Processing the NP- versus S-Coordination Ambiguity: Thematic Information Does not Completely Eliminate Processing Difficulty.

John C. J. Hoeks
Center for Language and Cognition Groningen and
BCN NeuroImaging Center, University of Groningen, The Netherlands
(j.c.j.hoeks@rug.nl)

Petra Hendriks
Center for Language and Cognition Groningen, University of Groningen, The Netherlands
(p.hendriks@rug.nl)

Wietske Vonk
Department of Linguistics, University of Nijmegen and
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
(vonk@mpi.nl)

Colin M. Brown
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

and Peter Hagoort
FC Donders Centre for Cognitive NeuroImaging, Nijmegen, The Netherlands
(peter.hagoort@fcdonders.kun.nl)

Address correspondence to:
Dr. John C.J. Hoeks
CLCG / BCN NeuroImaging Center
University of Groningen
PO Box 716, 9700 AS Groningen, The Netherlands
e-mail: j.c.j.hoeks@rug.nl
fax: +31 (0)50 3636855

Abstract

When faced with the NP- versus S-coordination ambiguity as in, for example, *The thief shot the jeweler and the cop* ..., readers prefer the reading with NP-coordination (e.g., "The thief shot the jeweler and the cop yesterday") over one with two conjoined sentences (e.g., "The thief shot the jeweler and the cop panicked"). A corpus study is presented showing that NP-coordinations are produced far more often than S-coordinations, which in frequency-based accounts of parsing might be taken to explain the NP-coordination preference. However, frequency differences only become apparent when coarse-grained measures are considered. We also describe an eye tracking experiment investigating how temporarily ambiguous S-coordinated sentences are actually processed, and, more importantly, how information regarding thematic fit is used by the parser in this kind of structure. The S-coordinated sentences such as *Jasper sanded the board and the carpenter laughed* that were tested in the present experiment, were predicted to cause only minimal processing difficulty, if any, because of the poor thematic fit between *carpenter* and *sanded* arguing against NP-coordination. Our results indicate that information regarding poor thematic fit was used rapidly, but not without leaving some residual processing difficulty. This is compatible with claims that thematic information can reduce but not completely eliminate garden-path effects.

KEYWORDS: sentence processing, coordination, corpus study, eye tracking

Introduction

Language users, be they readers or listeners, have to keep up with rapidly, sequentially delivered language input. This requires the human sentence processor to deal with ambiguity fast and effectively, which it succeeds to do remarkably. More than two decades of research into the mechanism by which ambiguity, and especially syntactic ambiguity, is resolved has not led to much consensus, but there have been important advances in the identification of the factors that play a role in the processing of ambiguous structures. Controversy remains, though, over whether these factors play a role before, during, or only after the actual resolution of syntactic ambiguity. The eye-tracking experiment presented in this paper will not adjudicate between the two processing frameworks that are still the most prominent and influential: garden-path theory and constraint-based models (see, e.g., Rayner & Clifton, 2002, for an excellent overview). Instead, our aim is to find out more about the time-course with which different sources of information are used by the on-line parser, while being relatively agnostic with respect to theoretical paradigm. The syntactic ambiguity that will be the focus of this paper is the NP versus S-coordination ambiguity. Consider, for instance, sentence 1a.

1a. The thief shot the jeweler and *the cop* ...

1b. The thief shot [the jeweler and *the cop*] during a robbery.

1c. [The thief shot the jeweler] and [*the cop* panicked].

When the NP *the cop* is read in (1a), it is unclear whether it should be read as part of the direct object of the verb *shot* as in (1b), or as the subject of a subsequent verb, as in (1c). In the first case *the cop* has to be conjoined with *the jeweler* into a complex NP (NP-coordination); in the latter case it becomes the subject of a conjoined sentence (S-coordination).

