The role of syntax in sentence and referential processing

Roger P.G. van Gompel

School of Psychology

University of Dundee

Juhani Järvikivi

Department of Linguistics

University of Alberta

Address correspondence to:

Roger P.G. van Gompel

School of Psychology

University of Dundee

Dundee DD1 4HN

United Kingdom

Tel: +44 1382 388117

E-mail: r.p.g.vangompel@dundee.ac.uk

Introduction

How language comprehenders process the syntactic structure of sentences and, to a somewhat lesser extent, how sentence structure affects referential processing have been important questions in language comprehension research. Results from studies using the *visual-world eye-tracking method* have yielded important insights regarding these issues (see Pyykkönen-Klauck & Crocker, and Spivey & Huette, this volume, for overviews of the visual-world method). This chapter reviews visual-world studies that have done this.

The use of context in the processing of syntactically ambiguous sentences

Following early work by Cooper (1974), Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy (1995) kick-started present-day research using the visual-world eye-tracking method. In this study, they contrasted two main accounts of how sentence structure is processed. According to modular accounts, sentence structure is initially processed using domain specific, structural information. Among modular accounts, the most influential account has been the garden-path theory (Frazier, 1979; Frazier & Rayner, 1982; Rayner, Carlson, & Frazier, 1983), which claims that in cases of structural ambiguity, language comprehenders adopt the analysis that is structurally least complex. Other potentially useful information such as context and semantics are only used during later processing. The assumption is that, for structural processing, the human sentence processing mechanism uses an autonomous processing module that is informationally encapsulated and therefore is not directly influenced by non-structural information. In contrast, a second type of account assumes that processing sentence structure involves the immediate use of various sources of constraining information such as context, semantics and the frequency of structures. In cases of structural ambiguity, these information sources simultaneously activate the different structural analyses, and the analysis that is activated most is adopted. These theories are generally referred to as constraint-satisfaction theories (MacDonald, Pearlmutter, & Seidenberg, 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Trueswell, Tanenhaus, & Garnsey, 1994).

Because listeners' fixations to the objects/pictures that are mentioned in a sentence provide a continuous, fine-grained record of auditory language processing,

the visual-world eye-tracking method has been a very fruitful way of testing these opposing accounts of sentence processing. A high temporal resolution is critical, because modular accounts assume that non-structural information can affect late processing stages, so it is essential that the method can distinguish between early and later processing.

Tanenhaus et al. (1995) investigated whether information from the concurrent visual context has an immediate effect on the processing of sentence structure. They tested temporarily ambiguous sentences such as (1a) and compared them with unambiguous sentences such as (1b).

1a. Put the apple on the towel in the box.
1b. Put the apple that's on the towel in the box.

Sentence (1a) is temporarily ambiguous because the prepositional phrase (PP) *on the towel* modifies *the apple*, but could initially also be analysed as the destination of the apple (as in *Put the apple onto the towel*). The latter analysis is ruled out by the subsequent PP *in the box*. In contrast, in sentence (1b), *that's on the towel* unambiguously modifies *the apple*. The question is whether structural processing of the temporarily ambiguous PP *on the towel* in (1a) is affected by the context (Crain and Steedman, 1985; Altmann & Steedman, 1988). If only one apple is in the context, a definite noun phrase (NP) with a modifier may be infelicitous, because there is no other entity to which *the apple* can refer, so a modifier would provide redundant information. However, if there is more than one apple, a modifier is needed to specify which apple is referred to. Therefore, if context influences structural analysis, then language comprehenders should initially adopt the (incorrect) destination analysis of *the towel* in a context with one apple (one-referent context), but adopt the modifier analysis in a context with two apples (two-referent context).

Prior to Tanenhaus et al. (1995), the effect of context on syntactic ambiguity resolution had been investigated in reading studies (e.g., Altmann & Steedman, 1988; Britt, 1994; Ferreira & Clifton, 1986; Murray & Liversedge, 1994; Van Berkum, Brown, & Hagoort, 1999). Typically in these studies, the target sentence was preceded by a linguistic context in which one or two referents were introduced. Because these studies had not shown consistent effects of context on syntactic ambiguity resolution, Tanenhaus et al. (1995) argued that context effects may have

been weak in some studies because the context was linguistic and therefore had to be kept in memory. Instead, in their visual-world study, they presented visual, real-world contexts while participants processed the temporarily ambiguous sentence.

While participants listened to sentences such as (1), they either saw a one-referent context containing an apple on a towel, an empty towel without an object on it, an empty box and a distractor object (a pencil) or a two-referent context containing an apple on a towel, another apple on a napkin, an empty towel, and an empty box. Their task was to act out the instruction (1a and 1b, above) using the objects in the scene. Participants' eye movements were recorded while they listened to the sentences. After hearing *towel* in (1a), participants looked at the empty towel in 55% of cases in the one-referent context. This indicates that they initially misinterpreted *on the towel* as the destination of the apple and looked at the towel because they thought they had to put it there. They never did this when hearing the unambiguous (1b). In contrast, in the two-referent context, participants rarely looked at the towel either when the sentence was ambiguous or unambiguous, but immediately looked at the box suggesting that in neither case did they consider the destination interpretation of *on the towel*.

Spivey, Tanenhaus, Eberhard, and Sedivy (2002) found similar results. This study also addressed a concern with the two-referent condition: Listeners might not have fixated the incorrect destination (the empty towel) because if they adopted the destination interpretation of *on the towel*, then *the apple* would not have a modifier, so it could refer either to the target apple (which is on a towel) or the distractor apple (which is on a napkin). Because this results in more fixations to the distractor apple, the incorrect destination (the empty towel) might have been fixated less. Therefore, Spivey et al. included a condition where the distractor consisted of three apples. In this condition, participants rarely fixated the three apples when they heard *the apple*, but despite this, they also rarely fixated the incorrect destination. This suggests that participants realised that *the apple* did not refer to the three apples, but nevertheless interpreted *on the towel* as a modifier because there was more than one apple in the visual context. These findings provide clear evidence that the visual context has a very rapid effect on the processing of sentence structure. This is consistent with constraint-satisfaction theories, but problematic for modular theories of sentence processing.
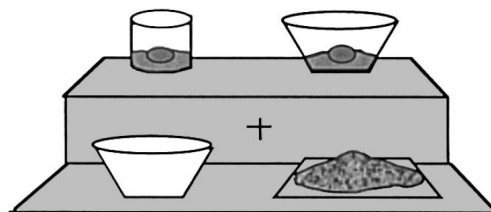
However, Chambers, Tanenhaus, and Magnuson (2004) argued that these results may still be consistent with a more general notion of modularity: Referential context may directly inform linguistic representations of discourse, and may therefore represent information that is intrinsic to the linguistic module. As a stricter test of modularity, they investigated the effects of action-based affordances on structural processing. For example, the act of pouring is compatible with the affordances of a liquid egg, but not with those of a solid egg.

Chambers et al. (2004) presented temporarily ambiguous sentences such as (2) while participants saw scenes that contained either two liquid eggs (Fig. 1a) or a liquid and a solid egg (Fig. 1b).

2a. Pour the egg in the bowl over the flower.
2b. Pour the egg that's in the bowl over the flower.

a. Two objects compatible with action   b. One object compatible with action
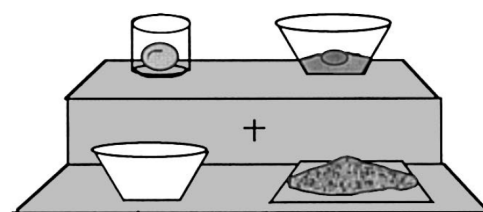
Fig. 1: Visual contexts in Chambers et al. (2004), Experiment 1.

The results from the conditions with two liquid eggs, which both afforded the pouring action mentioned in (2), were the same as the results from two-referent contexts in Tanenhaus et al. (1995) and Spivey et al. (2002): Participants fixated the incorrect destination (the bowl) no more often in the ambiguous than the unambiguous condition. But when there was only one liquid egg that afforded the action, the results were similar to those from one-referent contexts: Participants fixated the incorrect destination more often in the ambiguous than unambiguous condition, suggesting that they initially adopted the destination interpretation.

A second experiment manipulated affordances in a different way. Participants listened to sentences such as (3) while they saw displays such as Fig. 2.

3a. Put the whistle on the folder in the box.

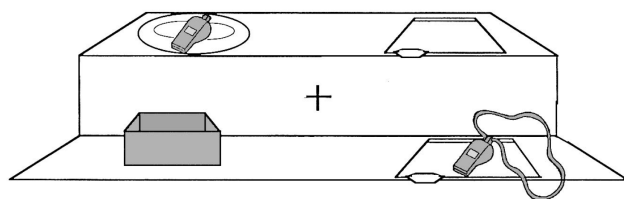3b. Put the whistle that's on the folder in the box.

Fig. 2: Visual context in Chambers et al. (2004), Experiment 2.

In one condition, participants had to use a hook to pick up the objects, whereas in another condition, they did not have a hook and had to pick up the objects by hand. One of the whistles had a string attached, so it was the only whistle that could be picked up in the "hook" condition, whereas both whistles could be picked up in the "no-hook" condition. This affordance manipulation had a clear effect on listeners' interpretation of *on the folder*: In the "no-hook" conditions, they looked no more often at the incorrect destination in the ambiguous than the unambiguous sentences, whereas in the "hook" conditions, they looked at the destination more often in the ambiguous sentences. These effects occurred rapidly, from about 200 ms after the onset of *folder*. Thus, listeners quickly adopted the destination interpretation in the ambiguous "hook" condition, but they adopted the modifier in the ambiguous "no hook" condition. Whether an object can be picked up with a hook is clearly not part of the linguistic representation of *hook* or *put*, so Chambers et al. concluded that this provides strong evidence that structural processing is not modular.

In sum, the results of visual-world studies provide strong evidence that both referential visual context and action-based affordances have very rapid effects on the interpretation of structurally ambiguous sentences. They support constraint-satisfaction theories, but are less compatible with modular theories. The latter type of theory would have to assume that the delay in using non-structural information is extremely short, and therefore not detectable with the visual-world method, even though it provides a very fine-grained temporal record of sentence processing. Given that differences in fixations to the incorrect destination in the ambiguous one- and two-referent conditions typically start arising during the noun in the temporarily ambiguous PP (e.g., *towel* in (1)), listeners would have to revise their initial

destination analysis in the two-referent conditions during the presentation of the preposition and article (*on the*).

Syntactic ambiguity resolution: Children vs. adults

However, there is evidence that young children are less sensitive to visual context. The visual-world method turns out to be ideally suited to study language comprehension in young children before they learn to read because it allows the investigation of spoken (rather than written) language comprehension and does not require the children to carry out a complex task. Trueswell, Sekerina, Hill, and Logrip (1999) were the first to test children using this method. They used the same structural ambiguity as Tanenhaus et al. (1995) and indeed found very similar results with adult participants. However, the results from 5-year-old children were different. Shortly following the temporarily ambiguous PP *on the towel*, children looked more often at the incorrect destination (the empty towel) in the ambiguous conditions than the unambiguous conditions, and most important, this effect was equally strong in the one- and two referent conditions. Thus, the children appeared to misinterpret the temporarily ambiguous phrase as the destination regardless of whether the context supported this interpretation or not. Furthermore, in about 60% of trials children failed to carry out the correct action (putting the apple that is on the towel into the box) in both the one- and two-referent ambiguous conditions, with no difference in the number of incorrect actions between these conditions. Thus, in both conditions children frequently failed to revise their initial destination interpretation of the temporarily ambiguous phrase *on the towel* into a modifier interpretation.

Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell (2000) showed that children are unable to use context during structural processing even though they *do* produce modifiers successfully in two-referent contexts. One obvious possibility is that children rely more exclusively on structurally-based processing strategies than adults, and therefore adopt the structurally least complex analysis regardless of context. Alternatively, the strong destination preference may not be due to structural complexity, but due to verb-specific, lexical biases. The verb *put* is virtually always immediately followed by a destination rather than a modifier PP, which may explain both why adults initially adopt this interpretation in the one-referent condition and why children adopt it in either context condition.

