Available online at www.sciencedirect.com
SCIENCE DIRECT®
COGNITION
ELSEVIER
Cognition 94 (2004) 113–147
www.elsevier.com/locate/COGNIT

The role of discourse context in the processing of a flexible word-order language

Elsi Kaiser[a,b,*], John C. Trueswell[a,b]

[a]*Department of Linguistics, University of Pennsylvania, 619 Williams Hall, 36th and Spruce St., Philadelphia, PA 19104 6305, USA*
[b]*Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA*

Received 5 July 2002; revised 19 November 2003; accepted 15 January 2004

Abstract

On-line comprehension studies of flexible word-order languages find that noncanonical ('scrambled') structures induce more difficulty than canonical structures [e.g., Hyönä & Hujanen, *Q. J. Exp. Psychol*. 50A (1997) 841–858], with this difference being attributed to the structural complexity/infrequency of these forms. However, by presenting sentences in isolation, most existing studies have not examined the extent to which processing patterns can be explained by discourse factors. In Finnish, an articleless flexible word-order language with canonical SVO order, word order can be used to encode the given/new distinctions: OVS marks the object as given and the subject as new; SVO is more flexible, being used in multiple contexts. Thus, the observed dispreference for noncanonical orders in previous work may arise from additional presuppositions needed to understand them out of context [cf. Crain & Steedman, *Natural Language Parsing*, (1985), 320–358]. Two experiments were conducted to examine these issues in Finnish. Experiment 1 used a self-paced reading task to show that the usual difficulty associated with noncanonical constructions is partially alleviated in the presence of appropriate discourse contexts. In Experiment 2, the eye gaze of listeners was tracked as they heard spoken descriptions of scenes, so as to test whether specific on-line referential processes are involved in SVO/OVS comprehension: Upon hearing an OV… sequence, Finnish listeners should expect the upcoming noun to be discourse-new, whereas an SV… sequence makes no such prediction. The results confirmed these predictions. As compared to SVO, OVS sentences showed anticipatory eye movements to a discourse-new referent at the second noun onset, even before participants had enough acoustic information to recognize this word. Our findings illustrate that in a flexible word-order language, a noncanonical order can result in anticipatory processes regarding the discourse status of a yet-to-be-heard constituent.

* Corresponding author. Address: Center for Language Sciences, 421 Meliora Hall, University of Rochester, Rochester, NY 14627, USA.
E-mail address: ekaiser@bcs.rochester.edu (E. Kaiser).

0022-2860/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2004.01.002

Keywords: Sentence processing; Word order; Scrambling; Discourse; Finnish

1. Introduction

Human languages differ in the amount of word order flexibility they permit. Some languages, including English, have fairly rigid word order. If the word order of an English sentence—e.g. *The bird ate a worm*—is changed, the meaning of the sentence also changes: *A worm ate the bird*. This is a consequence of English using word order to encode the grammatical relations between words, which help to constrain thematic role assignments (i.e. 'who did what to whom'). However, many other languages, such as Finnish, German and Russian, encode these grammatical relations not by means of word order, but by case marking on the nouns. Hence it is case marking rather than word order that contributes more to role assignment. In the Finnish sentence (1a) below, the bird (*lintu*) is the eater not because of its position in the sentence but because it is in nominative case rather than accusative case (accusative is marked with the ending '-n').[1]

1.	a.	*Lintu söi madon.*	b.	*Madon söi lintu.*
		bird-NOM ate worm-ACC		worm-ACC ate bird-NOM

Perhaps not surprisingly, languages with rich case-marking systems, such as Finnish, often have flexible word order. For instance, sentence (1b) means the same as (1a) in terms of who is doing what to whom even though the order has changed to Object–Verb–Subject (OVS).

The existence of word order flexibility raises questions about the relations between the different word orders that are possible in a language. On an intuitive level, native speakers often feel that one of the orders is the basic, 'default' order, while the other orders are perceived to be somehow more unusual. Even though English has rigid word order and thus cannot be used to illustrate this precise intuition, a similar kind of contrast can be made by comparing a basic sentence (e.g. 'The bird ate a worm') and a more complex syntactic structure such as an *it*-cleft ('It was a worm that the bird ate'). The latter structure feels more 'unusual' than the basic sentence.

From a syntactic perspective, it is often assumed that the more 'unusual' orders are generated from the structure of the basic order by some kind of syntactic reorganization, usually dubbed 'scrambling' (Ross, 1967).[2] Syntacticians are often interested in 'scrambled' sentences because they can provide insights into the grammatical representations that generally operate in a language, as well as universal constraints on

[1] Abbreviations used in this paper: NOM, nominative case; ACC, accusative case; GEN, genitive case; PART, partitive case.

[2] Work in different languages and within different linguistic traditions has sometimes used different terms for different kinds of word-order-related movement operations. For example, in German linguistics, the term 'scrambling' is used to refer to one kind of word order variation, and 'topicalization' is used for another kind of movement operation. Since we are focusing on Finnish and these kinds of terminological distinctions have not been made for Finnish, we will employ the term 'scrambling' in a rather general sense to mean all kinds of word-order related movement operations.

linguistic representation (e.g. Bailyn, 1995; Baker, 2001; King, 1993; Saito, 1985; Tada, 1993; Webelhuth, 1989).

From a discourse-pragmatic perspective, sentences with unusual orders are interesting because of their special discourse requirements. Typically, the 'basic' order can be felicitously uttered in the absence of any substantial context—the usual test being whether it can be uttered in response to a question such as "What happened?" (see e.g. Pinto, 1997). This order is often referred to as the 'canonical' order. In contrast, the more unusual, 'noncanonical' orders require a specific discourse context to be felicitous. Thus, discourse theorists and semanticists are typically concerned with exploring the kinds of contexts that make noncanonical orders felicitous, or to put it in terms we will adopt, they are interested in the discourse functions of noncanonical structures (e.g. Birner & Ward, 1998; De Hoop, 1992; Prince, 1998, 1999; Vallduví, 1992).

For the psycholinguist, one important consideration is understanding how readers and listeners go about parsing and comprehending noncanonical sentences. For instance, to what extent do complex syntactic operations and unique discourse functions contribute to the time needed to understand these sentences? And how do noncanonical structures contribute to referential processing of their constituent noun phrases? Our focus here will be on these questions, with a particular emphasis on how canonical and noncanonical sentences are processed on-line as each sentence is being understood. Investigating the on-line processing of these kinds of sentences relates to an issue that has received much attention in the sentence processing literature, namely, how and when various evidential sources are used to recover the relevant structure. Furthermore, in light of findings in the linguistic literature about the discourse functions of canonical and noncanonical sentences, we will be asking about the role these structures play in an ongoing discourse and hence how readers and listeners use structures to guide referential interpretation.

2. Processing noncanonical structures

The existing work on the processing of noncanonical structures has tended to approach the question from a syntactic perspective. It has been argued that noncanonical structures are inherently more difficult to process than canonical structures, with the central reason being that they are syntactically more complex and hence computationally more costly to represent than their canonical counterparts (e.g. De Vincenzi, 1991; Frazier & Flores d'Arcais, 1989). However, some psycholinguistic researchers have also noted that the use of noncanonical structures is motivated by discourse-pragmatic factors which need to be taken into account when comparing the processing of canonical and noncanonical orders, since canonical orders are not subject to the same degree of discourse dependence (e.g. Bader & Meng, 1999; Kaan, 2001). In the sections below, we review the existing work on the processing of noncanonical structures and the conclusions that have been drawn from it concerning the nature of the human sentence processing mechanism.

2.1. The role of syntactic and computational complexity

To date, numerous researchers have shown that native speakers of flexible word-order languages experience difficulty when reading noncanonical sentences in isolation (e g. Bader & Meng, 1999; Frazier & Flores d'Arcais, 1989; Hemforth, 1993). Many of these studies have examined situations in which a temporary ambiguity arises between a canonical and noncanonical structure. That is, on target trials, a noun early in the sentence carries a case marker that is consistent with either a canonical or a noncanonical order. Later in the sentence, another word resolves the ambiguity toward a particular order and interpretation. Studies of this sort have shown that readers pursue a canonical interpretation at the point of ambiguity, as evidenced by the increases in reading time when the sentence later resolves toward a noncanonical structure. In addition, even noncanonical sentences that are locally unambiguous have been found to increase reading times as compared to unambiguous canonical sentences (e.g. Hyönä & Hujanen, 1997; Miyamoto & Takahashi, 2002; Vasishth, 2002).

Explanations of the difficulty with noncanonical forms and the preference for canonical forms have often centered on the representational and/or computational complexity of scrambled constructions. For instance, the canonical preference observed for temporally ambiguous structures (e.g. Bader & Meng 1999, Frazier & Flores d'Arcais, 1989; Hemforth, 1993; Kamide & Mitchell, 1999; Stojanović, 1998) has been explained in terms of De Vincenzi's (1991) Minimal Chain Principle (MCP), which is a more general version of Frazier and Flores d'Arcais' (1989) Active Filler Strategy (AFS). The MCP assumes that in order to understand a sentence involving movement (e.g. Wh-questions in English like "Which ball did Mary believe John took?"), readers and listeners must first recover the syntactic "chain" between a moved constituent (i.e. "Which ball") and its original location in the sentence structure (i.e. as the direct object of "took"). The MCP states that unnecessary chain members should not be postulated, but required chain members should be postulated at the earliest possible point. If one assumes that a scrambled structure is generated via syntactic movement, then a principle such as the MCP predicts a preference for canonical orders over noncanonical orders whenever an ambiguity is present. In terms of processing timing, the MCP has traditionally been embedded in a two-stage theory of parsing, in which syntactic operations like the MCP temporally precede other operations, therefore predicting that the MCP will guide initial parsing decisions independent of context.

When the parser is faced with a constituent that is unambiguously marked as being in a noncanonical position, ambiguity resolution strategies such as the MCP or the AFS do not make any predictions. However, approaches such as Gibson's (1998) Syntactic Prediction Locality Theory (SPLT) predict that a scrambled constituent induces a greater processing load because of the associated memory cost. For example, a sentence-initial object in a language that is canonically Subject–Verb–Object (SVO) gives rise to the expectation that the subject will occur at some later point in the sentence. According to the SPLT, retaining this prediction in memory increases memory load, which is reflected in slower reading times and other processing difficulties (see also Miyamoto & Takahashi, 2002; Nakayama & Lewis, 2000; Vasishth, 2002).

In addition to these syntax-based strategies, the difficulty with noncanonical orders has also been attributed to the relative frequency of canonical structures and infrequency of

noncanonical structures (Hyönä & Hujanen, 1997, see also Gibson, 1998). For instance, Hyönä and Hujanen explained Finnish readers' preference for the canonical SVO order (as compared to noncanonical OVS) as most likely arising from the high frequency of the SVO order, which facilitates the parsing process.

2.2. Role of contextual factors

It has often been noted in the literature that scrambling is not random or arbitrary. Rather, scrambling appears to be driven by discourse-based factors, such as whether a certain entity has already been mentioned or whether an entity is in a set relation with something else in the discourse (e.g. Birner & Ward, 1998; Givón, 1984; Lambrecht, 1994; Prince, 1999, inter alia). (See also the related literature on the role of context in interpreting the English passive, e.g. Gourley & Catlin, 1978; Olson & Filby, 1972.) It is often the case that, across languages, entities that have not yet been mentioned in the discourse (new information) tend to occur toward the end of the sentence, whereas entities that have already been mentioned (old/given information) tend to occur toward the beginning of the sentence. We will discuss Finnish word order patterns in more detail in Section 3, but clearly the use of word order to encode discourse-related information is not restricted to Finnish. In other articleless scrambling languages such as Japanese and Russian, the discourse properties of constituents are also related to their positions in the sentence (Ishihara, 2001; Yokoyama, 1986, inter alia). Even in German, a scrambling language that has articles, the position of the arguments of the verb reflects their connection to the preceding discourse (see e.g. Lenerz, 1977, and others).