The first study looking into how readers resolve the NP/S-coordination ambiguity was reported by Frazier (1987a). She conducted a segment-by-segment self-paced reading experiment, with the last segment of the experimental stimuli disambiguating towards either NP- or S-coordination, as in 2a and 2b, respectively. Reading times for the final segment are added in parentheses; the literal translations of the Dutch materials are shown in apostrophes and slashes indicate segment boundaries.

2a. *NP-coordination*

Piet kuste Marie en / haar zusje / ook. (1222 ms)

'Piet kissed Marie and / her sister / too.'

2b. *S-coordination*

Piet kuste Marie en / haar zusje / lachte. (1596 ms)

'Piet kissed Marie and / her sister / laughed.'

The significantly longer reading times for the final frame in the S-coordinated sentences (2b) suggested that the ambiguous NP 'her sister' was initially interpreted as part of the direct object of the verb 'kissed', causing substantial processing difficulty when this ambiguous NP turned out to be the subject of a conjoined sentence, as signaled by the disambiguating verb 'laughed'. In other words, when faced with the NP/S-coordination ambiguity, readers prefer NP-coordination over S-coordination.

This apparent preference for NP-coordination was taken by Frazier as evidence for a sentence processing mechanism that is guided by principles of syntactic simplicity. In the framework of garden-path/construal theory, the syntactic description of NP-coordination is the simplest because it requires fewer nodes than S-coordination, and is therefore chosen by application of the minimal attachment strategy (Frazier, 1987b; Frazier & Clifton, 1996; 1997). However, this study, and also a follow-up mentioned in Frazier and Clifton (1997), suffered from a number of shortcomings. As these may have had an impact on the estimated processing difficulty and hence on the estimated strength of the NP-coordination preference, it is important to discuss them in some detail.

The most serious problem of Frazier's study is the fact that processing difficulty is estimated from the reading times on the final frames of the critical sentences. Reading the final frames of S-coordinations could very well take more time because it takes the reader longer to interpret the set of two events depicted in an S-coordinated sentence, instead of only one event as in NP-coordination (cf. Caplan & Waters, 1999). A second and related issue is that whenever two sentences are coordinated, readers may expect a temporal or causal relation to exist between the two conjoined clauses (Gibbs & Moise, 1997; Hendriks, 2004; Kehler, 2002; Mithun, 1988). Inferring this relation presumably takes extra time, especially if it is difficult to find out what kind of relation is intended, as in some of the sentence-coordinations used in Frazier’s experiment. Consider, for example, sentence 3.

3. Inge serveerde de erwtensoep en / de Quiche Lorraine / mislukte.
   ‘Inge served the pea-soup and / the Quiche Lorraine / went-wrong.’

Here, the reader may be looking in vain for a causal relation between the serving of soup and the going awry of a main dish. So it may very well be that some part of the processing difficulty observed in the final frame of sentences such as 3 should be attributed to the difficulty of interpreting the two conjoined sentences in a meaningful way. Finally, it is unclear whether the critical regions that were compared were matched on length and lexical frequency. In sum, the results from the studies mentioned above (Frazier, 1987a; Frazier & Clifton, 1997) cannot provide conclusive evidence for the existence of a conjoint NP preference in sentences containing an NP/S-coordination ambiguity.

A recent study by Hoeks, Vonk, and Schriefers (2002), however, addressed the shortcomings of these earlier experiments by using sentences such as (4). Here, the disambiguating verb *risked* is separated from the final word by at least three words, making it possible to disentangle processes of disambiguation from those involving sentence-final integration. To estimate processing difficulty due to the temporary ambiguity, unambiguous sentences (disambiguated by using a comma) such as (5) were used as controls.

4. The thief shot the jeweler and the cop *risked* his life during the ensuing fight.
5. The thief shot the jeweler, and the cop *risked* his life during the ensuing fight.

