (carrot) (M=34.18; SE=2.19) than to the plausible agent (M=25.20; SE=2.19) in the Nominative condition. This was true for both groups as there was no interaction between Case, Referent, and Group $F(1, 74) = .144$, $p = .705$, so both children and adults were able to look at the appropriate referent in both case conditions. In addition to this, there was also a significant Group by Referent interaction $F(1, 74) = 7.04$, $p = .010$: while children showed an overall greater Patient looks (M= 35.82; SE = 2.28) compared to agent looks (M= 28.14; SE= 2.71) adults did not show this pattern. We think this might be due to an overall agent-first bias in children but not in adults. This is fully in line with our analysis above with mixed-effects logistic regressions, where we found greater agent preference in the accusative condition compared to the nominative condition accompanied by a greater agent-first expectation in children.

We did one more ANOVA analysis using the log values of Agent-Patient proportion (i.e., ln(%AgentLooks/%PatientLooks)) with Case (Nominative, Accusative) as within-subjects factor and Group (Children, Adults) as between-subjects factor for the critical time window (TW7, Modifier Region), which is an analysis similar to the one used in Carminatti & Knoeferle (2013). In this analysis, similar to the analysis we conducted using mixed-effects logistic regressions, the dependent variable is symmetrical around zero: if there are equal looks to both referents the value is zero, if the agent preference is greater, the value is positive and if is low the value is negative. Therefore, we expect to see an effect of case if case marker had a significant influence on participants' predictive looks. There was indeed the effect of case $F(1, 74) = 11.50$, $p = .001$ and no interaction between group and case $F(1, 74) = .99$, $p = .3$. There was greater agent looks in the accusative condition (M=.52-; SE=.26) compared to the nominative condition (M=-.79; SE=.29).

To further see whether the looks to the Topic significantly differed for each condition for the critical time window, we conducted a one-way ANOVA with Case (Nominative, Accusative) as within-subjects subjects factor and Group (Children, Adults) as between-subjects factor for each experiment. There was no effect of Case $F(1, 74) = 3.05$, $p = .085$ and no interaction between Case and Group $F(1, 74) = .76$, $p = .38$ (Table 11). Thus, the amount of topic looks was the same for both conditions for the critical time window.

Table 11: Mean percentage of Topic fixations in the critical time window (TW7)

Condition	Group	Mean	Standard Deviation
Accusative	Children	28.94	18.61
	Adults	35.82	24.30
Nominative	Children	36.74	22.41
	Adults	38.43	21.19

Appendix 5: Alternative analyses for Experiment 2

For Experiment 2, we also conducted an ANOVA with Case (Nominative, Accusative) and Referent (Plausible-Agent –fox, Plausible-Patient –carrot) as within-subjects factor and Group (Children, Adults) as between-subjects factor for the critical time window. According to this, there was no main effect of the Case $F(1, 59) = .75$, $p = .38$ or the Referent $F(1, 59) = 1.28$, $p = .15$, but there was a significant interaction between Case and Referent $F(1, 59) = 16.49$, $p = .000$. Similar to Experiment 1, there were more looks to the plausible agent (fox) (M=41.38; SE=2.84) than to the plausible patient (carrot) (M=25.74; SE=2.65) in the accusative condition while there were more looks to the plausible patient (carrot) (M=36.17; SE=2.60) than to the plausible agent (M=28.32; SE=2.48) in the Nominative condition. There was no other interactions between Case and Group $F(1, 59) = 2.07$, $p = .15$, or between Case, Referent, and Group $F(1, 59) = .02$, $p = .88$.

The second ANOVA analysis using the log values of Agent-Patient proportion (i.e., ln(%AgentLooks/%PatientLooks)) with Case (Nominative, Accusative) as within-subjects factor and Group (Children, Adults) as between-subjects factor for the critical time window (TW7, Modifier Region) showed and effect of case $F(1, 59) = 9.88$, $p = .003$ and no interaction between group and case $F(1, 59) = .55$, $p = .4$. There was greater agent looks in the accusative condition (M=.84-; SE=.28) compared to the nominative condition (M=-.56; SE=.35).

Finally, as we did in Experiment 1, we analyzed whether the looks to the topic were significantly different for each condition for the critical time window, we conducted a one-way ANOVA with Case (Nominative, Accusative) as within-subjects subjects factor and Group (Children, Adults) as between-subjects factor. There was no effect of Case $F(1, 59) =$

.70, p = .4 or no Case by Group interaction Case F(1, 59) = 2.30, p = .13 (Table 12). Here too, the topic looks did not differ by condition.

Table 12: Mean percentage of Topic fixations in the critical time window (TW7)

Condition	Group	Mean	Standard Deviation
Accusative	Children	24.53	19.92
	Adults	35.64	22.41
Nominative	Children	31.64	21.10
	Adults	33.60	20.70

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 and TÜBİTAK 1001 Grant (110K424) to A. Küntay and D. Özge. We thank Deniz Özkan, Berna Uzundağ and Pelinsu Bulut for help with data collection; Deniz Vidinli for her help in data analysis of the relatedness study; Ilgar Veryeri for the visual material we used in our experiments; and Umut Özge for the data trimming program and for computing the corpus frequencies of word order variation in Turkish that is presented in Table 10.