It is therefore possible that much of the processing difficulty with noncanonical word orders is due not to the infrequency or complexity of these structures but rather to a violation of the discourse demands of these structures as compared to their canonical counterparts. Indeed, most comprehension studies of scrambling have presented noncanonical structures in isolation (i.e., without any preceding discourse), even though these structures can serve highly specific discourse functions. Thus, some of the difficulty observed with these structures could pertain to the need to presuppose a detailed discourse to support a noncanonical structure.

This line of reasoning stems directly from the Referential Theory of sentence parsing (Altmann & Steedman, 1988; Crain, 1980; Crain & Steedman, 1985). Many early studies of on-line sentence processing focused on how readers resolved temporary ambiguities in English. These studies commonly presented target sentences in isolation, and it was observed that readers almost uniformly preferred interpretations that were syntactically less complex. This preference was proposed to reflect a human parsing strategy to avoid syntactically complex alternatives (Minimal Attachment, Frazier & Fodor, 1978; Rayner, Carlson, & Frazier, 1983). However, Crain and colleagues noted that more complex syntactic alternatives usually have more specific discourse functions associated with them (i.e., they presuppose a more specific/complex discourse). They therefore suggested that readers' initial parsing preferences do not reflect consideration of syntactic complexity but rather discourse complexity: in isolation, certain interpretations are dispreferred because they require additional discourse presuppositions. The same argument is being made here

about noncanonical structures. They are dispreferred in isolation because they require more presuppositions than their canonical counterparts. Evidence to date on English syntactic ambiguity resolution does show that referential contexts exert the parsing preferences predicted by Crain and colleagues but that the availability of syntactic alternatives, as determined by lexical-biases, also plays a role (Altmann & Steedman, 1988; Britt, 1994; Spivey & Tanenhaus, 1998). This latter finding is consistent with multiple-constraint parsing theories that permit simultaneous exertion of bottom-up lexical information and top-down contextual information (e.g. MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994).

To the best of our knowledge, only two published studies of flexible word-order languages have examined the comprehension of noncanonical sentences when embedded in discourse contexts (Bornkessel, Schlesewsky, & Friederici, 2003; Sekerina, 2003), and two studies have taken the discourse-driven nature of scrambling into account for sentences in isolation, by looking at the processing of different referential expressions (pronouns vs. full NPs) in canonical and noncanonical structures (Bader & Meng 1999; Kaan, 2001). The Sekerina (2003) study examined Russians' reading time patterns to sentences that were preceded by a single-sentence context claimed to be appropriate for scrambled orders. Sekerina does not specifically discuss which kinds of discourse factors motivate scrambling, and the study used the same context for both canonical and noncanonical sentences. She found a facilitatory effect of context for all sentence types, but an analysis of the word-by-word reading times revealed that there was still an effect of movement on the reading times. Thus, although her study provides further support for the idea that sentences in general are read faster when preceded by a context than in isolation, it leaves open whether certain contexts might eliminate or reverse the canonical preference.

Recent ERP research by Bornkessel et al. (2003) investigated the processing of German word order variation (canonical subject–object order and noncanonical object–subject order) in embedded sentences. Target sentences in their experiment were preceded by different kinds of wh-question contexts. Bornkessel et al. found a brief facilitatory effect of context for object–subject order, but only under particular contextual circumstances—specifically, when the wh-question context prompted a prediction which perfectly matched the noncanonical object and structure. Bornkessel et al. suggested that context effects on the processing of noncanonical structures only arise when a very specific contextual requirement needs to be satisfied. However, it is not clear to what extent the ERP findings reflect discourse contributions or merely the repetition of identical lexical and noncanonical structural properties across the question context and target answer.[3]

[3] There is related experimental evidence suggesting this might be the case. In a Ph.D. dissertation, Weskott (2003) used self-paced reading methodology to investigate the effects of different contextual manipulations on the processing of SVO and OVS sentences in German. His results replicate earlier findings about the existence of a strong subject-first preference, but also show that the processing of noncanonical, object-initial sentences can be facilitated by the co-occurrence of certain strong contextual factors, especially the presence of an immediately preceding 'prime' sentence that also has OVS order. This suggests that structural repetition might be creating the facilitation effect. Crucially, the work we report here does not precede the noncanonical target sentences with a sentence of the same structure.

A study by Kaan (2001) took steps to incorporate the discourse-driven nature of scrambling into an experimental design involving sentences presented in isolation. Here she investigated the strength of the canonical order (subject–object) preference for locally ambiguous NP1–NP2–V sequences in Dutch relative clauses. On target trials, the second NP was always case-ambiguous (subject or object) but it varied between being a definite NP or the pronoun *jullie* (you). The idea behind this manipulation is that pronouns, which encode information that is highly salient at that point in the discourse, "occur in subject position more often than full definite NPs" (Kaan, 2001, p. 533, see also Prince, 1992). In NP–NP–V sequences where the second noun is a pronoun, one might expect readers to be sensitive to the correlation between pronouns and subjecthood and to no longer show a default canonical subject–object order preference. Indeed, Kaan (2001) found that definite NPs generated a canonical subject–object order preference, but pronouns reduced or eliminated this preference. Kaan concluded that the preference for the canonical 'subject-first' order is not as strong as has often been assumed and can be influenced by NP type, which in turn is related to the discourse status of the referent of the NP. Thus her results indirectly suggest that people's preference for the canonical word order can be swayed by discourse information.

A related experiment by Bader and Meng (1999) using a speeded-grammaticality judgment task compared different kinds of subject–object ambiguities in German, including sentences with two full noun phrases and sentences with a pronoun and a full noun phrase. In German, according to Bader and Meng, temporally ambiguous sentences with noun–noun sequences where the first noun (subject or object) is a pronoun are more flexible in terms of their discourse properties than sentences with a noun–noun sequence where both are full NPs. In the study, they found that NP–NP sequences resulted in a stronger preference for the canonical subject–object order than pronoun–NP sequences. Following a serial model of parsing, they interpreted these results as evidence that the garden-path is stronger with NP–NP sequences that are disambiguated as object–subject structures than pronoun–NP sequences that are disambiguated as object–subject structures, because the former also require that the information structure of the sentence be revised. Like Kaan's results, these findings indirectly indicate that discourse information plays an important role in on-line parsing of scrambled constructions. And both studies raise the question of whether a noncanonical sentence could become as easy to process as a canonical sentence when located in a supportive context.

In sum, the discourse functions of noncanonical structures and the contribution of discourse factors to their on-line processing have not been extensively explored in the experimental psycholinguistic literature. Aside from the important exceptions discussed above, researchers have tended to attribute the processing difficulty of these sentences to their syntactic complexity rather than the discourse violation of presenting these sentences in isolation.

3. Experimental preliminaries: Facts about Finnish

We report here two experiments that examine the processing of canonical and noncanonical structures in Finnish. In order to explain adequately our hypotheses, we first

provide some linguistic details about word-order variation and case marking in Finnish. The upshot of this section is that the noncanonical OVS order arises in specific discourse contexts in which the object is discourse-old and the subject is discourse-new (Section 3.1). In addition, the case marking on the initial object in OVS sentences is often temporarily ambiguous between being part of a noncanonical structure or part of a canonical structure (Section 3.2).

3.1. Word order variation

In Finnish, SVO is regarded as the canonical order and is also documented to be the most frequent order (Hakulinen & Karlsson, 1979). However, all six possible permutations of these three elements are grammatical in appropriate contexts (Vilkuna, 1989, 1995). For purposes of brevity, we focus here on the discourse requirements of SVO and OVS orders. For discussion of the discourse properties of the other orders, including the role of contrastive focus, see Hakulinen and Karlsson (1988), Kaiser (2000), Vilkuna (1989, 1995), Vallduví and Vilkuna (1998), inter alia.

As noted above, Finnish has no definite or indefinite articles, and it appears that the SVO/OVS alternation is used to make many of the distinctions that, in languages such as English, are accomplished by using articles. A subject noun that is located in its canonical preverbal position is usually interpreted as 'old' information, i.e. refers to an entity that has already been mentioned in the discourse. In contrast, subjects that occur after the verb (in a noncanonical position) are interpreted as being information that is new to the discourse (resembling indefinites in English). This is illustrated by the following Finnish translation of a sentence from Beatrix Potter's *The Tale of Mr. Jeremy Fisher* (1979) (example from Chesterman, 1991, p. 100).

2. a. *English: S V* **A great big water-beetle** came up underneath ***the lily leaf***.
   b. *Finnish: V S* ***Lumpeenlehden*** alla ui **iso vesikuorianen**.
   lily-leaf-GEN under swam big-NOM water-beetle-NOM.

Here, the indefinite noun phrase *a great big water beetle* is translated into Finnish by locating it after the verb *ui* 'swims'. Thus, Finnish reflects a tendency that appears to be common cross-linguistically, namely the placement of new information later in the sentence. However, new information—specifically, new subjects—can also occur preverbally if both the subject and object of a sentence are discourse-new. In this case, they tend to occur in the canonical SVO order.

The interpretation of the given/new status of objects is somewhat more complicated. Objects in their canonical postverbal position (e.g. SVO) can be new or old information with the distinction made based on context. However, when an object occurs in a preverbal, noncanonical position followed by a postverbal subject (OVS order), the object is interpreted as old information, as illustrated by the example below. The postverbal subject is interpreted as being new information, as mentioned above. Preverbal objects thus resemble preverbal subjects in this regard (see e.g. Chesterman, 1991).

3. OVS **Tiedotteen** välitti julkisuuteen kurdien uutistoimisto D.E.M.
Announcement-ACC transmitted public-to Kurds-GEN newsoffice-NOM D.E.M.
'**The announcement** was made public by the Kurdish newsoffice D.E.M.' (from *Aamulehti* 3/16/1999)

In sum, Finnish word order partially encodes the given/new distinction. OVS order typically marks the object as given, and the subject as new. SVO order is more flexible: it can be used when the subject is given and the object is new, and also when both are old or both are new. In other words, when compared to SVO order, OVS order has only a small set of pragmatic contexts in which it is felicitous, and it is not felicitous when both constituents are new (as is the case when this order is read in isolation).

3.2. Case marking

Finnish has a rich case-marking system, possessing a total of 15 cases. The subject of a sentence is usually in the nominative case. According to a corpus study by Hakulinen, Karlsson, and Vilkuna (1980), 89% of all Finnish subjects in their corpus are in the nominative case. However, subjects can also occur in the partitive (7% of all subjects in the same corpus). The partitive case is used with mass nouns (e.g. *vettä* water-PART 'some water') and bare plurals (e.g. *poikia* boys-PART 'boys') as illustrated in (4).

4. (a) *Partitive subject* **Poikia** leikki pihalla.
boys-PART played yard-in
'**Boys** were playing in the yard.'
(b) *Nominative subject* **Poika** leikki pihalla.
boy-NOM played yard-in
'**The/a boy** was playing in the yard.'