This issue was explored by Snedeker and Trueswell (2004). They investigated how adults and 5-year-old children process syntactic ambiguities such as (4).

4. Tickle/Choose/Feel the frog with the feather.

In these sentences, *with the feather* can be analysed as the instrument of *to tickle/choose/feel* or as a modifier of *the frog*. A sentence completion study showed that verbs such as *tickle* have an instrument bias, verbs such as *choose* have a modifier bias*,* and verbs such as *feel* have no clear bias. Snedeker and Trueswell tested all three verb types and also manipulated the visual context. In the one-referent context (Fig. 3a), participants saw one frog, which held a miniature feather (target animal), while in the two-referent context (Fig. 3b), they saw an additional frog without a feather.

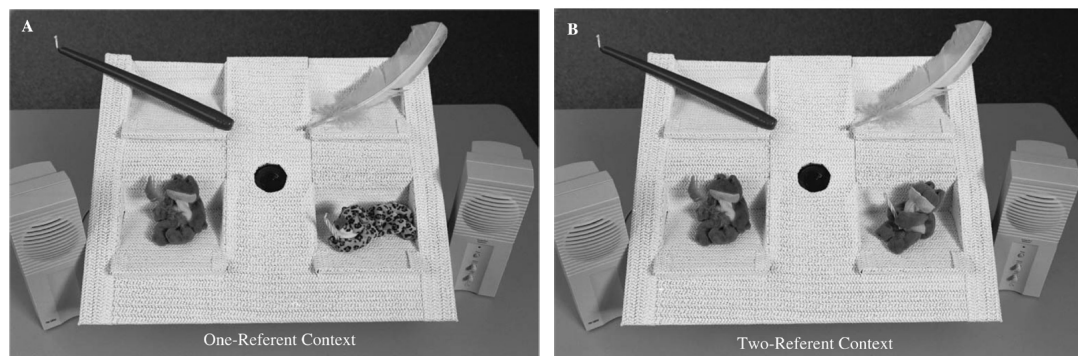a. One-referent context   b. Two-referent context

Fig. 3: Visual contexts in Snedeker and Trueswell (2004).

The results from adults showed that both verb bias and visual context affected structural ambiguity resolution. Their overall number of looks to the instrument (the large feather) was higher the more the verb favoured the instrument interpretation and was higher in the one- than two-referent context. Verb bias had a somewhat earlier effect on eye movements to the instrument (first appearing 200-700 ms after the onset of *frog*) than referential context (which first appeared 700-1200 ms after the onset of *frog*). The proportion of times that adults performed the correct instrument action (using the feather to tickle one of the frogs) was also affected by both factors. Similar to adults, children's overall number of looks to the instrument was higher the more the verb was instrument biased, and the effect of verb bias also affected the proportion

of times children carried out the correct instrument action. In contrast, and consistent with Trueswell et al. (1999), there was no referential context effect in the overall number of looks to the instrument or the proportion of instrument actions, suggesting that context did not affect children's final interpretation.

Children's actions in Snedeker and Trueswell's study show that 5-year-old children use lexical information even if it is inconsistent with the visual context and therefore results in a pragmatically infelicitous sentence. For example, if *with the feather* is interpreted as an instrument, then it is unclear which frog is referred to in Fig. 3b). Children appear unable to use contextual information in such situations even for their final interpretation (see also Kidd & Bavin, 2005; Trueswell et al., 1999). It is only between the ages of 5 and 8 that children start to become sensitive to information in the visual context (Weighall, 2008). Furthermore, Trueswell et al. (1999) showed that children often ignore syntactic constraints in their final interpretation: *on the napkin* is interpreted as the destination even though *into the box* is the real destination in *Put the apple on the napkin into the box*. Finally, Kidd, Stewart, and Serratrice (2011) showed that children rely less on plausibility information than adults, so they often try to use a candle rather than a knife when they hear *Cut the cake with the candle*.

Novick, Trueswell, and Thompson-Schill (2005) argued that children have difficulties revising their initial interpretation because their executive function and inhibitory control are not yet fully developed. As a result, when information later in the sentence disconfirms an initial interpretation, children fail to inhibit the initial interpretation. Choi and Trueswell (2010) obtained results from Korean destination/modifier ambiguities that are consistent with this explanation. In Korean, the temporarily ambiguous PP and the NP that it could potentially modify precede the verb. Therefore, verb bias information in Korean is not available until after the temporarily ambiguous PP, unlike in English. If children's strong reliance on verb information in English is due to the fact that the verb occurs first and children are subsequently unable to revise their verb-based interpretation, then Korean children should be less strongly affected by verb bias. This is indeed what Choi and Trueswell found. Following the sentence-final verb, children looked about equally often at the destination in a two-referent context regardless of whether the verb required a destination phrase (Korean translation of *put*) or did not allow it (Korean translation of *pick up*). Furthermore, even if the verb did not allow a destination phrase, children

frequently carried out a destination action (e.g., putting a frog on a napkin following the Korean equivalent of the sentence *Pick up the frog on the napkin*), clearly indicating that children ignored verb information. This contrasted with adults, who rarely looked at the destination and never carried out a destination action in such cases. Thus, these results are consistent with the idea that children often fail to revise their initial verb-based interpretation in English because the verb occurs before the point of ambiguity.

Interestingly, results by Novick et al. (2008) suggest that even adults sometimes fail to use contextual information when the verb strongly favours the destination interpretation, contrary to what was assumed on the basis of earlier studies. As in earlier studies, they tested sentences such as (5), which contained the verb *put*, which requires a destination PP and therefore strongly biases towards this analysis.

5a. Put the frog on the napkin into the box.
5b. Put the frog that's on the napkin into the box.

Participants saw either a one-referent (Fig. 4a) or two-referent (Fig. 4b) context.
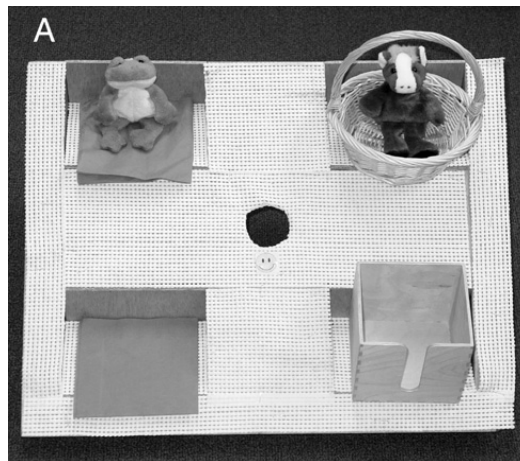
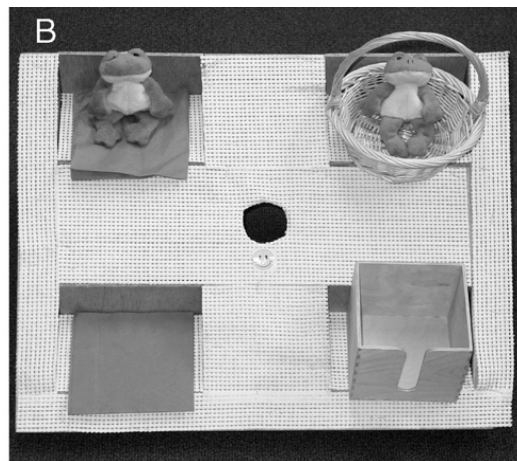a. One-referent context   b. Two-referent context

Fig. 4: Visual contexts in Novick et al. (2008).

These visual contexts were slightly different from earlier studies, in that the competitor was *in* a basket rather than *on* another object, so if *on the napkin* was interpreted as a modifier, then the preposition *on* immediately ruled out reference to the competitor frog (which was *in* a basket).

Consistent with previous studies, in one-referent contexts participants looked more often at the incorrect destination (the empty napkin) in the ambiguous than unambiguous condition shortly following *napkin*, but there was no such difference in two-referent contexts. This might suggest that in the two-referent context, participants did not consider the destination analysis. However, Novick et al. (2008) also analysed looks to the target and competitor frog in the two-referent conditions. If participants interpreted the temporarily ambiguous PP as a modifier of *the frog*, then participants should immediately look at the target frog (which is on a napkin), because the preposition *on* indicates that it cannot be a modifier of the competitor frog which is *in* a basket. But if they interpreted the PP as the destination, then *the frog* remains unmodified, so it is unclear to which frog it refers, and participants should look both at the target and competitor. Indeed, the preference to look at the target frog was less pronounced in the two-referent ambiguous condition (1a) than unambiguous condition (1b). This suggests that in some instances, the temporarily ambiguous PP was initially interpreted as a destination, and therefore, it was unclear which frog was referred to.

Further evidence came from analyses of looks to the correct destination (the box) shortly following *into*. Participants looked less often at the correct destination in the two-referent ambiguous than unambiguous condition, again consistent with the idea that in some instances, they misanalysed *on the napkin* as the destination in the ambiguous condition. Finally, on 8% of trials, participants carried out the action incorrectly in the two-referent ambiguous condition (they put the frog on the empty napkin rather than in the box), again indicating that participants adopted the destination analysis even in the two-referent context.

In sum, Novick et al.'s (2008) study suggests that adults do not always use contextual information during online sentence processing. Strikingly, their actions indicate that they sometimes fail to use both context and grammatical constraints even for the final interpretation of the sentence: They sometimes interpret *on the napkin* in (5a) as the destination, even if this is pragmatically infelicitous (in the two-referent condition, it is unclear which frog is referred to) and even though this is ungrammatical (*into the box* is the correct destination). This is consistent with other findings that adult language comprehenders sometimes fail to reanalyse temporarily ambiguous sentences (e.g., Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Van Gompel, Pickering, Pearson, & Jacob, 2006). All these findings support the idea that comprehenders' final representation of the sentence is often based on non-

syntactic, "good-enough" heuristics and therefore syntactically not fully specified (Ferreira, 2003; Townsend & Bever, 2001).

The use of prosodic cues in syntactic ambiguity resolution

Prosodic information often marks syntactic boundaries in a sentence, so visual-world eye-tracking research has tried to establish whether language comprehenders use it during syntactic ambiguity resolution. Before the introduction of the visual-world method, researchers investigating the use of prosody mainly relied on offline methods, which do not provide a moment-to-moment record of the time course of sentence processing. Snedeker & Trueswell (2003) were the first to use the visual-world method to investigate prosodic effects. One participant (the speaker) was shown an action, which they had to describe to another participant (the listener) whose eye movements to two-referent visual contexts similar to Fig. 3b were monitored while they listened to the speaker's instruction. Speakers either instructed the listener to perform an instrument action (e.g., tapping a frog using a feather) or a modifier action (e.g., tapping a frog that has a feather without using an instrument). Their instructions were generally ambiguous (e.g., *Tap the frog with the feather*), but acoustic and prosodic analyses showed that for instrument instructions, speakers tended to lengthen the direct object noun (*frog*) and paused between this noun and the preposition *with*. They also often put an intonational phrase break after the direct object noun and used a pitch accent for the preposition. For modifier instructions, they tended to lengthen the verb and produced a pause after it. They also tended to shorten the PP and often put an intonational phrase break after the verb.

Listeners used the prosodic information provided by the speaker very rapidly when they heard the ambiguous utterance. When they heard the direct object *the frog* in an instrument instruction, they looked at both frogs equally often, indicating that they interpreted *with the feather* as the instrument, and therefore it was unclear which frog was referred to. But in modifier instructions, they mainly looked at the frog that had the feather, indicating that they analysed *with the feather* as a modifier. Furthermore, prosody affected listeners' fixations to the instrument when hearing the ambiguous PP, with more looks to the instrument in the instrument than modifier instructions.