The only difference between (4) and (5) is the comma attached to the object noun *jeweler*, which makes it impossible for the reader to conjoin *the cop* with *the jeweler*. And because (4) and (5) are identical in terms of lexical items, sentence-level meaning, and syntactic structure, differences in processing difficulty can only be attributed to the temporary NP/S-coordination ambiguity and not to differences in interpretive complexity. Hoeks et al. found evidence for the predicted NP-coordination preference in two on-line reading experiments (i.e., self-paced reading and eye-tracking). Sentences such as (4), embedded in small story-like texts, were read significantly more slowly than the unambiguous controls at the disambiguating verb *risked* and/or at the post-disambiguation region (e.g., *his life*). Readers evidently assumed that the ambiguous NP *the cop* was part of the direct object and thus incurred processing difficulty when this NP turned out to be the subject of a conjoined sentence.

Though this result is fully compatible with the prediction of garden-path theory, Hoeks et al. argued that readers prefer NP-coordination not for reasons of syntactic simplicity, but because NP-coordination is simpler than S-coordination in terms of *topic-structure*. Topic-structure can be loosely defined as describing the relation between the *topic* of a sentence, that is, the element referring to an entity about which information is given, and the *information* that is expressed by a sentence (see Lambrecht, 1994, for a critical discussion of the notions sentence topic, discourse topic, and topic/focus-structure). In NP-coordinations there is only one topic, which is supposed to be the default and most frequently occurring situation, whereas S-coordinations contain an additional topic *the cop*. Having more than one topic, Hoeks et al. argue, is highly unexpected and will lead to processing difficulty, as readers will have to accommodate the second, unintroduced, entity as a topic in their mental model of the discourse (e.g., Crain & Steedman, 1985; Lambrecht, 1994).

This line of thinking was supported by the results of a crucial second condition in the Hoeks et al. experiments, where context sentences introducing two (simultaneous) topics, such as (6), were shown to effectively eliminate the processing difficulty associated with S-coordination.

6. When they saw the jeweler pulling a gun, the thief and the cop jumped up immediately.

By using a presentational device called a *cataphor-construction*, in which a pronoun (e.g., *they*) precedes the occurrence of the actual referents (e.g., *the thief* and *the cop*), both of these referents are placed into the center of attention, which has the effect of making them very likely topics in the next sentence (cf., Lambrecht, 1994). And indeed, readers did not show any processing difficulty in a subsequent S-coordinated sentence where these same entities served as grammatical subjects of the conjoined sentences.

Though Hoeks et al. argued for a pragmatic origin of the NP/S-coordination ambiguity, other accounts cannot be ruled out. For instance, the apparent absence of processing difficulty in these experiments can still be explained within garden-path theory as the result of very rapid and relatively cost-free reanalysis after an initial misparse. The results from the topic-structure manipulation are also compatible with the other class of processing models: the constraint-based models. These models view sentence interpretation as a constraint-satisfaction process in which a multitude of different factors, including discourse/pragmatic information, can provide different degrees of support for one or the other alternative structure (MacDonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus & Trueswell, 1995). Under most instances of this model, syntactic alternatives enter into a competition process that is assumed to be lengthy and troublesome if the alternatives receive equal support, but which can also be very short if one of the alternatives is strongly biased. Typically, lexically based factors such as the frequency with which a word is used in a specific structural constellation in a given language, but also higher level information from the discourse context may immediately and simultaneously affect the comprehension process. In the NP/S-coordination ambiguity at hand, the presence or absence of topic-structure information from the preceding context must be a factor of importance, but perhaps also the frequency with which the connective *and* is used to either coordinate NPs or Ss. To investigate whether frequency biases involving the connective in the (Dutch) language can be responsible for the NP-coordination preference, we undertook a corpus study (presented below) of a thousand occurrences of 'en', the Dutch equivalent of *and*.

In the Hoeks et al. (2002) experiments, topic-structure was identified as a factor that plays an important role in ambiguity resolution. However, this 'high-level' pragmatic information is not the only source of information that can be used by the parser to resolve the coordination ambiguity (and neither is lexical frequency, for that matter). Indeed, as the next example sentence illustrates, it is crucial for the parser to take more 'local' lexico-semantic factors into account too. Consider, for instance, sentence (7).