In addition, partitive case is also used with so-called experiencer verbs (e.g. *frighten*, *worry*), as shown in (5), modified from Hakulinen and Karlsson (1988, p. 99). In these kinds of constructions, the experiencer is marked with partitive case (with some verbs also with allative or some other case) and the cause of the experience ('this waiting' in (5)) is marked with nominative case. Importantly, the order shown in (5), with the partitive-marked experiencer located preverbally and the nominative causer in a postverbal position, is the unmarked order for this construction and follows the discourse patterns of other unmarked constructions as described above (see Hakulinen & Karlsson 1988, p. 99).

5. *Naista väsyttää tämä odottaminen*.
woman-PART tires this-NOM waiting-NOM.
'Woman is tired by this waiting.'

In addition to being used for certain kinds of subjects, the partitive case is also routinely used to mark objects. In fact, it is the most common case for objects: 58% of all objects in the corpus of Hakulinen et al. (1980) were in the partitive case. This means that a sentence-initial NP marked with partitive case generates a temporary ambiguity, as it could belong to more than one structure; for example, it could be part of a noncanonical OVS sentence or a canonical sentence involving an experiencer verb.

The second most common object case is accusative (20% of the objects in the Hakulinen et al. corpus). Like the partitive, the accusative case is also ambiguous and does not, on its own, provide sufficient information to ascertain the grammatical function of the noun. This is because for singular nonpronominal count nouns, accusative case is morphologically identical to the genitive case. For example, *miehen* can be interpreted as man-GEN (man's) or as man-ACC.

Thus, in sum, the most common object case markings (partitive and accusative) generate temporary ambiguities when they appear on a sentence initial noun: only the type of verb will determine if the sentence is intended to be canonical or noncanonical.

4. General predictions

In the experiments presented in this paper, we investigate two hypotheses based on the discourse-driven nature of scrambling. First, we hypothesize that the increased processing load found in earlier scrambling studies—usually attributed to the complexity or relative infrequency of scrambled structures—is actually due to the fact that the stimuli were presented without any context. Thus, the establishment of a discourse that satisfies the presuppositions of such structures ought to greatly mitigate any processing difficulty. Second, we hypothesize that the discourse information carried by word order in Finnish is used by comprehenders incrementally during processing. Thus, the presence of a noncanonical OV… order ought to lead to the expectation that the upcoming subject is discourse-new, an entity not yet mentioned in the discourse.

Experiment 1 tested the first hypothesis by using a self-paced reading task to investigate how felicitous and infelicitous discourse contexts influence the processing difficulty associated with scrambled sentences. Experiment 2 tested the second hypothesis by using an eye-movement-during-listening task to investigate how the word order of a sentence conveys given-new status to listeners who are engaged in a referential task.

5. Experiment 1: Self-paced reading

In this experiment, native speakers of Finnish participated in a self-paced reading experiment that contained noncanonical (OVS) and canonical (SVO) sentences. These target sentences appeared in discourse contexts that either did or did not support the typical discourse functions of these structures. In supportive contexts, OVS sentences appeared with the object already given (i.e. already introduced in the preceding story) and with the subject new (i.e. introduced as a new referent in the target sentence).

In the unsupportive contexts, the opposite was true: the object was new and the subject was given. Similarly, SVO sentences appeared either in a supportive (subject given, object new) or an unsupportive (subject new, object given) context.

The initial noun in OVS sentences contained a case marker which in most cases made it ambiguous between an OVS structure and a more canonical structure (predominately partitive and accusative case markers, see discussion above in Section 3). In addition, even those sentence-initial nouns that had case markers which did not exhibit this particular kind of ambiguity could be interpreted as being part of more than one structure. For example, they could be ambiguous between and OVS structure and a nominalized, prenominal relative-clause type construction.

If the previously observed processing difficulty and dispreference for noncanonical structures is due to the target sentences having been presented without an appropriate discourse context, then we expect OVS sentences in a felicitous context to be easier to process. If, on the other hand, it is the structural complexity or infrequency of noncanonical structures that makes them hard to process, we do not expect a context manipulation to have any effect on the ease of processing either SVO or OVS sentences.

In particular, if one adopts a syntax-first approach to parsing, which claims that only syntactic information is available to the parser in the very earliest stages of processing (Ferreira & Clifton, 1986; Frazier & Fodor, 1978), then one predicts that even if context can have a facilitatory effect on the processing of noncanonical structures, this effect will not be present at the initial stages of processing. In contrast, if one adopts a referential approach, difficulty with OVS is expected to be eliminated. A final possibility, that of simultaneous effects of context and structure, would be expected under constraint-satisfaction theories of parsing (MacDonald et al., 1994; Trueswell & Tanenhaus, 1994). As already mentioned, it has been found in English that referential effects depend upon the availability of syntactic alternatives (Spivey & Tanenhaus, 1998; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993).

5.1. Method

5.1.1. Participants

Forty-four native Finnish speakers, mainly students at the Helsinki University of Technology, participated in this experiment. Participants received approximately $4 for participation in the experiment.

5.1.2. Materials

The critical materials manipulated the given/new status of the subject and object in SVO and OVS sentences. Twenty critical items were designed, and each item consisted of a two-sentence context, followed by the third (target) sentence. Critical sentences are provided in Appendix A. An example item is provided in (6) below. The contexts were used to establish either the subject or object of the third sentence as given. The other argument was introduced in the target sentence. This resulted in four conditions: canonical sentences in supportive contexts [$S_{given}VO_{new}$]; canonical sentences in unsupportive

contexts [$S_{new}VO_{given}$]; noncanonical sentences in supportive contexts [$O_{given}VS_{new}$]; and noncanonical sentences in unsupportive contexts [$O_{new}VS_{given}$].

6. (a) **Context Sentences (and Literal Translations)**
Lotta etsi eilen sieniä metsässä.
(Lotta looked-for yesterday mushrooms forest-in.)
Hän huomasi heinikossa hiiren/jäniksen joka liikkui varovasti eteenpäin.
(S/he-NOM noticed grass-in mouse-ACC/hare-ACC that was moving carefully forward.)

(b) **Target Sentences (and Literal Translations)**
SVO: *Hiiri seurasi jänistä ja linnut lauloivat.*
(Mouse-NOM followed hare-PART and birds were-singing.)
OVS: *Jänistä seurasi hiiri ja linnut lauloivat.*
(Hare-PART followed mouse-NOM and birds were-singing.)

The target sentences introduced the subject, verb and object in the first three words. The words that were used as subjects and objects were all singular count (nonmass) nouns. The frequencies of the subject and object in each sentence were matched as closely as possible—that is, a noun with a certain frequency co-occurred with another noun with a similar frequency (based on the frequencies reported in Saukkonen, Haipus, Niemikorpi, & Sulkala, 1979). The nouns were also matched in terms of their level of animacy/sentience: either two animals or two humans were used.

Four presentation lists were constructed by randomly combining the 20 target stories with 35 filler stories. Within a presentation list, 10 of the target trials appeared with the SVO structure and 10 appeared with the OVS structure. For each of these sentence structure types, five appeared in supportive contexts and five appeared in unsupportive contexts. Each target item was then rotated through these four conditions, generating four different presentation lists.

The filler stories were similar in length to the critical items. The fillers did not have any OVS orders in the third sentence. This was done in order to make the overall SVO/OVS distribution in the experiment mimic the relatively low frequency of OVS sentences compared to SVO sentences found in corpus studies of Finnish (Hakulinen & Karlsson, 1979).

5.1.3. Procedure

The experiment was run on a PC using the DMASTR software developed at Monash University and at the University of Arizona by K.I. Forster and J.C. Forster. Each trial consisted of the participant silently reading the story and answering a yes/no comprehension question. For each trial, the participant pressed a button on a button box to reveal a portion of text such that each button press revealed further text and masked the previously revealed text.

On both fillers and target trials, the first two sentences of each item were presented one phrase at a time. The third (final) sentence of each item was then presented word-by-word, and the reading time for the first five words of the target sentence (the NP–V–NP-sequence and the first two subsequent words) was recorded. We opted for this presentation style, instead of word-by-word presentation for all three sentences, because of concerns

about button-pressing fatigue. The lists were preceded by five practice trials, and participants were told to read the sentences for comprehension. Word-by-word self-paced reading methodology is frequently used in the literature to measure localized processing difficulty in comprehension (e.g. Altmann & Steedman, 1988; Gennari & Poeppel, 2003; Gibson, 1998; Hsiao & Gibson, 2003; Just, Carpenter, & Woolley, 1982; MacDonald, 1994; McElree, Traxler, Pickering, Seely, & Jackendoff, 2001; Sedivy, 2002; Spivey-Knowlton & Sedivy, 1995; Warren & Gibson, 2002). In addition, several studies have compared directly word-by-word reading times to eye movement reading times and found comparable patterns of localized difficulty (e.g. Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Ferreira & Henderson, 1990, Trueswell, Tanenhaus, & Kello, 1993).

All trials were followed by a simple yes/no comprehension question, which the participant answered by pressing buttons labeled *kyllä* 'yes' or *ei* 'no'. In order to encourage the participants to focus equally on all parts of each story, a randomly chosen one-third of the comprehension questions were based on the first sentence, a third on the second sentence, and a third on the last sentence. This was done on the set of all items (critical items and fillers). The computer provided the participants with feedback immediately after the participant had answered each question. After answering the question and receiving feedback, the participants then pressed a button to move onto the next item. Only answers to questions following critical trials were recorded.

5.1.4. Data analysis

We analyzed participants' reading times and question-answering accuracy. The raw reading times (per position) were trimmed in the following way. First, all reading times below 100 ms were dropped. Then, for each position (combining all conditions), the mean reading time and the SD were computed. Reading times that were more than three SDs away from the mean were adjusted to mean plus three SDs. In total, this leads to less than 2.2% of the data being adjusted.[4]

All reading time analyses were conducted on reading times adjusted for string length. Reading times for each participant were entered into separate regression analyses with reading time as the dependent variable and string length as the independent variable (see Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994). Residual reading times were calculated by taking the reading time for a word and subtracting off the reading time predicted by the linear regression given the word's string length.

5.2. Results

5.2.1. Question-answering accuracy

On average, participants answered 85% of the questions (on critical trials) correctly, and all participants answered at least 70% of the questions correctly. As mentioned earlier, a randomly chosen third of the comprehension questions were based on the first sentence, a third on the second sentence, and a third on the last sentence. Thus, only the questions about the third sentence are likely to be affected by the factors being tested (Structure and

[4] In addition, due to a programming error, a random 3% of reading times were recorded incorrectly. These data were dropped from further analyses.

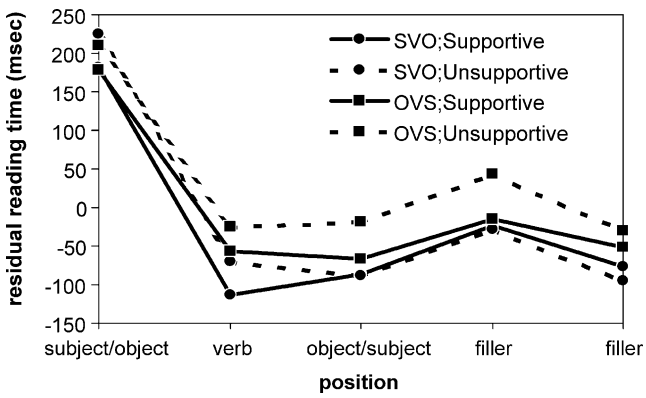

Fig. 1. Mean residual reading times as a function of word position. (Reading times are corrected for string length.)

Context), since the other questions did not vary across conditions. Out of the questions that were about the third (critical) sentence, 95% were correct. Overall, half of all the incorrect answers (53%) were in response to three particular questions, none of which were about the critical third sentence. Thus, it seems that, for independent reasons, these three questions (which were identical in all conditions) were difficult to answer. Given the low number of relevant errors (5%), a breakdown of word-by-word reading times into correct and incorrect question trials was not done.