Snedeker and Yuan (2008) found that children also used prosody during the processing of sentences such as *you can feel the frog with the feather*, but the effect of prosody on looks to the instrument (feather) appeared somewhat later than effects of verb bias. This contrasted with the pattern in adults, where the effect of prosody was earlier and appeared in the same analysis regions as the verb bias effects. Snedeker and Yuan suggested that the difference in time course with children may arise because lexical information becomes available earlier, at the verb, than prosodic information, which may not be a useful cue for the interpretation of the ambiguous PP until the preposition *with* is heard. However, this does not explain why children's use of prosody was delayed relative to when adults used it and why prosody affected adults' eye fixations while they heard the direct object *the frog* in Snedeker and Trueswell (2003). Another possibility is that children initially relied on lexical information and used prosodic information later.

Processing structurally ambiguous sentences: Summary

To summarise visual-world eye-tracking research investigating syntactic ambiguity resolution, most studies suggest that during the processing of structurally ambiguous sentences, adults rapidly integrate various types of information, including information about the visual context, action-based affordances, lexical biases and prosody. Despite the fact that the visual-world method provides a very fine-grained temporal record of sentence processing, there is no clear evidence that adult language comprehenders initially adopt the structurally least complex analysis and subsequently revise this interpretation in the light non-structural information, as claimed by modular sentence processing theories. The findings are more consistent with constraint-satisfaction theories. In contrast, the results from young children suggest that they strongly rely on verb bias information and are also sensitive to prosodic information, but they often fail to use referential context. One possibility, suggested by Novick et al. (2005) is that this is because the verb occurs first in English, and children have difficulty revising their verb-based initial analysis in the light of subsequent information.

Structural effects on anticipatory eye movements

The studies discussed in the previous section indicate that both adults and children process sentence structure incrementally. For example, when they hear *Put the apple on the towel in the box* in a one-referent context, they look at the destination (an empty towel) as soon as they hear *towel*; they do not postpone their structural analysis until the end of the sentence. Further evidence for incrementality comes from a series of studies on referential processing.

Eberhard et al. (1995) gave participants instructions such as (6) while they saw visual displays containing geometrical shapes. In the early-disambiguation display, there was only one plain object (all other objects had stars on them), so the sentence was disambiguated at *plain*. In the intermediate disambiguation display, all objects were plain, but only one object was red, so the word *red* disambiguated, while in the late-disambiguation display, the word *square* provided the disambiguating information. In all conditions, there was only one object that was both plain, red and a square. Participants fixated this target object faster the earlier the disambiguation occurred. Importantly, target fixations generally occurred before participants heard the noun, indicating that they semantically interpreted the NP with respect to the visual context before they heard the syntactic head of the NP structure.

Sedivy et al. (1999; Sedivy, 2003) showed that listeners also rapidly establish referential contrasts. When participants heard *Touch the tall glass*, they fixated the target glass faster when there was another, shorter glass in the display than when the distractor object was not a glass. Eye fixations indicated that they used the visual contrast information before they heard the head noun. Furthermore, the taller the glass, the faster they fixated it, indicating that they incrementally interpreted the adjective *tall* relative to the typical properties of the head noun. Again, analyses of fixation behaviour indicated that they did this before they heard the noun.

Altmann and Kamide (1999) argued that sentence processing is not just incremental, but that language comprehenders project upcoming argument roles even before the argument is heard. Participants listened to sentences such as (7) while their eye movements to scenes such as Fig. 5 were monitored.

7a. The boy will eat the cake.
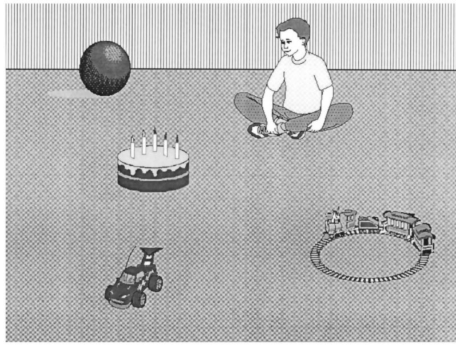7b. The boy will move the cake.

Fig. 5: Scene in Altmann and Kamide (1999).

Before they heard *cake*, participants looked more often at the cake in (7a) than (7b). Altmann and Kamide suggested that listeners predict that *eat* will be followed by an edible noun and therefore make *anticipatory eye movements* to objects that are edible. In contrast, *move* can be followed by any object in the scene, so listeners do not anticipate edible objects any more than other movable objects.

Various factors appear to affect anticipatory eye movements. Kamide, Altmann, and Haywood (2003) found that anticipatory eye movements are not just affected by how well an object satisfies the semantic constraints of the verb, but also by how well it satisfies combined constraints of both the verb and the subject. Participants tended to look at a carousel when they heard *The girl will ride …*, whereas they looked at a motorbike when they heard *The man will ride … .*

As discussed in more detail by Knoeferle (this volume), Knoeferle, Crocker, Scheepers, and Pickering (2005) showed that depicted actions influence anticipatory eye movements in German sentences that are ambiguous between agent-verb-patient and patient-verb-agent order: If the depicted action suggested that the NP after the verb was a patient, participants tended to look at a potential patient shortly after the verb, before they heard the NP, but when the action suggested that the NP was an agent, they tended to make anticipatory eye movements to the agent. Knoeferle and Crocker (2007) showed that depicted actions can even override semantic constraints that determine how plausible an object (or person) is as an agent or patient. Prosody also appears to affect anticipatory eye movements in this type of ambiguity: Weber, Grice, and Crocker (2006) showed that before participants heard the postverbal NP, they fixated a potential patient when sentence stress was on the verb, but fixated a potential agent when stress was on the NP preceding the verb. This suggests that participants analysed the first NP as an agent when stress was on the verb (and

15

therefore anticipated a patient following the verb), whereas they analysed the first NP as a patient when stress was on this NP (and therefore anticipated an agent).

Given the focus of this chapter, the question of most interest is whether anticipatory eye movements are influenced by syntactic factors. Results by Boland (2005) indicate that anticipatory looks to an object are not only affected by how compatible it is with the semantics of the verb, but also by the syntactic status of the upcoming phrase in which it may be mentioned. In one of her experiments, she tested sentences such as (8).

8a. One window was broken, so the handyman mentioned it right away to the owners.
8b. One window was broken, so the handyman fixed it hurriedly for the owners.

Corpus counts showed that *owners* follows *mentioned* and *fixed* equally frequently. However, *owners* is a syntactic argument of the verb *mentioned*, whereas it is an adjunct in the sentence with *fixed*. The results showed that argument status of an upcoming phrase affected participants' anticipatory eye movements: They were more likely to fixate a picture of a group of owners shortly after hearing the verb *mentioned* (8a) than *fixed* (8b).

Anticipatory looks are also influenced by syntactic information preceding the verb. In a study by Kamide, Scheepers, and Altmann (2003), German participants listened to sentences such as (9) while they saw scenes containing a hare, a fox, and a cabbage.

9a. Der Hase frißt gleich den Kohl.
   The hare (subject) eats shortly the cabbage (object).
   "The hare will shortly eat the cabbage."
9b. Den Hase frißt gleich der Fuchs.
   The hare (object) eats shortly the fox (subject).
   "The fox will shortly eat the hare."

In (9a), syntactic case marking on the article indicates that *Hase* is the subject, whereas in (9b), it indicates that it is the object. This affected anticipatory eye movements: During the presentation of *gleich* ("shortly") participants looked more at

the fox in (9b) than (9a), whereas they looked slightly more at the cabbage in (10a) than (9b).

Similar effects were observed with English active and passive structures (10).

10a. The hare will eat the cabbage.
10b. The hare will be eaten by the fox.

During the presentation of the verb *eat/be eaten by*, participants looked more at the fox in (10b) than (10a), whereas they looked slightly more at the cabbage in (10a) than (10b). Thus, the results from German show that listeners use syntactic case marking when making anticipatory eye movements, while the results from English indicate that they use information about whether the sentence is active or passive.

Kamide, Altmann, and Haywood (2003) investigated whether listeners make anticipatory eye movements before they hear the verb. The verb is the syntactic head of the sentence on which other grammatical elements of the sentence such as the subject and object depend, so anticipations may occur when listeners hear the verb, but not before it. To examine this possibility, Kamide et al. investigated Japanese sentences such as (11).

11a. Weitoresu-ga kyaku-ni tanosigeni hanbaagaa-o hakobu.
     Waitress-nom customer-dat merrily hamburger-acc bring.
     "The waitress will merrily bring the hamburger to the customer."
11b. Weitoresu-ga kyaku-o tanosigeni karakau.
     Waitress-nom customer-acc merrily tease.
     "The waitress will merrily tease the customer."

In (11a), the dative case marked NP kyaku-ni "customer-dat" strongly suggests that the customer is a recipient, and therefore another NP is needed that indicates what the transferred object (or theme) is. Because Japanese is verb final, this NP should precede the verb. In contrast, in (11b), the accusative case marked NP *kyaku-o* "customer-acc" makes it unlikely that it is followed by another NP, because *kyaku-o* can be analysed as the object of a monotransitive verb. The results indeed showed that listeners anticipated an upcoming theme: During the presentation of *tanosigeni* "merrily", participants fixated a picture of plausible theme (a hamburger) more often

17

in (11a) than (11b). Kamide et al. concluded that listeners do not postpone syntactic analysis until the head of the sentence, but anticipate upcoming arguments before they hear the verb.

The results from Kamide, Scheepers and Altmann (2003) and Kamide, Altmann, and Haywood (2003) also show that fixations to unmentioned objects are not just due to semantic priming by the verb or the subject, as argued by Kukona et al. (2011). Kukona et al. showed that when participants heard (12), they not only fixated a picture of a criminal, which is likely to be mentioned after *arrested*, but also frequently fixated a picture of a policeman, even though a policeman is not a plausible object.

12. Joe arrested the criminal.

They argued that participants looked at both the criminal and the policeman because both are semantically primed by the verb *arrested*. Participants looked only slightly more often at the criminal than the policeman, suggesting that anticipation of upcoming information only had a weak effect. However, the results by Kamide and colleagues cannot be due to semantic priming: In (9a) and (9b), the first NP and the verb are identical, and in (11a) and (11b) the first two NPs are identical. Together, the studies by Kukona et al. and Kamide and colleagues therefore suggest that both linguistic anticipation and semantic priming affect fixations to unmentioned objects.

Structural priming effects on comprehension

Research has also used anticipatory eye movements to investigate structural priming, that is, how the structure of one sentence affects the processing of a subsequent sentence. Although there is much evidence that structural priming affects language production processes (Pickering & Ferreira, 2008), until recently there was only limited evidence that it affects comprehension of sentences. Recent findings from both the visual-world method and other methods (Branigan, Pickering, & McLean, 2005; Ledoux, Traxler, Swaab, 2007; Tooley, Traxler, & Swaab; 2009; Traxler, & Tooley, 2008) have changed this.

In a study by Scheepers and Crocker (2004), German participants first read aloud one of the prime sentences in (13).

13a. Der Regisseur lobte insbesondere den Produzenten.

The director (subject) commended in particular the producer (object).

13b. Den Regisseur lobte insbesondere der Produzent.

The director (object) commended in particular the producer (subject).

In (13a), the first NP (*Der Regisseur*) is the subject and agent, whereas the second NP (*den Produzenten*) is the object and patient. In (13b), the grammatical and thematic roles are reversed. Thus, (13a) and (13b) are semantically very similar, but differ structurally. Next, participants listened to one of the target sentences in (14) while they saw a picture consisting of a sportsman pushing a nurse who was blow-drying a priest. Eye-movements to these three characters were analysed.

14a. Die Krankenschwester föhnt offensichtlich den Priester.

The nurse (ambiguous) blow-dries apparently the priest (object).

14b. Die Krankenschwester schubst offensichtlich der Sportler.

The nurse (ambiguous) pushes apparently the sportsman (subject).

The sentences in (14) are temporarily ambiguous, because case marking of the first NP (*Die Krankenschwester*) does not disambiguate it towards either a subject or object. The first point of disambiguation is at the verb: In (14a), the verb in combination with the picture indicates that *Die Krankenschwester* "the nurse" is the subject of the blow-drying action. In (14b), she is the object of the pushing action.