7. Jasper sands the board and the carpenter scrapes the paint from the doors.

Here, the same NP/S-coordination ambiguity is present as in example sentence (4), but in this case there is a conflict between the animacy feature of the ambiguous NP *the carpenter* and the thematic requirements of the verb *sands*. At some point, then, the parser must reject *the carpenter* as part of the complex object and analyze it instead as the subject of a conjoined sentence. The use of thematic role information in parsing has been studied extensively (e.g., Clifton, Traxler, Mohammed et al., 2003; Just & Carpenter, 1992; Ferreira & Clifton, 1986; McRae, Feretti, & Amyote, 1997; McRae, Spivey-Knowlton, & Tanenhaus, 1998; McElree & Griffith, 1995; Stowe, 1989; Tanenhaus, Carlson, & Trueswell, 1989; Trueswell, Tanenhaus, & Garnsey, 1994; see also Pickering & Traxler, 1998). One of the best known sentences in this context is undoubtedly (8), taken from Ferreira and Clifton (1986).

8. The evidence examined by the lawyer turned out to be unreliable.

In this sentence, the verb *examined* is used as a past participle introducing a reduced relative clause, but it could also be a tensed main verb, which is the generally preferred reading. The first NP *the evidence*, however, is inanimate and thus a poor AGENT of *examined*, which could lead to some kind of processing difficulty if a main verb reading is preferred. Indeed, Ferreira and Clifton found increased reading times for *examined* in sentences such as (8), as compared to unambiguous controls (e.g., The evidence that was examined by the lawyer ..), indicating that readers were aware of this fact. However, this did not lead readers to abolish the main clause reading; they showed as much processing difficulty when reading the disambiguating by-phrase *by the lawyer* as when the first NP was an animate entity that easily could fulfil the AGENT role (e.g., *the defendant*). Trueswell et al. (1994) challenged this finding by pointing out that some of the 'poor' AGENTS used by Ferreira and Clifton were not that poor at all. For instance, *the car* in "The car towed .." can very well play the role of INSTRUMENT in a towing event. With improved materials Trueswell et al. showed that little or no trace of processing difficulty remained in sentences headed by inanimate NPs that were really poor AGENTs.

These results were recently called into question by Clifton et al. (2003), using a subset of the Trueswell et al. materials. Processing difficulty for temporarily ambiguous sentences headed by an inanimate NP was shown to be (numerically) smaller than when the first NP is animate, but still significantly larger than for unambiguous controls. Clifton et al. stress that these results are not that different from those found by Trueswell et al. (1994), at least as far as first-pass reading times are concerned, but that their extended analyses of the data contradict the claim that animacy can completely eliminate processing difficulty. For instance, Clifton et al. found the largest effects, especially for inanimate first NPs, in a measure called *regression-path duration*, which reflects the time spent on fixating a given region for the first time *plus* all the time spent making regressive fixations to earlier parts of the sentence (e.g., Brysbaert & Mitchell, 1996; Hoeks et al., 2002; Konieczny, Hemforth, Scheepers, & Strube, 1997; Liversedge, Paterson, & Underwood, 1997; Murray, 2000; Rayner & Duffy, 1986; Traxler, Pickering, & Clifton, 1998). This measure was not in general use at the time Trueswell et al. published their paper (i.e., in 1994).