5.2.2. Word-by-word reading times

Fig. 1 shows residual reading times (adjusted for string length) for each of the five word positions. (Raw reading times are given in Appendix B.) At the first word, the reading times were all fairly slow,[5] with given (i.e. supportive context) nouns (solid lines) read faster than new (i.e. unsupportive context) nouns (dashed lines). Thus, readers appear to be showing immediate sensitivity to the given/new distinction: Nouns that were mentioned in the preceding sentence were read faster than new nouns. However, this pattern may have a lower-level explanation since repeated words are typically read more quickly than unrepeated words (e.g. Traxler, Foss, Seely, Kaup, & Morris, 2000, and references therein). At the verb (word 2), a canonical word order preference appears, such that both SVO conditions (marked with circles) were read faster than the OVS conditions (marked with squares). However, this effect of structure all but disappears for the supportive contexts in the remaining word positions. The typical difficulty with noncanonical structures is reserved for unsupportive contexts.

To test the statistical significance of these patterns, analyses of variances (ANOVAs) were performed on participant and item means at each word position having the following factors: Context (supportive, unsupportive) and Structure (canonical, noncanonical) and the grouping factor (four lists in the participant analysis and four item groups in the item analysis).

[5] The slower reading times at the first word as compared to the other words are a result of the word-by-word presentation mode just having started. Participants are often rather slow at the first word of such a word-by-word self-paced reading task.

At the first noun (word position 1), nouns that had been mentioned in the preceding sentence were read significantly faster than new nouns that were being mentioned for the first time. This effect of Context was significant in the participant analysis and marginal in the item analysis ($F1(1,40) = 8.68$, $P < 0.01$; $F2(1,16) = 3.50$, $P = 0.08$).[6] No other effects or interactions were significant.

At the verb (word position 2), SVO structures were read significantly faster than the OVS structures ($F1(1,40) = 15.18$, $P < 0.01$; $F2(1,16) = 17.85$, $P < 0.01$). A Context effect was also observed: Verbs in supportive contexts were read significantly faster than those in unsupportive contexts ($F1(1,40) = 13.32$, $P < 0.01$; $F2(1,16) = 6.49$, $P < 0.05$). The interaction between Structure and Context was not significant at this word position ($F < 1$). Planned comparisons showed the effect of Structure to be significant in unsupportive contexts ($F1(1,40) = 6.09$, $P < 0.05$; $F2(1,16) = 7.61$, $P < 0.05$) and supportive contexts ($F1(1,40) = 12.17$, $P < 0.01$; $F2(1,16) = 8.18$, $P < 0.05$).

By the second noun (word position 3), processing difficulty appears to be reserved for the noncanonical structure in the unsupportive context. Indeed, although there was a reliable effect of Structure ($F1(1,40) = 20.21$, $P < 0.01$; $F2(1,16) = 7.53$, $P < 0.05$) and a marginal effect of Context ($F1(1,40) = 3.53$, $P = 0.07$; $F2(1,16) = 3.24$, $P = 0.09$), these two factors were found to interact in the participant analysis though this effect was not significant in the item analysis ($F1(1,40) = 4.95$, $P < 0.05$; $F2(1,16) = 2.44$). Planned comparisons showed the effect of Structure to be significant in the unsupportive contexts ($F1(1,40) = 21.72$, $P < 0.01$; $F2(1,16) = 6.79$, $P < 0.05$) but not in the supportive contexts ($F1(1,40) = 1.93$; $F2(1,16) = 0.66$).

At word position 4, processing difficulty again appears to be reserved for the noncanonical structure in the unsupportive context. There was a reliable effect of Structure by participants, marginal by items ($F1(1,40) = 7.00$, $P < 0.05$; $F2(1,16) = 4.06$, $P = 0.06$) and a marginal effect of context by both subjects and items ($F1(1,40) = 2.95$, $P = 0.09$; $F2(1,16) = 3.19$, $P = 0.09$). These two factors again interacted in the participant analysis ($F1(1,40) = 5.83$, $P < 0.05$; $F2(1,16) = 2.36$). And, planned comparisons showed the effect of Structure to be significant in the unsupportive contexts ($F1(1,40) = 7.94$, $P < 0.01$; $F2(1,16) = 11.00$, $P < 0.01$) but not in the supportive contexts ($F < 1$).

A similar statistical pattern is seen at word position 5, where there was a reliable effect of Structure ($F1(1,40) = 11.24$, $P < 0.05$; $F2(1,16) = 11.53$, $P < 0.01$) and no effect of Context ($F < 1$). These two factors were found to interact ($F1(1,40) = 7.42$, $P < 0.05$; $F2(1,16) = 4.41$, $P = 0.05$). And, planned comparisons showed the effect of Structure to be significant in the unsupportive contexts ($F1(1,40) = 18.42$, $P < 0.01$; $F2(1,16) = 18.81$, $P < 0.01$) but not in the supportive contexts ($F1(1,40) = 2.12$; $F2(1,16) = 1.67$). The pattern of means suggests, however, that the reliability of the interaction may in part be due to a reversal of the context effect in SVO sentences as compared to OVS.

[6] There were reliable effects involving control variables (e.g. list and item group) in some of the analyses we conducted in this paper. Because we believe that they have no bearing on the proposals that we will be making, these effects will not be reported.

5.3. Discussion

The results of the self-paced reading study show that that the processing of noncanonical structures is indeed facilitated by the presence of an appropriate discourse context. When a discourse was established that satisfied the referential presuppositions of the noncanonical (OVS) structure, reading times for the noncanonical structure were only slightly longer than for the canonical (SVO) version, with the significant effect being limited to the verb (see solid lines of Fig. 1). When discourse properties did not support these referential presuppositions, as was the case for OVS sentences in unsupportive contexts, considerable difficulty was observed for OVS sentences relative to SVO sentences in similar contexts (dashed lines, Fig. 1).

As we argue here, this reading time pattern supports parsing theories that propose a central influence of both referential and syntactic factors on early parsing procedures, either via simultaneous exertion of these factors or via a mechanism of rapid revision.

First, the data can easily be explained by constraint-satisfaction theories in which both referential and syntactic factors simultaneously influence parsing choices (MacDonald et al., 1994; Trueswell & Tanenhaus, 1994). Consider first the unsupportive contexts, in which the initial noun is discourse-new. This situation is analogous to previous sentence processing studies of OVS and SVO sentences read in isolation, where the initial noun was also discourse-new. Like those studies, a large garden-path effect is observed for OVS sentences as compared to SVO. This effect arises at the verb in OVS because this is where the structure becomes unambiguously noncanonical.[7] Readers showed prolonged difficulty for OVS over the course of the sentence, again consistent with most prior reading time studies of OVS sentences in isolation. In supportive contexts (the solid lines of Fig. 1), the garden-path pattern is partially alleviated, with statistically reliable difficulty for OVS over SVO localized to the verb. This localized difficulty is consistent with the frequency-based complement of these models: Here, a potentially ambiguously marked noun (the O… in OVS) leads to consideration of noncanonical form because the discourse supports this interpretation but transient difficulty is encountered because the parser must battle against the more common canonical alternative. This explanation of the data is in line with the interpretation of syntactic ambiguity resolution in English, where very similar reading time patterns have been observed (Britt, 1994; Spivey & Tanenhaus, 1998). In those studies, contextual factors interacted with the frequency of the structural alternatives. In particular, contextual support of a garden-path sentence reduced but did not completely eliminate processing difficulty when the intended structure was a less common, subordinate syntactic alternative, suggesting that the relative accessibility of parsing alternatives also contributes to ambiguity resolution. The small difficulty we observed in OVS sentences in supportive contexts could be interpreted in a similar way since the noncanonical interpretation must compete with a canonical one. Small accessibility effects of this sort are predicted by multiple constraint theories of parsing

[7] Recall that although the case-markers used on the initial noun of OVS sentences were consistent with the noncanonical OVS structure, many of them also permitted canonical structures, if for instance the verb turned out to be an experiencer verb rather than a simple action verb. In addition, even nouns that had case markers which did not exhibit this particular kind of ambiguity could still be interpreted as being part of multiple structures.

and have been supported by parametric manipulations of lexico-syntactic frequency (Garnsey et al., 1997; Trueswell, 1996).

However, our data can also be explained by the rapid-revision (syntax-first) theory sketched earlier (Frazier & Clifton, 1996; Mitchell, 1987). Under this account, the localized difficulty at the verb in OVS sentences in supportive contexts is seen as a garden-path/revision effect. That is, readers may have initially pursued the canonical SVO parse in this condition, but the context generated a quick rejection of this parse at the verb. Indeed, small bumps in processing difficulty like those found in our supportive context condition have been interpreted in other studies as 'unconscious' garden-path effects within theories of this sort (Frazier & Clifton, 1996). According to this theory, the OVS in the unsupportive context generated a large conscious garden-path because it is only the linguistic material that could disambiguate the parse.

Thus the data rule out theories that assume that most of the difficulty with noncanonical structures is syntactic in nature, since we clearly observe that discourse modulates parsing processes of OVS forms very quickly. Indeed the two competing accounts sketched above both assume a central role for discourse in ongoing interpretation, and both assume that syntactic factors play a brief transient role in reading time difficulty in the presence of contextual support.

Despite these similarities, these two theories make fundamentally different assumptions about the architecture of the comprehension system: one proposes rule-based structure building whereas the other proposes simultaneous constraints acting on the detection of structure. Thus, a small increase in reading times can be explained by radically different systems. This conundrum reflects a debate currently going on in the sentence processing literature over what are appropriate explanations of processing difficulty in theories of parsing. Some researchers view increased reading times as evidence for revision, whereas others view these reading times as evidence for competition/uncertainty. This has even resulted in a situation in which reading studies with very similar stimuli and reading time patterns (Clifton, Traxler, Mohamed, Williams, Morris, & Rayner, 2003; Trueswell et al., 1994) have been interpreted as evidence either for constraint-satisfaction (Trueswell et al., 1994) or rapid revision (Clifton et al., 2003).

There is, however, a third explanation of the data pattern, which we would like to put forth. It may be the case that readers in the OVS supportive condition 'know' they are in an OVS structure but they experience processing difficulty because of it. In particular, the referential context may be guiding readers toward an interpretation that increases predictive processing of upcoming constituents. That is, Finnish readers at OV… in supportive contexts know this is an OVS structure and as a result begin to predict the properties of the upcoming Subject, including the fact that it should be discourse-new. Here the linking assumption would be analogous to one recently put forth by Hale (2003). According to Hale's model of ambiguity resolution, a sudden drop in parsing uncertainty leads to a processing slowdown, precisely because the system has further work to do to specify the representation. In this case, such an account would be compatible with referential, interactive theories of processing because the predictive processing is partially driven by referential factors.

This alternative explanation of the data, which makes a different linking assumption (namely that localized difficulty could arise because the reader is anticipating a shift to

a discourse-new referent) seems worth pursuing and possibly differentiating from other accounts. One effective means for resolving debates over linking assumptions has been to adopt a multiple-methods approach, in which we examine how predictions of a theory hold up against a different measure, a measure that has different and perhaps more transparent linking assumptions. This is the approach we have adopted here. In the next experiment, we examine the comprehension of OVS in spoken Finnish, asking whether listeners predict the upcoming constituent (the S, at OV…) to be discourse-new. The measure will be eye gaze to potential referents in a scene. Here, the linking assumption we adopt is that listeners will look to entities they are currently considering as possible referents.