The structure of the prime affected anticipatory eye movements to the priest and sportsman while participants heard *Die Krankenschwester*. After subject-verb-object primes, participants looked longer at the priest, the patient in the picture (being blow-dried by the nurse), than the sportsman, the agent in the picture (pushing the nurse). This suggests that they anticipated that the priest was going to be mentioned next. In contrast, after object-verb-subject sentences, they looked slightly longer at the sportsman than the priest. Thus, participants used information about the structure of the prime in their anticipations of whether the postverbal NP was an agent or a patient.

Arai, Van Gompel, and Scheepers (2007) conducted a study to compare structural priming effects in comprehension more directly with previous results from production studies. They used ditransitive structures such as (15) and (16), which

have also been extensively investigated in production (e.g., Bock, 1986; Bock & Loebell, 1990; Pickering & Branigan, 1998).

15a. The assassin will send the parcel to the dictator.
15b. The assassin will send the dictator the parcel.

16a. The pirate will send the necklace to the princess.
16b. The pirate will send the princess the necklace.

Participants first read aloud either a prepositional object (PO) prime sentence (15a) or double object (DO) prime (15b) and then listened to a PO (16a) or DO (16b) target sentence while they saw scenes such as Fig. 6.
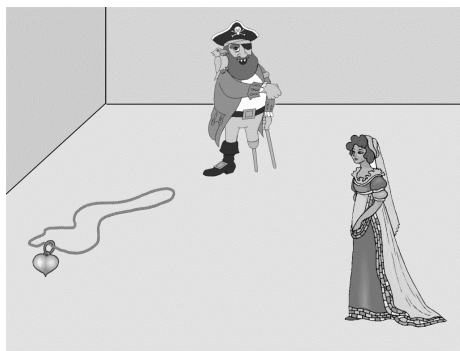

Fig. 6: Scene in Arai et al. (2007).

In their first experiment, Arai et al. used the same verb in prime and target.

Shortly following the verb onset and before the first postverbal noun, participants looked more often at the necklace when the prime was a PO than DO structure, whereas they looked more at the princess after a DO than PO structure. This indicates that participants used information from the prime structure to anticipate whether the first noun following the verb was a theme or recipient. However, in a second experiment, Arai et al. found no evidence for priming when the verb in prime and target was different. This contrasts with results from production (e.g., Pickering & Branigan, 1998), which show that priming occurs when the prime and target verb are different, though priming is stronger when the verb is the same. Thus, structural priming in comprehension occurs with the same ditransitive structures as in production, but the effect appears to be more strongly lexically driven. Arai et al.

suggested that this may be because in comprehension, people process the ditransitive verb before the recipient and theme NPs, whereas in production, people may often determine the order of the recipient and theme before they access a specific verb (e.g., *give* rather than *hand*). The absence of structural priming when the verb is different in prime and target is consistent with findings from reading studies investigating syntactic ambiguity resolution (Branigan et al., 2005; Ledoux et al., 2007; Tooley et al., 2009; Traxler, & Tooley, 2008).

Results by Carminati, Van Gompel, Scheepers and Arai (2009) indicate that these priming effects are due to anticipation of the argument role of the first postverbal NP rather than due to anticipation of semantic properties of the noun. When both the recipient and theme were animate (17a, b), priming was equally strong as when the recipient was animate and the theme was inanimate (17c, d).

18a. The general will send the messenger to the king.
18b. The general will send the king the messenger.
18c. The general will send the telegram to the king.
18d. The general will send the king the telegram.

This suggests that comprehenders use information from the prime to anticipate whether the first postverbal NP is an indirect object recipient or direct object theme rather than whether it is animate or inanimate.

Carminati and Van Gompel (2010) showed that structural priming is not just a short-lasting effect: The priming effects that Arai et al. (2007) and Carminati et al. (2009) showed also occurred when there were two intervening unrelated sentences between prime and target, and priming was not significantly reduced in such cases. The finding that structural priming effects on comprehension persist across intervening sentences suggests that accumulated, repeated exposure to a specific prime structure results in verb bias effects such as observed by Snedeker and Trueswell (2004) and in reading time studies (e.g., Trueswell, Tanenhaus, & Kello, 1993).

While Arai et al. (2007) and Carminati et al. (2009) investigated structural priming in cases where the verb was the same in prime and target, Thothathiri and Snedeker (2008a) priming in the absence of verb repetition using a method in which participants had to act out ditransitive target instructions following a prime sentence.

For example, one prerecorded voice told a short story in which (18) was the last sentence, and then a second voice said *It's my turn. Are you ready?* before giving the instruction in (19).

18a. Then I read a story to the boy.
18b. Then I read the boy a story.

19a. Now you can send the horse the book.
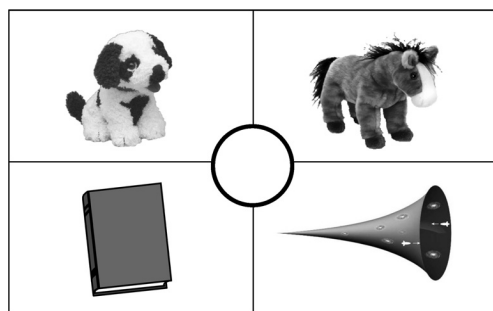19b. Now you can send the horn to the dog.

Fig. 7: Visual display in Thothathiri and Snedeker (2008a).

Participants acted out the instruction using the objects in Fig. 7. The beginning of the first postverbal noun in the target sentence was ambiguous (*hor…*), which permitted analysis of looks to the recipient (*horse*) and theme (*horn*) that occurred during the ambiguous part of the postverbal noun. These analyses showed that the preference to look at the recipient rather than the theme was stronger after double object (18a) than prepositional object (18b) primes. Thus, structural priming occurred across two intervening sentences, as in Carminati and Van Gompel (2010), but in contrast to Arai et al. (2007), priming occurred even though the verb was different in prime and target.

Thothathiri and Snedeker (2008b) observed similar results with both 3 and 4-year-old children. They investigated priming both when the verb was the same in prime and target, and when it was different, but priming was not significantly affected by verb repetition. They concluded that children use abstract, lexically-independent structural representations during sentence comprehension rather than rely on lexically specific representations associated with specific verbs.

The reason why Thothathiri and Snedeker (2008a, b) observed lexically-independent priming effects, whereas Arai et al. (2007) did not is most likely due to

differences in methodology. One possibility is that lexically-dependent and independent priming have a different time course. Arai et al. analysed looks from the verb onset, whereas Thothathiri and Snedeker analysed looks from the first postverbal noun onset, so it is possible that lexically-dependent priming exerts earlier effects, at the verb, than lexically-independent priming. Priming in Thothathiri and Snedeker's experiments may in fact have arisen during reanalysis. On many trials, participants may initially have anticipated the dog or the book in Fig. 7 at the verb, and then had to revise this analysis when they heard *hor...* . They may have used lexically-independent information from the prime during this reanalysis. Carminati et al. (2008) suggested a different explanation: The act-out task that Thothathiri and Snedeker used involved picking up an object (the theme) and moving it towards another object (the recipient), regardless of the specific verb that was used. Thus, participants may first have determined which object should be moved and where it should go before they determined the action indicated by the verb (e.g., giving or throwing). By prioritizing information relevant to the act-out task, participants may initially have determined the argument structure of the sentence independently of the target verb, resulting in lexically-independent priming. This explanation implies that the presence or absence of lexically-dependent priming depends on the listeners' task.

In sum, visual-world studies have demonstrated that syntactic factors play an important role in anticipatory eye-movement behaviour. In particular, listeners use information about case marking and active/passive voice when making anticipatory eye movements, and these eye movements are also affected by whether the upcoming NP is an argument or adjunct of the verb. Research on anticipatory eye movements has also provided evidence for structural priming in structures that are similar to those used in production, and these studies are beginning to shed light on the circumstances under which structural persistence occurs.

Visual-world eye tracking as a method for investigating referential processing

A striking feature of the visual-world paradigm is its sensitivity to reference. The studies discussed in the previous sections have shown that people immediately look at entities that are mentioned and also anticipate the upcoming referent. It may not come as a big surprise, then, that the paradigm has recently been applied to investigating some of the central issues in the processing of referential expressions, such as

pronouns, demonstratives, and reflexives. The focus has been on the role of structural heuristics in the processing of ambiguous pronouns and their relation to other non-structural factors, and on the role of binding constraints in the processing of pronouns and reflexives.

The role of structural heuristics in adults

Prior research using methods other than the visual-world method has shown that many factors affect pronoun resolution at some stage during the comprehension process, such as sentence and discourse focus, semantics, gender, and number (see Garnham, 2001 for an overview). Among the factors that have received considerable attention in ambiguous pronoun resolution are structurally-related heuristic strategies used by the comprehender to decide on the preferred antecedent. The two most prominent factors affecting this choice are the syntactic role of the antecedent and the position it occupies in the sentence. Many researchers have argued that the preferred antecedent of an ambiguous pronoun is the grammatical subject of the preceding clause or sentence (e.g., Frederiksen, 1981; Crawley, Stevenson, & Kleinman, 1990; Gordon, Grosz, & Gilliom, 1993). In contrast, others have argued that the first-mentioned noun phrase of the previous clause or sentence is the preferred antecedent, regardless of its grammatical role (Carreiras, Gernsbacher, & Villa, 1995; Gernsbacher & Hargreaves, 1988; Gernsbacher, Hargreaves, & Beeman, 1989).

The first study to use the visual-world method to examine pronoun resolution, Arnold, Eisenband, Brown-Schmidt and Trueswell (2000), investigated the role of gender information (whether the pronoun's gender is consistent with one or two potential antecedents) and order-of-mention/grammatical role (whether the preferred antecedent is the 1st mentioned subject or 2nd mentioned object) in English. Participants were presented with auditory texts such as (20a,b) and their eye movements were recorded time locked to the pronoun *he* while they looked at visual scenes showing a picture of Donald Duck and Mickey/Minnie Mouse (Fig. 8).

20a. Donald is bringing some mail to Mickey, while a violent storm is beginning. He's carrying an umbrella.
20b. Donald is bringing some mail to Minnie, while a violent storm is beginning. He/She's carrying an umbrella.
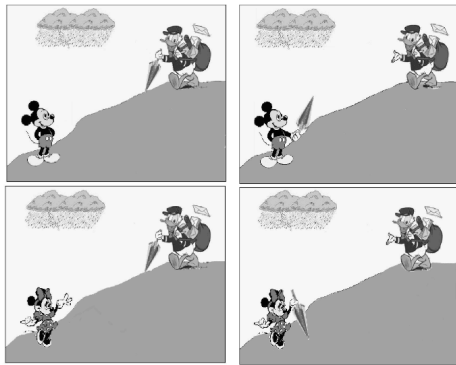
Fig. 8: Scenes used in Arnold et al. (2000).

When gender information disambiguated the pronoun (20b), participants' eye movements were directed towards the gender-matching referent from 200 ms after the onset of the verb (*carrying*), both when the pronoun referred to the first (*he*) and the second-mentioned character (*she*). Interestingly, when gender did not disambiguate the pronoun (20a), participants used order-of-mention as a cue: When the meaning of the verb (*carrying*) was consistent with the picture of the first-mentioned character (*Donald*), they looked more often at this character than the second-mentioned character (*Mickey*), starting 200 ms after the verb onset, whereas they looked equally often at both characters when the verb was consistent with the second-mentioned character. A second experiment showed that when the first-mentioned character was made even more salient by pronominal reference to it (*Donald is bringing some mail to Mickey. He's sauntering down the hill while a violent storm is beginning*), participants looked more often at the first-mentioned character even if the verb was inconsistent with it, starting 400 ms after verb onset.