Summarizing, there is still controversy over when and how thematic information is used by the human sentence processing mechanism. There are basically three positions. First of all, the results of Ferreira and Clifton (1986) suggest that thematic information does not help at all in overcoming a garden-path effect. A weaker position is taken by Clifton et al. (2003) who suggest that thematic information may cause a (slight) reduction of the garden-path effect, though it does not completely eliminate processing difficulty. On the other hand, Trueswell et al. (1994) and others claim that thematic information should be able to completely neutralize a garden-path effect, provided that the thematic information is sufficiently strong. The main aim of the present experiment now is to find out which of these three claims best describes the effect of thematic fit in processing temporarily ambiguous S-coordinations. We will use sentences such as (7), in which the ambiguous NP (e.g., *carpenter*) provides the parser with all information necessary to reject NP-coordination and, instead, to adopt S-coordination. The eye-tracking study reported here will investigate at what point in time thematic information is used in the resolution of the NP- versus S-coordination ambiguity.

First, though, we will first present the results of the corpus study to see whether the preference for NP-coordination is supported by a frequency bias (i.e., involving the connective *en*) that could be present in the Dutch language. This is an important issue, because constraint-based models (e.g., such as the one proposed in Trueswell et al., 1994) make predictions about the strength of structural preferences, and about the processing difficulty that results when these preferences are violated, on the basis of the strength of the factors supporting either of the two candidate structures, in this case NP-coordination and S-coordination. We have already seen that one of these factors is the pragmatic bias towards sentences with one and only one topic; the strength of the frequency factor will be determined through a corpus study.

Corpus Study

In a corpus of 1000 occurrences of the connective *en*, taken from the Dutch daily newspaper TROUW, NP-coordination is by far the most frequent structure: it is present in 46 % of all cases. VP-coordination, as in, for instance, *The thief shot the jeweler and made off with the loot*, is the next most frequent structure, with 15 %. Almost as frequent are adverb / adjective coordinations, as in *nice and smooth*, with 14 %. S-coordinations make up 10 % of the total number of occurrences with *en*, and finally coordination of prepositional phrases, as in *for better and for worse* occur in 3 % of all cases. All other cases were assigned to the 'rest' category.

In addition to these so-called 'coarse-grained' frequency measures, we were also interested in the frequencies at a finer grain, that is, we also wanted to find out how often structures that were structurally and semantically similar to our experimental materials would appear in the corpus. To that end we started by making a subdivision of all NP-coordinations in terms of syntactic function, animacy and definiteness. See Table 1 for this frequency information regarding NP-coordinations.

(Table 1 about here)

Firstly, we wanted to see how frequently NP-coordinations occur as grammatical objects. Of the 461 NP-coordinations in the present corpus, most conjoined NPs, 262, served as arguments of prepositions, such as, e.g., *of mice and men* (i.e., 57 % of all NP-coordinations), 138 as grammatical subjects of a sentence (i.e., 30 % of all NP-coordinations), and only 61 of the NP-coordinations (i.e., 13 % of all NP-coordinations) occurred as grammatical objects, which is 6 %

of all instances of *en*. The cases where the conjoint NPs are grammatical objects, were further investigated with respect to animacy and definiteness.

As to animacy, 51 of the 61 'object' NP-coordinations consist of conjoint inanimate entities (i.e., 84 % of all object NP-coordinations). In 8 cases, the conjoint NPs consist of animate entities (i.e., 13 % of all object NP-coordinations), and 2 cases are of mixed animacy (1 animate-inanimate and 1 inanimate-animate). A second important factor in our materials is *definiteness*: In all sentences both the object NP and the ambiguous NP are definite (see, e.g., (4)). Conjoint NPs occurring as grammatical objects are frequently indefinite, namely 39 out of 61 (i.e., 64 % of all object NP-coordinations), against 17 cases where both NPs are definite (i.e., 28 % of all object NP-coordinations), and 5 where definites and indefinites are mixed (i.e., 8 % of all NP-coordinations). And finally, if we count the cases where conjoint NPs functioning as grammatical objects are both definite and animate this amounts to 5, which is 8 % of all object NP-coordinations, and less than 1 % of all cases containing *en* in the present corpus.