6. Experiment 2: eye tracking during listening

As discussed above, the self-paced reading experiment shows that discourse context facilitates the processing of scrambled sentences but leaves open the question of what is accounting for the small difficulties observed with OVS sentences in contextually appropriate contexts. In this second experiment, in order to explore this issue, we tested whether participants show incremental sensitivity to the discourse functions of canonical and noncanonical word orders—more specifically, whether people will engage in predictive processing at the point where the word order information they have already encountered is informative about the discourse-newness of upcoming referents.

In order to investigate this question, we took advantage of recent findings from Altmann and Kamide (1999), Kako and Trueswell (2000), and Kamide, Scheepers, Altmann, and Crocker (2002), who have shown that the semantic and syntactic restrictions of verbs can cause listeners to anticipate upcoming referents that have not yet been uttered by the speaker. For example, the semantic restrictions of a verb can induce anticipatory eye movements to objects that are semantically appropriate direct objects even at a point in time when only the subject and the verb have been heard. Altmann and Kamide (1999) compared participants' eye movements when listening to sentences like 'The boy will move the cake' and 'The boy will eat the cake', while viewing a picture of a boy surrounded by a cake and some toys. In the first sentence, the verb does not predict the object in any way, since various objects in the scene can be moved, but in the second sentence, there is a more specific relation between the verb and the object, since only certain kinds of physical objects are edible. Altmann and Kamide found that saccadic eye movements to the object of the verb were launched significantly earlier in the 'eat' condition than in the 'move' condition, prior to hearing "cake".

In light of the self-paced reading results discussed above, this raises the question: Do anticipatory effects arise also on the basis of discourse status as encoded in word order? Recall that OVS order is used in Finnish when the object is old and the subject new, whereas SVO is used when the subject is old and the object is old *or* new. Thus, the OV… configuration *predicts* that the postverbal subject will be new information, whereas the SV… configuration has no such predictive power. In the current study, we wanted to see if this information is used during on-line processing, i.e. if the OVS order induces anticipatory looks to the discourse-new referent, in comparison to the SVO order.

6.1. Method

6.1.1. Participants

Sixteen native Finnish speakers, mainly students at the Helsinki University of Technology, participated in this experiment. Participants received approximately $5 for participation in the experiment.

6.1.2. Procedure

An eye-movement-during-listening paradigm was employed in which participants heard descriptions of clip-art generated pictures (similar to Altmann & Kamide, 1999; Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000). Participants were shown large color pictures of simple scenes involving human or animal characters and heard a brief pre-recorded story about the scene. Participants were told that some of the stories may not match the pictures, and that their task was to correct any errors they notice. Participants were told to do this verbally, i.e. to simply say out loud whether the story was correct or contained a mistake, and if it contained a mistake, to verbally correct it.

A participant's eye movements were recorded using a simple video-based apparatus. In particular, on each trial the participant was presented with a printed color image, which was placed directly below a camcorder that used SONY DVcam digital videotapes with audio-lock recording. The camera was centered above the image, and trained on the participant's face. The pre-recorded sound files were played by a Dell laptop PC over external stereo speakers. The participant's face and eyes, the auditory stimuli, and the participant's vocal responses were all recorded to the DVcam digital videotape. Analysis of the eye movements and speech onsets, described below, were done by hand on the videotapes at a later date, using a SONY DSR-30 digital VCR with jog-shuttle control. This video-based eye gaze process was used because the data were collected in Finland, where neither author had access to a head-mounted eye tracking system.

6.1.3. Materials and design

The visual stimuli for this experiment consisted of large color pictures generated from clip art, printed on 11×16 in. paper from a high-resolution color-ink-jet printer. At a typical viewing distance of 1 m, the visual angle of this scene subtended approximately 22°. Pictures typically contained three to five animate entities (people or animals) and other objects that made up a coherent scene. These images were generated from a large repository of purchased clip-art images and arranged and edited using Adobe Photoshop.

Brief verbal passages were prepared that described a simple story involving the participants shown in the picture. In many cases, the final sentence contained a statement that was inconsistent with the scene, which the participant was to correct (e.g. the man was wearing a hat, when in fact the woman was). The sound files were recorded using the Syntrillium CoolEdit 2002 program on a laptop PC. The same female native Finnish speaker's voice was used for all sound files.

A total of 16 target items (i.e. scene–story pairs) were prepared. These scenes always consisted of three easily identifiable characters, such as the example item seen in Fig. 2a.

(a)

(b)

Fig. 2. Sample item: (a) unambiguous-referent scene; (b) ambiguous-referent scene.

The three characters were of approximately the same size and positioned such that one was to the left side of the image, one was in the middle, and one was to the right. (Pilot testing was done to determine that the characters were far enough apart to enable coders to easily distinguish eye movements to each character from the video record.)

Fig. 2a presents an example of an unambiguous-referent scene, in which three different types of characters were presented (a patient, a doctor, and a nurse). For this scene, the following verbal passage, translated here to English, was provided. (Critical sentences appear in Appendix C.)

(7). *On the hospital reception desk are leaning a doctor and a nurse, and it is almost two o'clock. After a moment,*

(a) *Doctor–subject glances-at patient–object…* (SVO)

(b) *Doctor–object glances-at patient–subject…* (OVS)

This patient is holding a pair of scissors.

In these items, prior to the critical sentence (7a or 7b), two of the characters, namely the doctor and the nurse near the desk, had already been mentioned in the story, and the third character (e.g. the patient), had not been mentioned and, hence, was at that moment discourse new.

The critical sentence had either SVO or OVS order, as illustrated in (7a, 7b). The first, preverbal argument (subject or object) always referred to a discourse-old referent (the doctor), whereas the second, postverbal argument, referred to the previously unmentioned character (the patient). Thus, both the SVO and OVS versions of the critical sentence were felicitous in Finnish. The verbs used in critical sentences were like *glanced at*, in that all had few semantic restrictions on their direct objects and critically permitted animate/sentient direct objects, i.e. a character could glance at something or be glanced at. The critical sentence was followed by a sentence about the postverbal referent. This sentence was always incorrect given the scene, and participants were expected to correct it (saying, e.g. *No, she's holding a rose*).

The ambiguous-referent version of this scene is shown in Fig. 2b. Here, instead of a patient, another nurse is in the display. In this condition, the story context was the same, but the critical target sentences were changed to those in examples (8a) and (8b).

(8). (a) *Doctor–subject glances-at nurse–object…* (SVO)

(b) *Doctor–object glances-at nurse–subject…* (OVS)

This nurse is holding a pair of scissors.

The referent for the second NP in these sentences (*nurse*) was potentially ambiguous, in that it could refer to the discourse-new nurse (on the left) or the discourse-old nurse (on the right). The follow-up sentence was designed to elicit a correction from the participant which would reveal the referent that he/she had assigned to the ambiguous noun (e.g. *No, she's holding a tray*. Or, e.g. *No, she's holding a pencil*.). Position of the new referent was counter-balanced such that on half of the trials the new-referent was on the right, and on half of the trials the new-referent was on the left.

The unambiguous-referent SVO and OVS sound files were created from the ambiguous-referent SVO and OVS sound files by carefully splicing the postverbal noun into the critical sentence. This was done to ensure that the first part of the sentence, [SV…] or [OV…], did not vary across the ambiguous and unambiguous conditions. Participants who were asked about the sound files after participating in the experiment said they had not noticed that some involved splicing.

Thirty-two fillers were also constructed. The fillers were designed to vary in the number of characters in the picture and in whether or not all the characters were mentioned in the story. The characters used in the fillers and the critical items were all different from each other, such that each character was seen only once by a given participant. Most of

the fillers used canonical word order, so as to make sure that the stimuli as a whole match the word order patterns observed by Hakulinen and Karlsson (1979). Twenty-four filler stories had no incorrect descriptions, and eight contained an incorrect description in the final sentence. These numbers were chosen so that there were approximately equal numbers of trials with and without 'mistakes'.

Four presentation lists were constructed by randomly combining the 16 target stories with 32 filler stories. Each target item was separated by at least one filler item. Within a presentation list, eight of the target trials appeared with the SVO structure and eight appeared with the OVS structure. For each of these sentence structure types, four appeared in the unambiguous-referent condition and four appeared in ambiguous-referent condition. Each target item was then rotated through these four conditions, generating four different presentation lists. Reverse order lists were also generated to control for trial order.

6.1.4. Analysis of video record

A native speaker of Finnish went through the audio portion of each videotape and determined the onset of critical target sentences, marking the frame at which they occurred. The video record was analyzed frame-by-frame (without sound) from the marked onset of the critical sentence until the end of the trial. Coding consisted of recording frame-by-frame whether the participant was looking to the left, right, center, or elsewhere. Because the audio was turned off, coders were blind to experimental condition. These data were then used to determine which characters had been fixated over time, relative to the onset of key target words. This eye gaze technique has been used previously in our lab, working quite successfully on both adult and child participants (see Snedeker, Thorpe, & Trueswell, 2001).[8]

6.1.5. Predictions

Let us now consider how we predict people's eye-movements to pattern in the different conditions. In light of the discourse properties of SVO and OVS order, we predict that the OV… configuration will encourage anticipatory looks to the discourse new referent, since postverbal subjects in Finnish are discourse-new. Thus, we predict that, even before participants have heard the postverbal subject noun, they will know, on the basis of OV word order, that it will be a previously unmentioned referent. Moreover, we predict that SV… order will delay consideration of the new referent until the object is heard, because postverbal objects can be old or new information in Finnish and thus the word order is not informative in this respect. In other words, we predict that it is not until participants hear the noun itself that they know what the postverbal object refers to.

For the ambiguous-referent items, we predict that participants' ultimate interpretation of the ambiguous noun will be guided by the word order of the target sentence. We predict that the OV… order will prompt anticipatory looks to the discourse-new referent, in this case the not-yet-mentioned second nurse. Moreover, we predict that upon hearing the ambiguous noun, participants will prefer the discourse-new referent (the new nurse)

[8] To determine the reliability of the eye gaze coding, the video record of three participants was fully double coded. The two scorers were in agreement on over 93% of the video record.

over the discourse-old referent (the already-mentioned nurse). In contrast, the SV… order should result in less consideration of the discourse-new nurse than the OV order, because postverbal objects in Finnish can be old or new, whereas postverbal subjects are new information. It is possible that for the SVO order, participants will prefer the discourse-old nurse simply to maintain discourse continuity, i.e. to continue talking about the same nurse (see e.g. Walker, Joshi, & Prince, 1998, on discourse continuity).

Unambiguous-referent sentences (e.g. with *patient*) should result in sharp increases in looks to the discourse-new referent because it is the only patient in the scene; this should happen only after hearing the noun in the SVO condition, but should happen much earlier in the OVS condition because of anticipatory eye movements.