Arnold et al.'s (2000) results suggest that in the absence of other cues, participants preferentially interpret a pronoun as coreferent with the first-mentioned character. However, because the first-mentioned character was also the subject, as is often the case in studies conducted in English (e.g., Garnham, Traxler, Oakhill, & Gernsbacher, 1996; Gernsbacher, 1989; McDonald & MacWhinney, 1995), it is unclear whether it was first-mention or subjecthood that affected the participants' eye movements. Therefore, some recent studies using the visual-world method have turned to languages with more flexible word order to investigate this question.

Järvikivi, Van Gompel, Hyönä and Bertram (2005) contrasted first-mention and subjecthood in Finnish, a language that allows manipulation of word order without changing the morphosyntactic marking of the subject and object. Finnish is a flexible word-order language with a gender-neutral 3rd person singular pronoun *hän* 'he/she' referring to both male and female referents. Grammatical roles are indicated through morphosyntactic marking, with the subject typically in nominative (nom) and the object in partitive (ptv), as shown in (21).

21a. Tony Blair kätteli George Bushia valkoisessa talossa. Hän halusi keskustella Irakin tilanteesta. (Tony Blair (`nom-sub`) shook hands with George Bush (`ptv-obj`) in the White House. He wanted to discuss the situation in Iraq.)
21b. George Bushia kätteli Tony Blair valkoisessa talossa. Hän halusi keskustella Irakin tilanteesta. (George Bush (`ptv-obj`) shook hands with Tony Blair (`nom-sub`) in the White House. He wanted to discuss the situation in Iraq.)

In contrast to the subject-verb-object (SVO) order in (21a), the order of the subject and object can be reversed (OVS order) without any change in the inflectional marking of the subject, object or verb, as shown in (21b).

Järvikivi et al. (2005) presented participants with spoken sentences such as (21) above and pictures presenting the two characters (e.g., Bush and Blair), and the location mentioned in the prepositional phrase (e.g., the White House). The location was mentioned to draw the participants' eyes away from the critical characters at the onset of the pronoun. The study found two main effects: An early advantage for subject over object antecedents (starting 480 ms after the pronoun onset) and a somewhat later effect of first-mention that appeared 690ms after pronoun onset. Based on the main effects in the absence of interactions the authors concluded that both grammatical role and order of mention had independent effects on ambiguous pronoun resolution in Finnish, the subjecthood effect becoming significant earlier than first-mention in the eye movement record.

Ellert and colleagues (Ellert, 2010; See Ellert, Roberts, & Järvikivi, 2011) tested whether there is a first-mention preference in cases where the first- and second-mentioned referent have the same grammatical role marking. They investigated the resolution of German and Dutch ambiguous masculine pronouns (*er* 'he/it', *hij* 'he/it')

or d-pronouns/demonstratives (*der* 'this', *die* 'this') using sentences in both languages that consisted of comparative structures with animate or inanimate entities followed by a sentence beginning with an ambiguous personal or d-pronoun (22). They presented scenes such as Fig. 9.

22. Der Schrank ist schwerer als der Tisch. Er/Der stammt aus einem Möbelgeschäft in Belgien.
'The cupboard is heavier than the table. It originates from a furniture store in Belgium.'
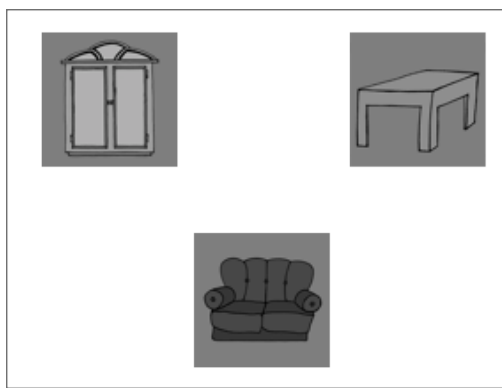

Fig. 9: Example scene from Ellert (2010).

Ellert observed a strong first-mention preference for personal pronouns starting 400ms after the pronoun onset in Dutch and slightly later in German, whereas she found a second-mention preference for d-pronouns, which appeared slightly later in both languages. This suggests that in the absence of grammatical role information, personal pronouns in both German and Dutch follow the first-mention principle (cf. Bouma & Hopp, 2007, who found no clear evidence for order-of-mention in German in an offline task using nominative marked subjects and accusative and dative marked direct and indirect objects). The results also indicate that the order-of-mention preference is dependent on the type of pronoun (personal vs. d-pronoun).

Kaiser and Trueswell (2008) argued that different anaphoric forms do not just have different preferences, as observed by Ellert and colleagues, but are also sensitive to different saliency factors (see also Brown-Schmidt, Byron, & Tanenhaus, 2005). They investigated the resolution of the Finnish personal pronoun *hän* and contrasted it with the demonstrative *tämä* 'this', which can be used to refer to animate entities

including humans. They presented a discourse context (23) followed by an SVO (24a) or OVS (24b) antecedent sentence, which was in turn followed by a sentence starting with the critical pronoun *hän* or *tämä* (25).

23. Liisa astuu sisään erään suuren firman päätoimistoon. Hän huomaa sihteerin, joka puhuu puhelimessa.
'Liisa steps into the main office of a big company. She notices a secretary who is talking on the phone.'
24a. Hetken päästä sihteeri moittii juuri sisään tullutta liikemiestä samalla kun printterit tulostavat päivän raportteja.'
'After a moment the secretary (`nom-sub`) criticises a businessman (`ptv-obj`) who has just walked in while the printers are churning out the day's reports.'
24b. Hetken päästä sihteeriä moittii juuri sisään tullut liikemies samalla kun printterit tulostavat päivän raportteja. '
'After a moment the secretary (`ptv-obj`) criticises a businessman (`nom-sub`) who has just walked in while the printers are churning out the day's reports.'
25. Hän/Tämä seisoo valokopiokoneen lähellä.
'(S)he/This is standing near a photocopier.'

Sentence completions following (24) showed a subject preference for *hän* independent of word order, whereas *tämä* showed a second mention preference, which was somewhat stronger following SVO than OVS order. Their eye movement study showed that the pronoun *hän* was mainly influenced by grammatical role starting about 400 ms from the pronoun onset, although the subject preference in OVS appeared later and was not as pronounced as for SVO. For *tämä*, the pattern was less straightforward with an initial first-mention/subject preference, which later changed into a second-mention preference. Overall, the results suggest that Finnish *tämä* and *hän* are sensitive to different cues, consistent with Kaiser and Trueswell's (2008) claim that different referential forms are sensitive to different salience factors and can therefore be influenced by syntactic, semantic and information structure in different ways.

Although the results of Kaiser and Trueswell (2008) and Järvikivi et al. (2005) are generally in line, it is possible that differences in the stimuli contributed to the slight differences in the results. On the one hand, it may be that isolated presentation

28

of the OVS sentences in Järvikivi et al. (2005) drew attention to the first-mentioned object, resulting in an effect of first-mention with *hän* that was not observed in Kaiser and Trueswell (2008). On the other hand, the context in Kaiser and Trueswell (2008) may have put the subject of the OVS in focus (Vilkuna, 1989), because it presents the subjects as new information, and may therefore have highlighted the preference for the subject.

In sum, these studies show that structural information, both order-of-mention and the grammatical role of the antecedents in the previous context influence pronoun resolution in various languages. Furthermore, properties of the different anaphoric expressions interact with structural cues in different ways, such that personal and demonstrative pronouns are sensitive to various sources of information to a different extent. This is in line with theories that predict that different referring expressions are cues for the listener as to the extent to which the intended antecedent is accessible in the current discourse (e.g., Ariel, 1988; 2001), and theories that predict a link between the type of referring expression and the cognitive status, givenness, of the antecedent (e.g., Gundel, Hedberg, & Zacharski, 1993).

However, a study by Järvikivi, Van Gompel and Hyönä (2015) suggests that the subject preference in the above studies may at least partly be due to the fact that the subject and object had different semantic roles. Järvikivi et al. counterbalanced the semantic roles of the verb by using Finnish stimulus-experiencer verbs such as *pelkäsi* ('feared') and experiencer-stimulus verbs such as *pelotti* ('frightened') that share the same stem (26). Participants saw scenes such as Fig. 10.

26. Vladimir Putin pelkäsi/pelotti George Bushia Valkoisessa talossa, koska hän oli kuluneen viikon aikana antanut useaan otteeseen ymmärtää, ettei maiden Irakin suhteissa olisi näkemyseroja.
'Vladimir Putin (`nom-sub`) feared/frightened George Bush (`ptv-obj`) at the White House, because he had during the past week given many times the impression that there would be no differences of opinion concerning the countries' relations with Iraq.'
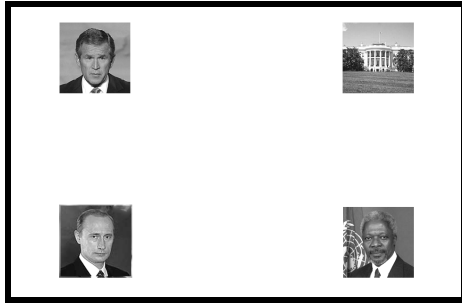
Fig. 10: Scene used in Järvikivi et al. (2015).

Research has shown that people tend to attribute causality to the stimulus role, and therefore, they have a preference to assign pronouns following the causal conjunction *because* to the stimulus (e.g., Garnham et al., 1996; Koornneef & Van Berkum, 2006; Stewart, Pickering, & Sanford, 2000). This preference is generally referred to as the *implicit causality bias*. Järvikivi et al. found that shortly following the pronoun *hän*, listeners looked more often at a picture of the stimulus than the experiencer, and there was no evidence that this implicit causality bias was delayed relative to the first-mention effect that they found in the same experiment. Importantly, there was no overall grammatical role preference. Thus, this experiment shows that semantics has an early effect on pronoun resolution, contra the integration hypothesis (Garnham et al., 1996; Stewart et al., 2000), which claims that implicit causality effects should appear late, when the second clause is integrated with the first. This conclusion is further supported by a similar visual-world study in Dutch (Cozijn, Commandeur, Vonk, & Noordman, 2011), which also showed very early implicit causality effects. Furthermore, Järvikivi et al.'s experiment showed that when grammatical role was counterbalanced, there was no subject preference, suggesting that the subject preference in many previous studies may have been in large part semantic in origin.

In a second experiment, Järvikivi et al. investigated whether structural first-mention and semantic verbs bias information affected the interpretation of personal (*hän*) and demonstrative (*tämä*) pronouns differently. The results indicated that both pronouns were similarly affected by verb bias, so that people looked more at the stimulus than the experiencer, but differed with respect to their preference for structural information, with *hän* preferring the first-mentioned subject and *tämä* the second-mentioned object, in line with Kaiser and Trueswell (2008). This suggests that

anaphor type interacts with structural information but not with semantic implicit causality information.

Pyykkönen and Järvikivi (2010) found that implicit causality information affects the activation of discourse entities even before listeners hear the pronoun and the conjunction (*koska* 'because'). Participants heard sentences containing either stimulus-experiencer or experiencer-stimulus verbs such as in (27) as part of a short story while they saw pictures of the subject (butler) and object (guitarist) antecedents and two additional pictures of entities or events in the story.

27. Hovimestari pelkäsi/pelotti kitaristia ravintolasalissa, koska koko päivän hän kummallista kyllä oli näyttänyt erittäin tyytymättömältä.
'The butler feared/frightened the guitarist in the dining room, because for the whole day he curiously enough had seemed extremely unhappy.'

They looked more frequently at the entity consistent with implicit causality information following the onset of the object in the first clause (*kitaristia*, 'guitarist'), indicating that implicit causality affected activation even before the pronoun (see Cozijn et al., 2011 for similar evidence from Dutch). At the pronoun, both semantic and structural constraints had an effect: In addition to a preference for the first-mentioned subject, the authors observed a significant interaction between verb semantics and grammatical role such that implicit causality information modulated the subject antecedents but not the object antecedents. They also found some evidence for a syntactic role parallelism effect (e.g., Smyth, 1994) during later processing: When the second clause contained an object pronoun (*hänet*), there were more fixations to object pictures than when it contained a subject pronoun (*hän*, as in (27)). In sum, semantic information increased the salience of the antecedent that was already prominent, namely the first-mentioned subject. Furthermore, pronouns with different grammatical roles are differently affected by the grammatical roles of the potential antecedents, but similarly affected by implicit causality information.