We also took a closer look at the S-coordinations that were present in the corpus. For one thing, we determined whether they had grammatical subjects referring to the same or to different entities. For instance, *The thief shot the jeweler and he made off with the loot* is formally an S-coordinated structure, but the subjects of the two conjoined sentences, *the thief* and *he*, refer to the same person (at least in the most plausible interpretation), and so these S-coordinations can be said to have only one topic. In that sense they are different from sentences such as (4), where the subjects refer to two different entities, and which hence have two topics. In 11 of the 96 S-coordinations (i.e., 11 % of all S-coordinations), the subject of the conjoined clause was a pronoun referring to the subject of the first clause (i.e., one-topic S-coordinations), leaving 85 instances (i.e., 9 % of all S-coordinated cases) to have two distinct subjects.

There were also other instances where the subject of the second clause consisted of a pronoun: in 19 cases (i.e., 20 % of all S-coordinations) the second clause was headed by a pronoun referring to the first clause as a whole (as in "Everyone could draw his own conclusions and *that* is exactly what happened"), and in 26 cases, the subjects of the second clause were various kinds of pronouns, none of which referred to the subject of the first clause (i.e,. 27 % of all S-coordinations). In the remaining 39 S-coordinations the subject of the second sentence was an NP (i.e., 41 % of all S-coordinations and 4 % of all cases with *en*). Of these sentences there were 14 with animate subjects, 26 with definite subjects, and only 10 with grammatical subjects that were both animate and definite, which amounts to 1 % of the total number of cases with *en*.

The reason for conducting this corpus analysis was to determine the strength of the frequency factor as it is crucial for constraint-based models in making predictions about processing preferences and associated processing difficulty. On the basis of our present results we may conclude that, while coarse-grained measures may establish a substantial bias for NP-coordination as compared to S-coordination (i.e., 46 % vs. 10 % of all occurrences of *en*), the difference is very much smaller (and reversed!) when more fine-grained frequencies are taken into account: Coordinated NPs as grammatical object occur 6 % of the time, whereas S-coordinations with two different subjects make up for at most 9 % of all cases. If we constrain the options further by stipulating that the grammatical object-NPs must be definite and animate, and the S-coordinations must have a definite, animate NP as subject of the second clause, the percentage of relevant NP- and S-coordinations *do not differ at all* (both cases about 1 % of the total number of occurrences). Thus, while the coarse-grained frequency count would certainly add to a strong NP-coordination preference, going in the same direction as the pragmatic 'one topic' bias, the more fine-grained count is expected to be ineffective in the NP- versus S-coordination ambiguity, as the specific NP- and S-coordinations that were used in the current experiment occurred equally often. See, for instance, Gibson, and Schütze

(1999), Pickering, Traxler, and Crocker (2000), and Rayner and Clifton (2002) for a critical discussion of the use of corpus-based statistics in sentence comprehension research (but see also Desmet, Brysbaert, & De Baecke, 2002; Mitchell, Cuetos, Corley, & Brysbaert, 1995).

Eye-Tracking Study

In the present study we used eye tracking, a technique that allows the observation of normal, uninterrupted reading, and that provides a time-sensitive measure of processing (e.g., Rayner, 1998). This is important because we want to find the earliest point in the sentence where thematic fit has an influence on processing the NP/S-coordination ambiguity.

Method

Participants. The participants were 26 undergraduate students from the University of Nijmegen, who were paid for participation. All had normal, uncorrected vision.

Materials. Two sets of 60 experimental sentences each were constructed, all of which were S-coordinations. This number of items was required because of a planned replication using exactly the same materials in an ERP-experiment (Hagoort, Brown, Vonk, & Hoeks, in preparation), where it is necessary to use relatively large numbers of items to obtain an acceptable signal-to-noise ratio in the EEG-signal. In the first set of 60 sentences, the ambiguous NP (e.g., *the cop*) is animate and fits well as PATIENT of the first main verb. These sentences comprise the condition of good thematic fit (hereafter: good fit). In the second set of 60 S-coordinations, all sentences contained a matrix verb with a clear *selectional restriction* against animate objects, as in sentence (7). Here, the first object NP (*the board*) is inanimate, and a perfectly plausible THEME of the matrix verb *sands*. The ambiguous NP (*the carpenter*), however, is not an acceptable THEME of *sands* because of its animacy. This set of sentences constitutes the poor thematic fit condition (hereafter: poor fit). Unambiguous control sentences were created by attaching a comma to the first object noun of the ambiguous sentence (see Hoeks et al., 2002). Example sentences are shown in Table 2.