6.2. Results

6.2.1. Summary of eye-movement patterns

Fig. 3 shows the proportion of looks to the new referent in all four conditions. Consistent with our predictions, ambiguous and unambiguous OVS sentences showed anticipatory eye-movements to the discourse-new referent at the onset of the second noun, in contrast to ambiguous and unambiguous SVO sentences. Crucially, as compared to SVO, OVS sentences showed anticipatory eye-movements to the discourse-new referent (e.g. other nurse/patient) at the very onset of the second noun, prior to a point at which listeners could have phonetically analyzed the second noun. Listeners in the SVO-unambiguous condition did not look to the new referent (the new patient) until well after the word-onset. Also as predicted, the SVO-ambiguous condition shows few looks to

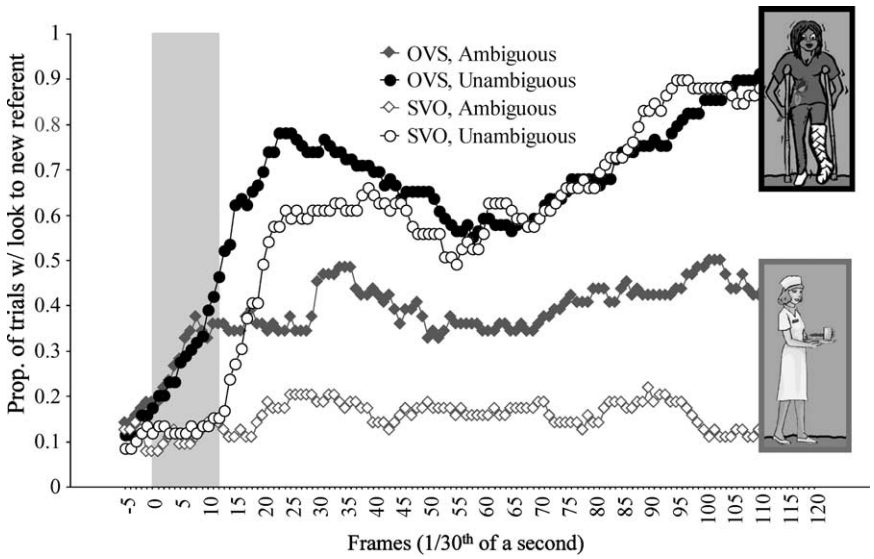

Fig. 3. Proportion of trials with a look to the new referent as a function of time. (The grey box indicates the first 400 ms, starting at the onset of the postverbal noun.)

the discourse-new referent. In contrast, the OVS-ambiguous condition shows substantial looks to the discourse-new nurse, but also later competition with the discourse-old nurse.

6.2.2. Statistical analysis of eye movements

In order to analyze the time-course of these eye movement patterns in more detail, we conducted ANOVAs on five 200 ms time-slices, starting with the onset of the postverbal noun and continuing for 1000 ms. For each time-slice, participant and item means of the proportion of looks to the new referent were entered into separate ANOVAs with four factors: Word order (SVO or OVS), Ambiguity (ambiguous or unambiguous), Order (forward or reverse list) and List (four levels) in the participant analysis and Item Group (four groups) in the item analysis.[9]

During the first two time-slices (0–200 ms; frames 0–5, and 200–400 ms; frames 6–11), there was a main effect of word order (Structure): OVS order prompted significantly more looks to the new referent than SVO order (first time-slice: $F1(1,8) = 10.14$, $P < 0.05$; $F2(1,12) = 9.06$, $P < 0.05$, second time-slice: $F1(1,8) = 18.72$, $P < 0.01$; $F2(1,12) = 12.63$, $P < 0.01$). This statistical pattern means that participants were beginning to launch eye movements to the discourse-new referent in OVS structures even before they could have phonetically analyzed the subject noun. Indeed, if one takes into account the approximate 100–150 ms for programming an eye movement of this sort (e.g. Matin, Shao, & Boff, 1993), eye movements to the discourse-new referent were being programmed near the end of the verb itself.

The strong effect of structure (SVO vs. OVS) persists in the next three time-slices that we analyzed, and we also see an effect of Ambiguity beginning at 400 ms (400–600 ms (frames 12–17) time-slice: Order $F1(1,8) = 30.21$, $P < 0.01$; $F2(1,12) = 17.54$, $P < 0.01$; Ambiguity $F1(1,8) = 16.58$, $P < 0.01$; $F2(1,12) = 7.55$, $P < 0.05$). The effect of Ambiguity arose because participants were more likely to look at the discourse-new referent in the unambiguous than in the ambiguous condition.

The observation that we get effects of both structure and ambiguity starting at 400 ms can be attributed to the fact that, at this point in time, participants have enough phonetic information about the postverbal noun to be able to look at the referentially unique, discourse-new referent in the unambiguous condition (e.g. the patient). The timing of this pattern for the SVO condition is analogous to that seen in word recognition studies using this technique (Allopenna, Magnuson, & Tanenhaus, 1998).

In the last two time-slices (600–800 ms; frames 18–23, and 800–1000 ms; frames 24–29), the strong effects of structure and ambiguity continue (600–800 ms: Order $F1(1,8) = 18.33$, $P < 0.01$; $F2(1,12) = 9.54$, $P < 0.01$; Ambiguity $F1(1,8) = 30.09$, $P < 0.01$, $F2(1,12) = 39.56$, $P < 0.01$; 800–1000 ms: Order $F1(1,8) = 21.255$, $P < 0.01$, $F2(1,12) = 9.58$, $P < 0.01$; Ambiguity $F1(1,8) = 42.19$, $P < 0.01$, $F2(1,12) = 48.38$, $P < 0.01$).

[9] Throughout the paper, whenever ANOVAs were conducted on proportions, identical ANOVAs were also conducted on an arcsin transformation of the data, arcsin $((2 \times p) - 1)$. This was done to adjust for the fact that the proportion p is bounded at 0 and 1. Unless otherwise noted, statistically significant effects that were found for the untransformed data were also significant for the transformed data.

6.2.3. Latency of looks

Although Fig. 3 is extremely informative, it cannot tell us about certain important aspects of the eye movement data, such as the mean latency with which participants moved their eyes to the discourse-new referent and even the proportion of participants who eventually did make such an eye movement. For instance, the OVS-ambiguous curve could reflect a situation in which roughly 40% of the participants moved their eyes to the discourse-new nurse and then held gaze there, or it could reflect a situation in which nearly all participants looked to the discourse-new nurse at different points in time—with no more than 40–50% ever fixating the nurse at the same moment. (The same is even true for SVO-ambiguous; it is possible that all participants did eventually look at the discourse-new nurse, but only 10–20% of them did so at any given moment.)

In order to get a better picture of how participants contributed to the proportion curves in Fig. 3, we computed the mean latency (and standard error) of first looks to the discourse-new referent and the proportion of participants who made such a look (see Table 1). Because of the anticipatory nature of the looks shown in Fig. 3, we began coding for this first look 200 ms *before* the onset of the critical noun, when relatively few participants were looking at the discourse-new referent and the rates were equal across the conditions (hence the latency 'timer' began 200 ms prior to the noun onset). Latencies reflect only trials in which participants were not looking at the discourse-new referent when the 'timer' had started and had eventually made an eye movement to the discourse-new referent later in the trial. The percentage of trials of this sort (i.e. trials where participants moved their eyes to the discourse-new referent) appears to the right in Table 1. For the sake of completeness, we also provide in parentheses the percentage of trials in which participants looked at the discourse-new referent from this onset time even if they were already fixating the discourse-new referent when the 'timer' had started.

As can be seen in the table, unambiguous-referent trials resulted in participants nearly always launching an eye movement to the discourse-new referent. This is not surprising because the noun itself uniquely specifies this referent. In addition, however, the mean latency of these eye movements is considerably and reliably faster in the OVS condition as compared to the SVO condition ($F1(1,8) = 24.57$, $P < 0.01$; $F2(1,12) = 5.94$, $P < 0.05$). This is because of the anticipatory eye movements shown in Fig. 3.

Table 1
Mean latency to look to the discourse-new referent, and the SE

Condition	Mean	SE	Proportion
Unambiguous-referent			
OVS	513	77	0.87 (0.97)
SVO	853	103	0.92 (1.0)
Ambiguous-referent			
OVS	917	117	0.83 (0.95)
SVO	1430	196	0.50 (0.61)

Proportion of participants who looked to the discourse-new referent (including trials where participants were already looking at the entity, in parenthesis). Latencies and proportions were computed from 200 ms prior to onset of critical noun.

Also shown in the table, ambiguous-referent conditions show large differences in the number of looks to discourse-new referent. Participants only looked at the discourse-new referent (the new nurse) on half the trials in SVO constructions, whereas OVS resulted in 83% looks. And, although on a smaller sample, we still see faster eye movements in the OVS structure. Interestingly, Ambiguous-referent latencies are generally much longer than Unambiguous-referent latencies, a point we return to later. (Statistical tests were not done on the RTs involving Ambiguous-referent items because of the lack of sufficient data in the SVO-Ambiguous-referent condition.)

6.2.4. Looks to discourse-old referents

Although not plotted in the figure, the pattern of looks to the discourse-old entities is in line with observations and interpretations drawn from the look-to-new data. In particular, during the early processing of the postverbal noun (0–400 ms, frames 0–11), listeners of SVO sentences showed a preference to look to the discourse-old character over the discourse new character, whereas OVS sentences generated the opposite pattern. For SVO in this time window, the proportion of time spent looking at the discourse-old character was 0.20 as compared to only 0.12 for the discourse-new character. For OVS, it was 0.28 for the discourse-new character and 0.20 for the discourse-old character. Thereafter, the Unambiguous stimuli of course generated large numbers of looks to the discourse-new entity regardless of structure because the noun unambiguously referred to that character (e.g. patient). For Ambiguous nouns, however (e.g. nurse), SVO had approximately three times as many looks to the *discourse-old* entity than the *discourse-new* entity, whereas OVS-Ambiguous generated roughly equal looks to both new and old. In other words, SV… structures generate an expectation/assumption by listeners that the discourse will *continue* with discussion of the same character whereas OV… structures generate the expectation/assumption that there will be a *shift* to a discourse-new entity.

6.2.5. Off-line referential judgments

Participants' off-line referential judgments matched the eye-movement patterns. As mentioned earlier, participants were told that some of the stories they hear (in fact all the critical items) may not match the pictures they see, and they need to correct the story where necessary. The participants' corrections of sentences such as 'This nurse/patient is holding a pair of scissors' thus provided an additional measure of how they interpret the postverbal nouns, since we can use the corrections (e.g. *No, she is holding a rose*) to determine which character the participant interpreted the postverbal noun in the target sentence as referring to. In the SVO ambiguous condition, only 11% of the responses[10] treated the postverbal object as referring to the new referent (e.g. the other nurse), whereas in the OVS ambiguous condition, 37% treated the postverbal subject as the new referent. In the SVO unambiguous condition, 98% interpreted the postverbal object (e.g. the patient) as the new referent, and in the OVS unambiguous condition, 95% did so.

[10] Some of the participants' corrections (about 9%) were ambiguous (e.g. 'No, she is not holding scissors') and did not distinguish between the two possible referents. The percentages reported here exclude such ambiguous responses. The basic pattern of percentages does not change if the ambiguous responses are included.

In the unambiguous condition, the postverbal noun (e.g. the patient) was not lexically ambiguous, and we would in fact expect all responses to treat it as new information.

6.3. Discussion

Overall, the results discussed above support the prediction that people use word order patterns to predict upcoming referents on the basis of discourse status. The analysis of the eye-movements shows that both ambiguous and unambiguous OVS sentences, in contrast to SVO sentences, prompt anticipatory eye-movements to the discourse-new referent at the onset of the second noun (during the first time-slice, 0–200 ms). Given that it takes about 100–150 ms to program an eye-movement (Matin et al., 1993), the presence of an effect of SVO vs. OVS order during the first time-slice shows that people are using the discourse-information encoded in the noncanonical OVS order to make predictions about the upcoming postverbal subject before they even hear that subject.