The role of structural heuristics in children

Recent visual-world experiments have also shed light on how children interpret pronouns. In general, they suggest that children may be affected by similar structural heuristics as adults.

Song and Fisher (2005) carried out a series of visual-world experiments investigating 3-year-old children's comprehension of English ambiguous pronouns. In their first visual-world experiment, they presented short stories together with pictures as shown in Fig. 11. The stories consisted of three context sentences that made either one of the other of two characters (the turtle or the tiger) more prominent by first mention, subject status, and pronominalization, followed by the critical sentence *now what does he have?* that contained the ambiguous pronoun *he*. Each sentence was presented together with two pictures; the question of interest was whether children would look at the picture with the tiger or turtle following the ambiguous pronoun in the critical sentence. Song and Fisher found that children looked at the more prominent character, the repeated first-mentioned subject, more often than the less prominent character, the second-mentioned object, between 3 and 4 seconds from the onset of the pronoun. In two further experiments, the pronoun in the third context sentence was replaced with a full noun (*And the turtle finds a box with the tiger*) or this sentence was removed altogether. The findings were similar in that the children preferred the most prominent character, except that this preference occurred earlier in the experiment where the third sentence was removed, perhaps because it reduced the complexity of the story.

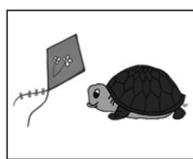See the turtle and the tiger./See the tiger and the turtle.
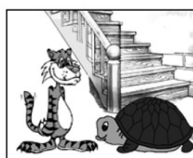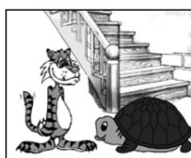
The turtle goes downstairs with the tiger./The tiger goes downstairs with the turtle.
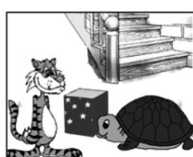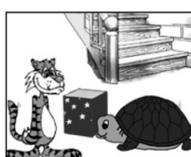
And he finds a box with the tiger./And he finds a box with the turtle.

Now what does he have? Look, he has a kite!

Fig. 11: Scenes and stories used in Song and Fisher (2005).

Arnold, Brown-Schmidt and Trueswell (2007; see also Arnold, Brown-Schmidt, Trueswell, & Fagano, 2005) investigated English 3-, 4-, and 5-five-year-olds' processing of ambiguous pronouns. An experimenter announced a story about two puppets (*this is a story about Froggy and Puppy*), placed the puppets on the table and named them, after which the child heard the story. The children were presented with stories such as (28), which were told by a puppet, Elmo. The authors manipulated the gender of the two puppets, whether the gender was different (28a, b) or the same (28c); and whether the referent of the pronoun was first (28a) or second (28b) mentioned. The children's task was to place the intended object (e.g., toy carton of milk) in front of the preferred puppet.

28a. Puppy is having lunch with Froggy. He wants some milk.
28b. Puppy is having lunch with Froggy. She wants some milk.
28c. Puppy is having lunch with Panda Bear. He wants some milk.

In their first experiment, the authors found that in the different-gender conditions (28a, b), the older children (4-5 years) behaved adult-like in that they almost invariably put the object in front of the gender-matching puppet regardless of the order-of-mention of the characters, whereas the younger children (3-4 years) were above chance in the different gender first-mention condition, but only marginally above chance if the gender-matching target was second-mentioned. Eye movements to the puppets showed similar results: Order of mention did not clearly affect the younger children's eye movements, whereas the older children tended to look at the gender-matching puppet from about 800 ms after pronoun onset, somewhat later than adults do. In the same-gender condition (28c), neither age group showed a first-mention preference in their actions, but the older children tended to fixate the puppet

that they chose for their action. Experiment 2 in Arnold et al. (2007) used the same materials as Arnold et al. (2000) (see example 20 and Fig. 8) and tested 5-year-old children. Five-year-olds looked at the gender-matching referent as quickly as adults (starting 400ms from the onset of the pronoun), but in line with the results of Experiment 1, there was no reliable evidence for a first-mention preference in the same gender condition. However, children had less than a second to show such an effect before disambiguating information at the verb became available. In the Song and Fisher studies, effects with gender-ambiguous pronouns appeared much later, between 3 and 4 seconds from the onset of the pronoun, and the first-mentioned referent was in fact made prominent by several means. Therefore, it is possible that with children, this preference can be seen only very late, which is possibly why Arnold et al. (2007), who only measured earlier eye movements, did not observe it. Recent results by Hartshorne, Nappa and Snedeker (2011) provide support for this. Like Arnold et al. (2005, 2007), Hartshorne et al. studied pronoun resolution in English speaking 5-year-olds. Their eye movement results confirmed that children resolved the pronoun towards the first-mentioned antecedent, but the effects did not appear until 1200-1400 after pronoun onset.

This also is in line with recent findings by Pyykkönen, Matthews and Järvikivi (2010) on English 3-year-old children's pronoun resolution. They manipulated semantic prominence by using stories such as in Fig. 12, in which the first sentence either had a high or low transitive verb (*hit* vs. *tease*). The degree of transitivity reflected whether the subject and object argument of the verb had more prototypical agent- and patient-like properties, such as volition, sentience, causation (subjects) and affectedness (objects), after Dowty (1991; see also Kako 2006). The third, critical sentence was always the same (*He did something very silly*) and contained the ambiguous pronoun *he*. After about 4000ms from the pronoun onset, the last sentence and disambiguating picture appeared. In Pyykkönen and colleagues' study the preference for first-mentioned subjects was found in the eye movement record 1200 ms after the onset of *he* in the third sentence, in line with the previous studies showing a late effect of structural prominence for young children in English. Importantly, however, children's attention to the discourse participants was rapidly modulated by semantic prominence in terms of the degree of verb transitivity: The effect of verb type appeared within the 200-720 ms window (about 400ms from the pronoun onset), showing more looks to both characters following high than low transitive verbs.

34

Importantly, however, the results showed that semantic and structural prominence interacted: The first-mention effect was more pronounced for low than high transitive verbs. The results suggest that 3-year-olds are already sensitive to the degree of semantic prominence of transitive verbs and indicate that semantic prominence is an important component of antecedent salience (see Järvikivi et al. submitted; Rose, 2005).

HIGH TRANSITIVITY CONDITION

The panda hit the parrot near the hut.
Do you know what happened next?
He did something very silly.

He climbed up the hut

LOW TRANSITIVITY CONDITION

The panda saw the parrot near the hut.
Do you know what happened next?
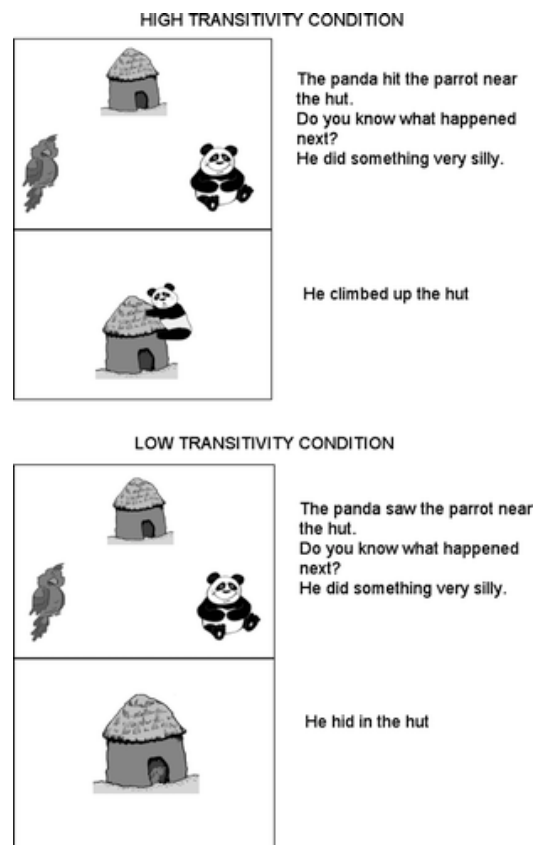He did something very silly.

He hid in the hut

Fig. 12. Example of the materials used in Pyykkönen et al. (2010).

Järvikivi, Pyykkönen-Klauck, Schimke, Colonna, & Hemforth (2014) investigated the impact of focus by syntactic clefts on 4-year-old German children's and adults' pronoun resolution preferences. The children watched animated videos showing two animal characters and a location (Fig. 13), while they listened to mini stories about the depicted characters. The stories manipulated focus (whether the sentence was cleft or not) and grammatical role (whether the cleft antecedent was the grammatical subject or the object) using SVO and OVS sentences, as in (29b-e), in the resolution of the ambiguous pronoun *er* (29f).
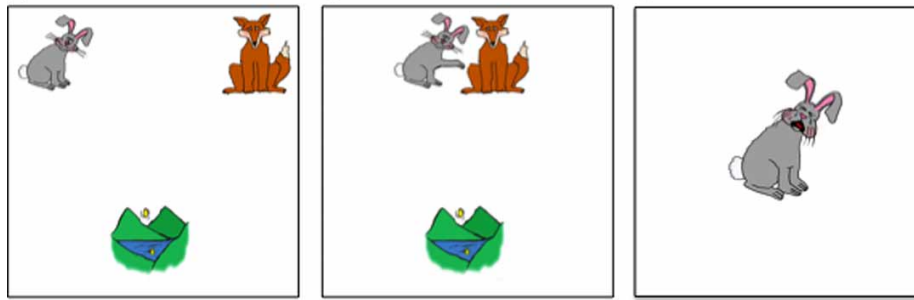
35

Fig. 13. Animated sequence used in Järvikivi et al. (2014), for the example mini story:

29a. Da sind der Hase und der Fuchs/Da sind der Fuchs und der Hase (here are the rabbit and the fox/fox and the rabbit)

> 29b. Der Hase kitzelt den Fuchs, an dem Bergsee (The rabbit tickles the fox at the mountain lake) [SVO, non-cleft]
>
> 29c. Es ist der Hase, der den Fuchs kitzelt, an dem Bergsee (It is the rabbit who tickles the fox at the mountain lake) [SVO, cleft]
>
> 29d. Den Fuchs kitzelt der Hase, an dem Bergsee (The fox (Obj) tickles the rabbit (Sub) at the mountain lake) [OVS, non-cleft]
>
> 29e. Es ist der Fuchs den der Hase kitzelt, an dem Bergsee (it is he fox (Obj) whom tickles the rabbit (Sub) at the mountain lake) [OVS, cleft]

29f. als *er* gerade etwas ganz besonderes Lustiges denkt (when he just something particularly funny thinks) [ambiguous masculine 3rd person pronoun]

29g. Doch dann muß der Hase plötzlich ganz furchtbar weinen (But then the rabbit suddenly rather terribly cries).

In line with previous research (e.g., Kaiser, 2011; See Kaiser in this volume), the adults' gaze data showed an overall preference for subjects over objects but no difference between the cleft and the non-cleft conditions. The children, however, showed an additional interaction between focus and grammatical role, with more looks to the focused (cleft) than unfocused (non-cleft) subjects, but no effect of focus on the object antecedents. This suggests that 4-year-olds are quantitatively largely adult-like in their preferences, but are still unable to weight the different information sources in an adult-like manner, whether due to still limited cognitive control skills for integrating various cues or more limited input/experience – or both – compared to adults.

In sum, the studies from children's pronoun resolution so far suggest that even 3- to 5-year-old children use similar information as adults to determine the referent of a personal pronoun. However, although most of the studies reviewed above have observed effects of structural prominence in children (cf. Arnold et al., 2007), in all of these, the effects appear later in eye movements than in adult studies, whereas young children rapidly appear to use disambiguating morphosyntactic information (e.g., gender) in an adult-like manner at least from 5 years on.

Structural binding constraints

Recently, the visual-world method has also been used to study structural principles in binding theory (e.g., Chomsky, 1986; Reinhart, 1983) that are assumed to constrain the distribution of pronouns and reflexives. Binding theory claims that the distribution of reflexives such as *himself* and pronouns such as *him*, is complementary in that the antecedent for reflexives should be found in a local domain (Principle A), whereas the antecedent for a pronoun is to be found outside the local domain (Principle B). This can be exemplified with so-called picture noun phrases, as in (29), which are often used to assess binding theory.

(29) Harry saw Ken's picture of himself/him.

Structural binding principles predict that *himself* should be taken to refer to *Ken* and *him* to *Harry*. Visual-world studies have investigated the time course of binding principles; whether they are the only constraints that are used during early processing to select the referential domain of a pronoun or reflexive or whether other factors also play a role (see Sturt, 2013, for a comprehensive overview).

Binding theory principles: Adults

Runner, Sussman, and Tanenhaus (2006; Runner, Sussman, & Tanenhaus, 2003) tested the predictions of binding theory in spoken language comprehension using picture noun phrases such as in (30).

30. Look at Joe. Pick up Ken. Have Ken touch Harry's picture of himself/him.

Participants sat in front of a visual display (Fig. 14) showing pictures of three dolls (Joe, Ken, and Harry) and manipulated the dolls according to the instructions.
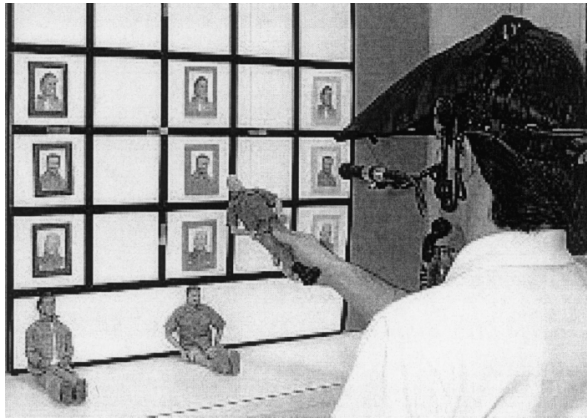

Fig. 14. Visual display used in Runner et al. (2003, 2006).

In contrast to the predictions of binding theory, participants chose the local antecedent (*Harry*) in only 68.9% of the cases when the anaphor was a reflexive (*himself*), going instead for the subject or the lead-in (*Ken* or *Joe*) in 31.1% of the time. In the pronoun conditions, they chose the binding-incompatible antecedent (*Harry*) less often (11.1%). Most crucially, the eye-movement results indicated that the initial interpretation of both anaphors was not fully constrained by binding principles: 300-1000 ms after the reflexive onset, participants looked equally often at the binding-inconsistent subject (*Ken*) as the binding-consistent possessor (*Harry*) and 300-1000 ms after the onset of the reflexive, they looked equally often at the binding-inconsistent possessor as the binding-consistent lead-in (*Joe*), though looks to the binding-consistent subject were most frequent. Thus, binding principles did not function as an early filter for anaphor interpretation. Furthermore, Runner et al. (2003) reported that order of mention of the subject and lead-in (*Pick up Joe. Look at Ken.* vs. *Look at Ken. Pick up Joe*) affected the proportion of looks in the pronoun but not in the reflexive condition, indicating that the domains for pronouns and reflexives are not complementary.

One explanation for why looks to the possessor were as frequent as for the lead-in in the pronoun condition is that the experiment had two possible referents for the pronoun according to binding theory (the subject and lead-in) and only one for the reflexive (the possessor), and so the looks to the binding-compatible antecedents were

split. A further experiment therefore mentioned only two referents by omitting the first lead-in sentence (*Look at Joe* in (30)). The results showed binding-compatible choices in 94% and 82% of the trials for pronouns and reflexives, respectively. The eye movement results showed that 300-1000 ms following either the pronoun or reflexive onset, participants looked more often at both binding theory compatible and incompatible antecedents than at another character (*Joe*) that was not mentioned. Furthermore, there was no difference in looks to the compatible and incompatible antecedents, providing evidence against the idea that binding constraints act as an early filter for antecedent selection.

Kaiser, Runner, Sussman and Tanenhaus (2009) investigated the use of structural and semantic information in the processing of pronouns and reflexives. It is assumed in syntactic accounts that the interpretation of pronouns and reflexives in picture noun phrases is sensitive to structural, but not semantic information. Kaiser et al. (2009) studied two kinds of picture noun phrases, those with (31) and without (32) a possessor:

31. Peter saw the picture of him/himself
32. Mary saw Lisa's picture of her/herself

According to classical binding theory, the local domain for (31) is the whole clause and the reflexive must take the subject (*Peter*) as its antecedent, whereas the pronoun has to refer to an antecedent outside this domain, thus to some other person outside the sentence. In (32), the reflexive must take the antecedent within the local domain, which is the domain of the possessor NP (*Lisa*), whereas the pronoun can now refer to the subject of the sentence (*Mary*). Kaiser et al. carried out two visual-world experiments investigating the influence of perspective, namely whether the potential antecedents were the 'source' or 'perceiver', in both possessorless and possessive picture noun phrases.

In the first eye movement experiment, participants heard sentences such as (33), while they saw scenes such as Fig. 15.

33. Peter told/heard from Andrew about the picture of him/himself on the wall.
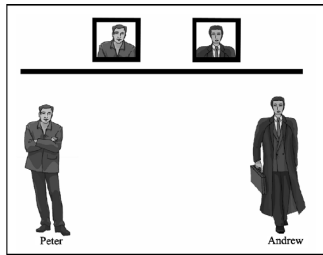
Fig. 15. Visual displays in Kaiser et al.'s (2009) first eye-movement experiment.

For pronouns, the authors found no overall structural preference for the subject (Peter) or object (*Mary*), but an early effect of verb semantics (200-600 ms after pronoun onset), with more looks to the subject when it was the perceiver (*hear*) than the source (*tell*), and the same effect also occurred during later processing, from 1400 ms following the pronoun. For reflexives, there was also some evidence that early interpretation preferences (200-600 ms after the reflexive onset) were affected by verb semantics, with more looks to the object when it was the source (*hear*) than the perceiver (*tell*) 200-600 ms after the reflexive onset, and more looks to the subject when it was the source (*tell*) than the perceiver (*hear*) during later processing (after 1400ms). In addition, reflexives showed an overall structural preference for the subject (*Peter*).

In a further experiment, the authors studied possessive picture noun phrases such as in (34) while participants saw Fig. 16.

34. Peter told/heard from Andrew about Greg's picture of him/himself on the wall.
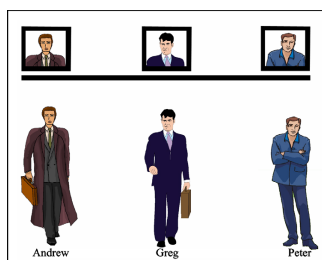

Fig. 16. Visual displays in Kaiser et al.'s (2009) second eye-movement experiment.

The eye movement record showed that participants looked much more often at the possessor for the reflexives than the pronouns from 600 ms after the anaphor onset. Pronouns also showed a late effect of verb semantics (after 1400ms), showing that the

participants looked more at the subject when it was the perceiver (*hear*) than the source (*tell*), but reflexives did not show a verb effect. Overall, the authors concluded that the interpretation of reflexives and pronouns is affected by both structural and semantic information, but that reflexives are more strongly affected by structural constraints, whereas pronouns are more affected by verb semantic information.

Binding theory principles: children

Child studies have shown that whereas children as young as three interpret reflexives like adults as referring to the local antecedent, the principle that pronouns do not take the local antecedent is learned later. The first eye-movement study on children's processing of reflexives and pronouns, Sekerina, Stromswold and Hestvik (2004), found that 4-7 year-old children preferred the local referent with reflexives when they had to choose one of the pictures in Fig. 17 following (35).

35. In these pictures, you see a boy, a man, and a box. The boy has placed the box on the ground. Which picture shows that the boy has placed the box behind him/himself?
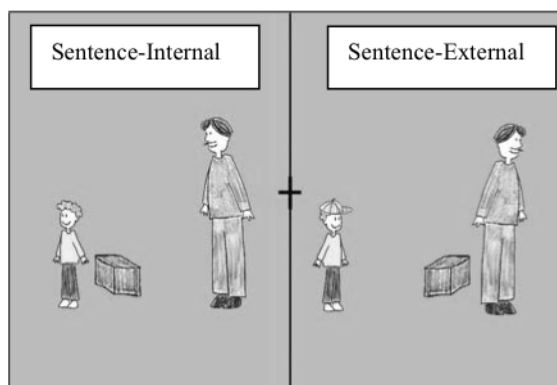

Fig. 17. Visual displays used in Sekerina et al. (2004).

With pronouns, both adults and children chose the picture where the man was the referent. In line with other studies in children's pronoun resolution (see above), this effect appeared later for children than adults.

Recently, Clackson, Felser and Clahsen (2011) further studied 6-9 year-old children's processing of reflexives and pronouns. Participants heard stories such as (36) and saw pictures such as in Fig. 18.

36. Peter was waiting outside the corner shop. He watched as Mr. Jones bought a huge box of popcorn for him/himself.

They also added a further condition where the pronoun and reflexive were disambiguated by gender by replacing *Peter* with *Susan*.
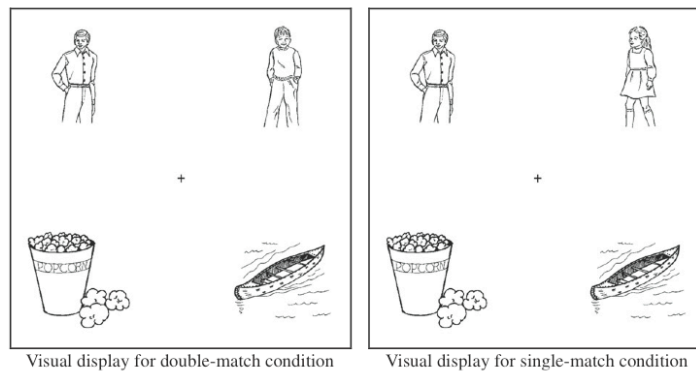

Fig. 18. Visual displays used in Clackson et al. (2011).

Interestingly, the eye-movement data showed that Children looked more frequently at the binding-incompatible antecedent (*Susan/Peter*) when the reflexive was gender ambiguous than unambiguous, indicating that during online processing, they considered the binding-incompatible antecedent when it matched the reflexive's gender. Adults did not consider the binding-incompatible antecedent. With pronouns, both children and adults looked more often at the binding-incompatible antecedent (*Mr. Jones*) in the ambiguous than unambiguous condition. Thus, both groups considered this antecedent during online processing, though the effect was stronger with children. Overall, the results from adults are consistent with Kaiser et al. (2009), who also found that the interpretation of pronouns is less affected by structural information than that of reflexives. The results from children showed that they rely less on structural binding constraints than adults for both pronouns and reflexives, even though they use binding principles for offline reflexive interpretation.

Conclusion

The research we have reviewed in this chapter shows that there are now a considerable number of studies that have used the visual-world method to investigate structural effects in both sentence and referential processing. These studies have

revealed much about online, moment-to-moment processing of structural information in both adults and children. Although some studies suggest that structural information plays a role in sentence and referential processing, a common finding in many studies is that various kinds of non-structural information have an early and strong effect. Research has shown that non-structural factors such as referential context, action-based affordances, verb biases and prosody all have rapid effects on how adults process structurally ambiguous sentences. In fact, the results from these visual-world studies can be explained without assuming the existence of purely structural processing strategies. Interestingly though, children appear to be more restricted in the information they use during structural ambiguity resolution, as they appear to ignore the referential context.