In contrast, SVO sentences do not induce anticipatory looks to the discourse-new referent, which makes sense in light of the observation that they lack the discourse-based predictive power of OVS order. In Finnish, a postverbal object can be old or new information, and thus a listener who has heard SV… cannot use word order to make any specific predictions about the discourse status of the upcoming object. Thus, listeners in the SVO-unambiguous condition do not look to the new referent (the new patient) until well after the word onset, i.e. after they have heard enough to recognize the word. Moreover, in the SVO-ambiguous condition, when they do hear the postverbal noun, listeners are more likely to look at the discourse-*old* nurse than at the discourse-*new* nurse—even though the word 'nurse' could refer to either nurse perfectly well. In fact, the data suggest that listeners anticipated that an SV… sequence would continue to talk about discourse-old entities whereas an OV… would shift to discussing a discourse-new entity.

It is well-known that maintaining discourse continuity (in the sense of continuing to talk about the same entity) is less costly than shifting the focus of attention to a new entity (e.g. Walker et al., 1998). SVO order allows people to follow this bias for discourse continuity, whereas OVS order suggests that the subject is new, and thus goes against the discourse continuity bias and supports the addition of a new referent to the discourse model. According to this view, the anticipatory looks triggered by SVO and OVS differ in that those in the SVO conditions reflect people's desire for discourse continuity, whereas the anticipatory looks in the OVS conditions are prompted by the referential properties of OVS order. In contrast to SVO the OVS-ambiguous condition does show substantial anticipatory looks to the discourse-new nurse, but also substantial competition later with the discourse-old nurse, due to the ambiguous lexical cue. The difference between ambiguous OVS and unambiguous OVS—i.e. that there is a higher proportion of looks to the new referent in the unambiguous condition than in the ambiguous condition—is attributed to the absence of a clear lexical cue in the OVS ambiguous case.

Indeed, the difference between the ambiguous and unambiguous OVS versions suggests that word-order is one of several probabilistic determinants of discourse-status in Finnish; listeners do not treat OVS order as requiring the postverbal referent to be new, as also can be seen in the off-line judgments. We strongly suspect that participants were behaving in a Gricean fashion, reasoning that the speaker would have most likely said

"another nurse" if she had meant to refer to the discourse-new nurse (i.e. "The doctor glanced at another nurse."). In fact, Finnish has a quantifier very much like the English quantifier '(an)other': *toinen* 'other'. The preference to use such quantifiers both in English and Finnish suggests that both languages' grammatical cues to given-new status (articles in English, word order in Finnish) only partially predict discourse status, which under potentially ambiguous situations can get further linguistic support from the use of quantifiers (see Sedivy, 2002).

The pattern of anticipatory eye movements in the OVS conditions can be captured with a constraint-satisfaction model in which referential and contextual factors play an immediate role. Just as Altmann and Kamide's (1999) results showed anticipation based upon semantic restrictions, our findings illustrate that anticipation also arises on the basis of discourse-status information, as encoded by word order, even when this order has been treated as structurally complex. It is important to note that OVS order, which has often been claimed to be significantly harder to process than SVO order, is in some sense more helpful to the processor than SVO order.

It must be noted that a syntax-first account also exists for this data, but the theory would have to assume very rapid revision of structure. During the perception of the verb, the initial parse (as SV…) would need to be rejected, and the new parse (as OV) would need to be computed along with the discourse implications necessary to predict the upcoming discourse-new S. This assumes a theory of revision that can happen during the perception of the verb, and that this disambiguation allows for simultaneous computation of discourse properties of the alternative parse. If such a position is adopted, however, the notion of an encapsulated initial syntactic stage becomes less clear because the stage's proposals can be rejected as soon as they are made. Similarly though, it is not entirely satisfying to say that predictive processing of this sort arises during the *competition* between SVO and OVS structures, which would be assumed under the constraint-satisfaction approach. We address this issue in Section 7 where the findings of Experiments 1 and 2 are integrated.

7. General discussion

We have presented two experiments that investigate the processing of canonical and noncanonical word orders in Finnish. The results of the first experiment revealed that processing difficulty associated with noncanonical structures could be partially mitigated by discourse contexts that support the referential function of such structures. In particular, processing difficulty for noncanonical OVS order was limited to the verb when the object was discourse-old and the subject was discourse-new. The findings of the second study show that comprehenders make incremental, efficient use of the discourse-information encoded in Finnish OVS order. Listeners who heard a sentence-initial discourse-old object followed by a verb showed signs of anticipating that the upcoming subject will be discourse-new. No such anticipation for the discourse-new referent was found for SV… structures. Even referentially ambiguous nouns showed sensitivity to this word-order cue, with listeners being more likely to infer the referent of the ambiguous noun to be the discourse-new entity when it followed an OV… sequence as compared to a SV… sequence.

The relationship between the findings of Experiments 1 and 2 raises an important issue regarding the interpretation of local slowdowns during processing. The finding that there was a slowdown at the verb with felicitous OVS orders (Experiment 1), combined with the finding that the verb in OVS order is the point at which listeners engage in predictive processing of a shift to discourse-new referents (Experiment 2), suggests that localized processing difficulty could in principle be related to reductions in ambiguity and the need to specify further an interpretation (in the sense of Hale, 2003). This proposal shares some properties with Gibson's (1998) Syntactic Prediction Locality Theory (SPLT) in which slowdowns can arise from the positing of upcoming structure. We would argue that this complexity is not simply a reflection of syntactic complexity but referential complexity as well: listeners do more than predict syntactic structure; they predict the referential properties of that structure. In other words, establishing a new referent would induce localized RT slowdowns in reading (Experiment 1) but manifest itself as an anticipatory shift in attention to the discourse new character (Experiment 2).

This proposal would be a departure from a particular assumption present in many constraint-satisfaction theories of parsing, namely that uncertainty increases processing difficulty and certainty decreases it. Here the opposite is proposed: certainty (about the discourse-newness of upcoming referents) increases processing load because it provides opportunity for further processing, even predictive processing. Patterns of local difficulty in subordinate (nondominant) structures in English parsing studies might reflect the very sort of processing we see here, e.g. the biasing of a relative clause over a main clause interpretation may lead to transient difficulty because of localized referential processing. Indeed, Trueswell et al. (1994) did observe some small transient difficulty for unambiguous reduced relative clauses ("The cake eaten by the family was…") as compared to unambiguous unreduced relatives ("The cake that was eaten by the family was…"), which would suggest such an account.

Clearly, further thinking and observations on this matter are necessary to understand exactly which conditions generate these kinds of patterns. However, it does seem that neither a rapid revision/syntax-first approach nor a constraint-based model offers a fully satisfying way of capturing the data patterns we see in Experiments 1 and 2. As mentioned earlier, a syntax-first approach would have to assume a very fast revision and recomputing stage, which raises questions about the need for an encapsulated syntax-first stage. From a constraint-satisfaction perspective, however, it seems surprising that a condition that is hypothesized to have competition between the dominant (SVO) and subordinate (OVS) structure should trigger anticipatory processing. The third option—based on Hale's (2003) hypothesis that a sudden drop in parsing uncertainty leads to a processing slowdown because the system has further work to do to specify the representation—seems to offer the best account of the data presented here. To be clear, this explanation requires a central role for discourse/pragmatic factors in ongoing interpretative processes and hence is most amenable to interactive accounts of processing in which multiple evidential sources contribute to on-line parsing decisions.

In closing, we reiterate that our present findings tell us a great deal about information use and structure building during language comprehension: Discourse factors that determine structural considerations in languages with free word order can lead listeners to

have expectations about the structure and discourse status of upcoming constituents. Experiment 1 showed that when the presuppositions are met regarding the referential functions of noncanonical structures, processing difficulty could be substantially reduced. Experiment 2 showed that Finnish listeners are sensitive to the discourse information carried by word order and in fact use this information to make predictions about the discourse status of upcoming referents. Thus we believe it is important for theories of language comprehension to take into account the discourse-functions of syntactic structures, since it seems very likely that, during on-line processing, hypotheses regarding an utterance's discourse function are made *in tandem* with hypotheses regarding the utterance's propositional content.

Acknowledgements

We are grateful to Ellen Prince, Lila Gleitman and the "Cheese" seminar participants for helpful advice during the development of this research. We also thank Irina Sekerina for assistance with the DMASTR program used in Experiment 1, and Kelly Rulison for assistance in the stimulus preparation and coding of Experiment 2. This research was partially supported by a grant from the National Institutes of Health (1-R01-HD37507) to the second author. Elsi Kaiser is now at the Center for Language Sciences at the University of Rochester.

Appendix A. Experimental materials for Experiment 1

For the first item, we present the full experimental paradigm. For the remaining critical items, only the SVO version of the target sentence is given.

1. Lotta etsi eilen sieniä metsässä. Hän huomasi heinikossa hiiren/jäniksen joka liikkui varovasti eteenpäin. Hiiri seurasi jänistä ja linnut lauloivat/Jänistä seurasi hiiri ja linnut lauloivat.
2. Mies katseli naista ja aurinko paistoi.
3. Poro puri lammasta ja ihmiset yllättyivät.
4. Kana näki hevosen kun tallin ovet avattiin.
5. Kissa tapasi suden ja muut metsäneläimet ilahtuivat.
6. Sotilas jutteli nuorukaiselle ja autot ajoivat heidän ohitseen.
7. Isä jutteli äidille ja illallinen oli valmiina pöydällä.
8. Pappi puhui opiskelijalle kun maisemat vilisivät ohi.
9. Koira tuijotti lintua ja pari lehteä kahisi tuulessa.
10. Lehmä rakastui karhuun ja sadun muut eläimet kauhistuivat.
11. Tyttö halasi poikaa ja sitten bussi saapui.
12. Orava tarkasteli kaniinia ja pari pulua lenteli lähistöllä.
13. Mies suuteli naista juuri kun aurinko meni pilveen.
14. Turisti puhui koululaiselle kun autot hurahtelivat ohi.

15. Lääkäri hymyili sairaanhoitajalle mutta sitten kuului vauvan itkua.
16. Pöllö seurasi yökköä ja pari varista lensi niiden takana.
17. Hanhi nokki sorsaa kun eläimille tuotiin ruokaa.
18. Tiikeri pakoili leijonaa ja kaikki yllättyivät.
19. Merja oli kertonut siitä Ollille koulussa.[11]
20. Jaakko kompastui Hannuun ja muut kauhistuivat.

Appendix B. Word-by-word raw reading times (in milliseconds) for each of the five word positions (Experiment 1)

Order	Context	Position	Mean RT (ms)
SVO	Supportive	Subject	656
SVO	Supportive	Verb	374
SVO	Supportive	Object	416
SVO	Supportive	Filler	433
SVO	Supportive	Filler	404
SVO	Unsupportive	Subject	701
SVO	Unsupportive	Verb	417
SVO	Unsupportive	Object	411
SVO	Unsupportive	Filler	427
SVO	Unsupportive	Filler	385
OVS	Supportive	Object	673
OVS	Supportive	Verb	432
OVS	Supportive	Subject	409
OVS	Supportive	Filler	445
OVS	Supportive	Filler	431
OVS	Unsupportive	Object	709
OVS	Unsupportive	Verb	460
OVS	Unsupportive	Subject	458
OVS	Unsupportive	Filler	498
OVS	Unsupportive	Filler	448

[11] One of the critical items involved a ditransitive verb, and the constituents that occur in the first five word positions in the noncanonical version of this item do not fully match the noun–verb–noun configuration of the other items. An additional set of analyses with this item excluded did not affect the results in any significant way and does not change the conclusions drawn from the experiment.

Appendix C. Experimental materials for Experiment 2

For the first item, we present the full paradigm. For the remaining critical items, we present only the SVO version of the critical sentence with the nouns used in the ambiguous and unambiguous versions.