Anticipatory eye movements to objects that are likely to be mentioned next in the sentence also appear to be strongly affected by semantic information, but information that is often seen as syntactic, such as case marking, also plays a role. Furthermore, there is now a series of studies that have shown structural priming effects on anticipatory eye movements, suggesting that that the structure of a preceding sentence affects linguistic anticipations.

Finally, visual-world studies of referential processing have shown that adults and children use both semantic information (implicit causality, semantic agent and patient properties) and structural heuristics such as the first-mention and subject advantage. Even in cases where binding theory postulates structural constraints on the interpretation of pronouns and reflexives, children and adults sometimes ignore them, and at least for adults, there is evidence that semantic constraints play a role.

In sum, visual-world eye-tracking research has significantly advanced our knowledge of how various structural and non-structural factors interactively constrain moment-to-moment sentence and referential processing. This research has only lifted a tip of the veil, so future visual-world research is likely to provide further important insights.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419-439.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30,* 191-238.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73,* 247-264.

Arai, M., van Gompel, R. P. G., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology, 54,* 218-250.

Ariel, M. (1988). Referring and accessibility. *Journal of linguistics, 24*(01), 65-87.

Ariel, M. (2001). Accessibility theory: An overview. In Sanders, T., Schilperoord, J., & Spooren, W. (Eds.), *Text representation: Linguistic and psycholinguistic aspects*. Amsterdam: Benjamins, 29-87.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition, 76,* B13-B26.

Arnold, J. E., Brown-Schmidt, S., Trueswell, J., & Fagnano, M. (2005). Children's use of gender and order of mention during pronoun comprehension. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Processing world-situated language: Bridging the language-as-product and language-as-action traditions.* Boston: MIT Press.

Arnold, J. E., Brown-Schmidt, S., & Trueswell, J. (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes, 22,* 527-565.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18,* 355-387.

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition, 35,* 1-39.

Boland, J. E. (2005). Visual arguments. *Cognition, 95,* 237-274.

Bouma, G., & Hopp, H. (2007). *Coreference preferences for personal pronouns in German.* the Paper presented at the Conference on Intersentential Pronominal Reference in Child and Adult Language.

Branigan, H. P., Pickering, M. J., & McLean, J. F. (2005). Priming prepositional-phrase attachment during comprehension. *Journal of Experimental Psychology-Learning Memory and Cognition, 31,* 468-481.

Britt, M. A. (1994). The interaction of referential ambiguity and argument structure in the parsing of prepositional phrases. *Journal of Memory and Language, 33,* 251-283.

Brown-Schmidt, S., Byron, D. K., & Tanenhaus, M. K. (2005). Beyond salience: Interpretation of personal and demonstrative pronouns. *Journal of Memory and Language, 53*, 292-313.

Carminati, M. N., van Gompel, R. P. G., Scheepers, C., & Arai, M. (2008). Syntactic priming in comprehension: The role of argument order and animacy. *Journal of Experimental Psychology-Learning Memory and Cognition, 34*, 1098-1110.

Carminati, M. N., & van Gompel, R. P. G. (2010). *Structural priming is not short lived*. Unpublished manuscript.

Carreiras, M., Gernsbacher, M. A., & Villa, V. (1995). The advantage of first mention in Spanish. *Psychonomic Bulletin and Review, 2*, 124-129.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language, 47*, 30-49.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology-Learning Memory and Cognition, 30*, 687-696.

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

Chomsky, N. (1986). *Barriers*. Cambridge, MA: MIT Press.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42*, 368-407.

Clackson, K., Felser, C., & Clahsen, H. (2011). Children's processing of reflexives and pronouns in English: Evidence from eye-movements during listening. *Journal of Memory and Language, 65*, 128-144.

Cooper. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*, 84-107.

Cozijn, R., Commandeur, E., Vonk, W., & Noordman, L. G. M. (2011). The time course of the use of implicit causality information in the processing of pronouns: A visual world paradigm study. *Journal of Memory and Language, 64*, 381-403.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological syntax processor. In D. R. Dowty, L. Karttunen & A. M. Zwicky (Eds.), *Natural Language Parsing: Psychological, Computational and Theoretical perspectives* (pp. 320-358). Cambridge, England: CUP.

Crawley, R. A., Stevenson, R. J., & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research, 19*, 245-264.

Dowty, D. R. (1991). Thematic Proto-Roles and Argument Selection. *Language, 67*, 547-619.

Ellert, M. (2010). *Ambiguous pronoun resolution in L1 and L2 German and Dutch.* Wageningen: Ponsen & Looijen.

Ellert, M., Roberts, L., & Järvikivi, J. (2011). Verarbeitung und Disambiguierung pronominaler Referenz in der Fremdsprache Deutsch: Eine psycholinguistische Studie. Krafft, A. & Spiegel, C. (Eds.), *Sprachliche Förderung und Weiterbildung - Transdisziplinär*. Berlin: Peter Lang, 51-68.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*, 348-368.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 164-203.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies.* Unpublished Ph.D. Dissertation. Indiana University Linguistics Club, University of Connecticut.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178-210.

Frederiksen, J. R. (1981). Understanding anaphora: Rules used by readers in assigning pronominal reference. *Discourse processes, 4*, 323-347.

Garnham, A., Traxler, M., Oakhill, J., & Gernsbacher, M. A. (1996). The locus of implicit causality effects in comprehension. *Journal of Memory and Language, 35*, 517-543.

Garnham, A. (2001). *Mental models and the interpretation of anaphora.* Hove: Psychology Press.

Gernsbacher, M. A., & Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language, 27*, 699-717.

Gernsbacher, M. A., Hargreaves, D. J., & Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language, 28*, 735-755.

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition, 32*, 99-156.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention. *Cognitive Science, 17*, 311-347.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 274-307.

Hartshorne, J. K., Nappa, R., & Snedeker, J. (2011. *Ambiguous pronoun processing development: Probably not u-shaped.* In N. Danis, K. Mesh, & H. Sung (Eds.), *BUCLD 35: Proceedings of the 35th annual Boston University Conference on Language Development*, 272-282.

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

Järvikivi, J., Pyykkönen-Klauck, P., Schimke, S., Colonna, S., & Hemforth, B. (2014). Information structure cues for 4-year-olds and adults: tracking eye movements to visually presented anaphoric referents. *Language, Cognition and Neuroscience, 29*(7), 877-892.

Järvikivi, J., Van Gompel, R. P. G., Hyönä, J., & Bertram, R. (2005). Ambiguous pronoun resolution - Contrasting the first-mention and subject-preference accounts. *Psychological Science, 16*, 260-264.

Järvikivi, J., Van Gompel, R., & Hyönä, J. (2015). *The interplay of implicit causality, structural heuristics, and anaphor type in ambiguous pronoun resolution*. Unpublished manuscript.

Kaiser, E. (2011). Focusing on pronouns: Consequences of subjecthood, pronominalization and contrastive focus. Language and Cognitive Processes, 26, 1625-1666

Kaiser, E., & Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes, 23*, 709-748.

Kaiser, E., Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2009). Structural and semantic constraints on the resolution of pronouns and reflexives. *Cognition, 112*, 55-80.

Kako, E. (2006). Thematic role properties of subjects and objects. *Cognition*(101), 1-42.

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

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*, 133-156.

Kidd, E., & Bavin, E. L. (2005). Lexical and referential cues to sentence interpretation: an investigation of children's interpretations of ambiguous sentences. *Journal of Child Language, 32*, 855-876.

Kidd, E., Stewart, A. J., & Serratrice, L. (2011). Children do not overcome lexical biases where adults do: The role of the referential scene in garden-path recovery. *Journal of Child Language, 38*, 222-234.

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

Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. *Journal of Memory and Language, 57*, 519-543.

Koornneef, A. W., & Van Berkum, J. J. A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language, 54*, 445-465.

Ledoux, K., Traxler, M. J., & Swaab, T. Y. (2007). Syntactic priming in comprehension: Evidence from event-related potentials. *Psychological Science, 18*, 135-143.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676-703.

McDonald, J. L., & MacWhinney, B. (1995). The time course of anaphor resolution: Effects of implicit causality and gender. *Journal of Memory and Language, 34*, 543-566.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 38*, 283-312.

Murray, W. S., & Liversedge, S. P. (1994). Referential context effects on syntactic processing. In C. Clifton., L. Frazier & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 359-388). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive Affective & Behavioral Neuroscience, 5*, 263-281.

Novick, J. M., Thompson-Schill, S. L., & Trueswell, J. C. (2008). Putting lexical constraints in context into the visual-world paradigm. *Cognition, 107*, 850-903.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39*, 633-651.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin, 134*, 427-459.

Pyykkönen, P., & Järvikivi, J. (2010). Activation and Persistence of Implicit Causality Information in Spoken Language Comprehension. *Experimental Psychology, 57*(1), 5-16.

Pyykkönen, P., Matthews, D., & Järvikivi, J. (2010). Three-year-olds are sensitive to semantic prominence during online language comprehension: A visual world study of pronoun resolution. *Language and Cognitive Processes, 25*, 115-129.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior, 22*, 358-374.

Reinhart, T. (1983). *Anaphora and semantic interpretation*. London: Croom Helm.

Rose, R. L. (2005). *The relative contribution of syntactic and semantic prominence to the salience of discourse entities.* Unpublished PhD thesis, Northwestern University.

Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2003). Assignment of reference to reflexives and pronouns in picture noun phrases: evidence from eye movements. *Cognition, 89*, B1-B13.

Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2006). Processing reflexives and pronouns in picture noun phrases. *Cognitive Science, 30*, 193-241.

Scheepers, C., & Crocker, M. W. (2004). Constituent order priming from listening to comprehension: A visual-world study. In M. Carreiras & C. Clifton (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERP, and Beyond* (pp. 167-185). New York: Psychology Press.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*, 109-147.

Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistic Research, 32*, 3-23.

Sekerina, I. A., Stromswold, K., & Hestvik, A. (2004). How do adults and children process referentially ambiguous pronouns? *Journal of Child Language, 31*(1), 123-152.

Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research, 23*, 197-229.

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language, 48*, 103-130.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*, 238-299.

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

Song, H. J., & Fisher, C. (2005). Who's "she"? - Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language, 52*(1), 29-57.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology, 45*, 447-481.

Stewart, A. J., Pickering, M. J., & Sanford, A. J. (2000). The time course of the influence of implicit causality information: Focusing versus integration accounts. *Journal of Memory and Language, 42*, 423-443.

Sturt, P. (2013). Referential processing in sentences. In R. P. G. van Gompel (Ed.), *Sentence processing*. Hove: Psychology Press.

Tanenhaus, M. K., Spivey Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632-1634.

Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. *Cognition, 108*, 51-68.

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.

Tooley, K. M., Traxler, M. J., & Swaab, T. Y. (2009). Electrophysiological and Behavioral Evidence of Syntactic Priming in Sentence Comprehension. *Journal of Experimental Psychology-Learning Memory and Cognition, 35*, 19-45.

Townsend, D., & Bever, T. G. (2001). *Sentence Comprehension: The Integration of Habits and Rules*. Cambridge MA: MIT Press.

Traxler, M. J., & Tooley, K. M. (2008). Priming in sentence comprehension: Strategic or syntactic? *Language and Cognitive Processes, 23*, 609-645.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology Learning, Memory, and Cognition, 19*, 528-553.

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

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-
path effect: studying on-line sentence processing in young children. *Cognition,
73*, 89-134.

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context
effects in sentence processing: Evidence from event-related brain potentials.
Journal of Memory and Language, 41, 147-182.

van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Jacob, G. (2006). The
activation of inappropriate analyses in garden-path sentences: Evidence from
structural priming. *Journal of Memory and Language, 55*(3), 335-362.

Vilkuna, M., 1989. Free Word Order in Finnish. Its Syntax and Discourse Functions.
Suomalaisen Kirjallisuuden Seura: Helsinki.

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the
interpretation of structural ambiguities: As study of anticipatory eye
movements. *Cognition, 99*, B63-B72.

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