1. Sairaalan vastaanottotiskiin nojailevat lääkäri ja sairaanhoitaja, ja kello on jo melkein kaksi. Hetken päästä lääkäri katsahtaa sairaanhoitajaan/potilaaseen//lääkäriin katsahtaa sairaanhoitaja/potilas. Tämä sairaanhoitaja/potilas pitää kädessään saksia.
2. Sitten nainen tervehtii miestä/lasta.
3. Hetken päästä tyttö vilkaisee poikaa/vauvaa.
4. Yhtäkkiä palomies virnistää maanviljelijälle/lehtimiehelle.
5. Yhtäkkiä balettitanssija nauraa valokuvaajalle/kokille.
6. Hetken päästä kalastaja mulkoilee mekaanikkoa/kapteenia.
7. Hetken päästä kapellimestari hymyilee viulunsoittajalle/rumpalille.
8. Yhtäkkiä golfinpelaaja virnistää tenniksenpelaajalle/ohikulkijalle.
9. Hetken päästä kuningatar vilkaisee merirosvoa/noitaa.
10. Yhtäkkiä taidemaalari nauraa turistille/museovartijalle.
11. Sitten opettaja hymyilee oppilaalle/kirjastonhoitajalle.
12. Sitten salapoliisi tervehtii laboranttia/siivoojaa.
13. Sitten lentäjä moittii puutarhuria/lentoemäntää.
14. Sitten pelle katsahtaa jonglööriin/tulennielijään.
15. Yhtäkkiä jääkiekkoilija mulkoilee luistelijaa/hiihtäjää.
16. Sitten liikemies moittii sihteeriä/pikalähettiä.

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*(4), 419–439.

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

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

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.

Bader, M., & Meng, M. (1999). Subject–object ambiguities in German embedded clauses: An across-the-board comparison. *Journal of Psycholinguistic Research*, *28*, 121–144.

Bailyn, J (1995). *A configurational approach to Russian 'free' word order*. PhD dissertational, Cornell University.

Baker, M. C. (2001). *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.

Birner, B., & Ward, G. (1998). *Information status and noncanonical word order in English*. Amsterdam: John Benjamins.

Bornkessel, I., Schlesewsky, M., & Friederici, A. (2003). Contextual information modulates initial processes of syntactic integration: The role of inter-versus intrasentential predictions. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29*, 871–882.

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.

Chesterman, A. (1991). *On definiteness*. Cambridge: Cambridge University Press.

Clifton, C., Jr, Traxler, M. J., Mohamed, M. T., Williams, R., Morris, R., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language*, *49*(3), 317–334.

Crain, S (1980). *Pragmatic constraints on sentence comprehension*. Unpublished PhD dissertation, UC Irvine.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context in the psychological syntax processor. In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language parsing* (pp. 320–358). Cambridge: Cambridge University Press.

De Hoop, H (1992). *Case configuration and noun phrase interpretation*. Doctoral dissertation, University of Groeningen.

De Vincenzi, M. (1991). *Syntactic parsing strategies in Italian. The Minimal Chain Principle*. Dordrecht: Kluwer.

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

Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 555–568.

Frazier, L., & Clifton, C. (1996). *Construal*. Cambridge, MA: MIT Press.

Frazier, L., & Flores d'Arcais, G. B. (1989). Filler-driven parsing: A study of gap-filling in Dutch. *Journal of Memory and Language*, *28*, 331–334.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*, 291–325.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, *37*, 58–93.

Gennari, S., & Poeppel, D. (2003). Processing correlates of lexical semantic complexity. *Cognition*, *89*(1), B27–B41.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.

Givón, T. (1984) (*Vol. 1*). *Syntax. A functional–typological introduction*, Amsterdam: John Benjamins.

Gourley, J., & Catlin, J. (1978). Children's comprehension of grammatical structures in context. *Journal of Psycholinguistic Research*, *7*(6), 419–434.

Hakulinen, A., Karlsson, F (1979). *Kvantitatiivinen tutkimus suomen morfosyntaksista ja sen tekstuaalisista tekijöistä*. Ms, University of Turku.

Hakulinen, A., & Karlsson, F. (1988). *Nykysuomen lauseoppia*. Helsinki: Suomalaisen Kirjallisuuden Seura.

Hakulinen, A., Karlsson, F., & Vilkuna, M. (1980). *Suomen tekstilauseiden piirteitä: Kvantitatiivinen tutkimus [Features of Finnish text sentences: A quantitative study]*. Helsinki: Department of General Linguistics, University of Helsinki.

Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research*, *32*(2), 101–123.

Hemforth, B. (1993). *Kognitives parsing: repräsentation und verarbeitung sprachlichen wissens*. St. Augustin, Germany: Infix.

Hsiao, F., & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition*, *90*, 3–27.

Hyönä, J., & Hujanen, H. (1997). Effects of case marking and word order on sentence parsing in Finnish: An eye fixation analysis. *Quarterly Journal of Experimental Psychology*, *50A*(4), 841–858.

Ishihara, S (2001). *Stress, focus and scrambling in Japanese*. MIT Working Papers in Linguistics 39, A few from Building E39.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*(2), 228–238.

Kaan, E. (2001). Subject–object order ambiguities and the nature of the second NP. *Journal of Psycholinguistic Research*, *30*(5), 527–545.

Kaiser, E (2000). *The discourse functions and syntax of OSV word order in Finnish*. Papers from the Regional Meeting of the Chicago Linguistic Society 16 (pp. 179–194).

Kako, E., & Trueswell, J. C. (2000). Verb meanings, object affordances, and the incremental restriction of reference. *Proceedings of the Annual Conference of the Cognitive Science Society*, *22*, 256–261.

Kamide, Y., & Mitchell, D. C. (1999). Incremental pre-head attachment in Japanese parsing. *Language and Cognitive Processes*, *14*(5/6), 631–662.

Kamide, Y., Scheepers, C., Altmann, G. T. M., & Crocker, M (2002). *Integration of syntactic and semantic information in predictive processing: Anticipatory eye-movements in German*. Talk given at the 15th Annual CUNY Conference on Human Sentence Processing, New York, March 2002.

King, T (1993). *Configuring topic and focus in Russian*. PhD dissertation, Stanford University.

Lambrecht, K. (1994). *Information structure and sentence form*. Cambridge: Cambridge University Press.

Lenerz, J. (1977). *Zur Abfolge nominaler Satzglieder im Deutschen (Studien zur deutschen Grammatik 5)*. Tübingen: Niemeyer.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, *9*(2), 157–201.

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

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information processing time with and without saccades. *Perception & Psychophysics*, *53*(4), 372–380.

McElree, B., Traxler, M. J., Pickering, M. J., Seely, R. E., & Jackendoff, R. (2001). Reading time evidence for enriched composition. *Cognition*, *78*(1), B17–B25.

Mitchell, D. (1987). Lexical guidance in human parsing: Laws and processing characteristics. In M. Colthear (Ed.), *Attention and performance XII: The psychology of reading*. London: Erlbaum.

Miyamoto, E., & Takahashi, S. (2002). Sources of difficulty in processing scrambling in Japanese. In M. Nakayama (Ed.), *Sentence processing in East-Asian languages*. Stanford, KY: CSLI.

Nakayama, M., & Lewis, R. L. (2000). Similarity interference and scrambling in Japanese. *Korean Journal of Cognitive Science*, *12*, 39–53.

Olson, D., & Filby, N. (1972). On the comprehension of active and passive sentences. *Cognitive Psychology*, *3*, 361–381.

Pinto, M (1997). Licensing and interpretation of inverted subjects in Italian. UiL OTS Dissertation series, Utrecht.

Prince, E. F. (1992). The ZPG letter: subjects, definiteness, and information-status. In S. Thompson, & W. Mann (Eds.), *Discourse description: Diverse analyses of a fund raising text* (pp. 295–325). Amsterdam: John Benjamins.

Prince, E. F. (1998). On the limits of syntax, with reference to Topicalization and Left-Dislocation. In P. Culicover, & L. McNally (Eds.), *Syntax and semantics. The limits of syntax 29* (pp. 281–302). New York: Academic Press.

Prince, E. F (1999). *How not to mark topics: 'Topicalization' in English and Yiddish*. Ms, University of Pennsylvania.

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.

Ross, J (1967). *Constraints on variables in syntax*. PhD dissertation, MIT, Cambridge, MA.

Saito, M (1985). *Some asymmetries in Japanese and their theoretical implications*. PhD dissertation, MIT.

Saukkonen, P., Haipus, M., Niemikorpi, A., & Sulkala, H. (1979). *Suomen kielen taajuussanasto [Frequency dictionary of Finnish]*. Porvoo: WSOY.

Sedivy, J. C. (2002). Invoking discourse-based contrast sets and resolving syntactic ambiguities. *Journal of Memory and Language*, *46*(2), 341–370.

Sekerina, I. (2003). Scrambling and processing: Dependencies, complexity and constraints. In S. Karimi (Ed.), *Scrambling and word order*. Malden, MA: Blackwell.

Snedeker, J., Thorpe, K., & Trueswell, J. (2001). *On choosing the parse with the scene: The role of visual context and verb bias in ambiguity resolution. Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Lawrence Erlbaum.

Spivey, M., & Tanenhaus, M. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1521–1543.

Spivey-Knowlton, M., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, *55*(3), 227–267.

Spivey-Knowlton, M., Trueswell, J., & Tanenhaus, M. (1993). Context effects in syntactic ambiguity resolution: Discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology*, *47*, 276–309.

Stojanović, D. (1998). Theta-role assignment in on-line processing of a free word order language. In Z. Bošković, S. Franks, & W. Snyder (Eds.), *The Sixth Annual Workshop on Formal Approaches to Slavic Linguistics. The Connecticut Meeting 1997* (pp. 287–302). Ann Arbor, MI: Michigan Slavic Publications.

Tada, H (1993). *A/A′ partition in derivation*. PhD dissertation, MIT.

Traxler, M. J., Foss, D. J., Seely, R. E., Kaup, B., & Morris, R. K. (2000). Priming in sentence processing: Intralexical spreading activation, schemas, and situation models. *Journal of Psycholinguistic Research*, *29*(6), 581–595.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, *35*, 566–585.

Trueswell, J. C., & Tanenhaus, M. K. (1994). Towards a lexicalist framework for constraint-based syntactic ambiguity resolution. In C. Clifton, K. Rayner, & L. Frazier (Eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum.

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

Vallduví, E. (1992). *The informational component*. New York: Garland Press.

Vallduví, E., & Vilkuna, M. (1998). On rheme and kontrast. In P. Culicover, & L. McNally (Eds.), *The limits of syntax, syntax and semantics 29* (pp. 79–108). New York: Academic Press.

Vasishth, S (2002). *Working memory in sentence comprehension: Processing Hindi center embeddings*. Unpublished Ph.D. Dissertation, http://www.ling.ohio-state.edu/~vasishth/PhD/

Vilkuna, M. (1989). *Free word order in Finnish. Its syntax and discourse functions*. Helsinki: Finnish Literature Society.

Vilkuna, M. (1995). Discourse configurationality in Finnish. In K. Kiss (Ed.), *Discourse configurational languages* (pp. 244–268). New York: Oxford University Press.

Walker, M. A., Joshi, A. K., & Prince, E. F. (1998). *Centering theory in discourse*. New York: Oxford University Press.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, *85*(1), 79–112.

Webelhuth, G (1989). *Syntactic saturation phenomena and the modern Germanic languages*. PhD dissertation, Umass Amherst.

Weskott, T (2003). *Information structure as a processing guide*. PhD dissertation, University of Leipzig.

Yokoyama, O. (1986). *Discourse and word order*. Amsterdam: John Benjamins.