(Table 2 about here)

In constructing both sets of sentences (i.e. good fit and poor fit), care was taken that the actions depicted in the first and the second clause of the conjoined sentence were semantically compatible, in order to avoid complications with sentence interpretation. If possible, a disambiguating verb was chosen that was synonymous or closely semantically related, at least intuitively, to the first main verb. This 'plausibility' constraint, together with the fact that a great number of items was needed in each thematic fit condition for the replication with ERP-measurement, made it virtually impossible to construct good and poor fit sentences in a within-item design; instead, a between-item design was used, with different sentences for the good fit and the poor fit condition.

It may be important to note that a between-item design does not prevent us from drawing valid conclusions from the *interaction* that is predicted to occur between thematic fit and ambiguity. Such an interaction will show whether the effect of ambiguity (i.e., ambiguous vs. control) is the same or different for the two thematic fit conditions. To assess the ambiguity effect in each of the thematic fit conditions, reading times at the disambiguating verb in the ambiguous sentence will be compared to the reading times at the disambiguating verb in the control sentence, which is exactly the same word. To put it differently, since each word serves as its own control in the assessment of the interaction, it is not too much of a problem that thematic fit conditions are constructed in a between-item design. There may be a problem if the disambiguating verb would be very different in length (or other relevant characteristics) between the two thematic fit conditions. This might make it more difficult to find an effect in one condition as compared to the other. However, if we look at the time people take to read the critical verb in the *unambiguous* versions of good and poor fit sentences, we can see that these were almost identical (e.g., first-pass reading times: poor fit: 283 ms; good fit: 281 ms, see Table 3). Thus, the conditions seem largely comparable. Another problem of the between-items design is that it makes it difficult to interpret any main effect of thematic fit, but as we are mainly concerned with the effect of ambiguity and the interactions between ambiguity and thematic fit, this shortcoming does not seem to be crucial either.

Eighty filler sentences with conjoined object NPs were added to the experimental S-coordinations, so as to minimize the chance of participants developing processing strategies. In half of these fillers both object nouns were animate, as in (9). Importantly, in the other half the first object noun was *inanimate* and the second one animate, mimicking the order of inanimate-animate nouns in the poor fit condition, as in (10), making it impossible for the reader to accurately predict the syntactic structure of the upcoming sentence from the mere inanimacy of the first postverbal NP.

9. The sultan expelled the rebel and his helper to a deserted island.
10. The company sent a computer and a programmer to solve the problem.

Added to these fillers were 120 sentences from an unrelated experiment containing relative clauses. An example of an unrelated filler is given in (11).

11. De studenten, die de professor gegroet hebben, gaan morgen op vakantie.
    (lit. The students, who the professor greeted have, go tomorrow on holiday.)
    The students who have greeted the professor will go on holiday tomorrow.

Design. To avoid mental fatigue and loss of concentration on the part of the participants, the experiment was run in two separate sessions, each consisting of 15 practice items followed by two blocks of 83 sentences. As we mentioned above, thematic fit (good fit vs. poor fit) was a between-items factor. Each experimental item appeared in two versions: ambiguous (without comma) and control (with comma). Two experimental lists were created using a Latin Square, with equal numbers of items occurring in each condition on each list, and no list containing more than one version of a given item. The order in which experimental and filler items appeared was determined semi-randomly (i.e., allowing maximally three experimental items in consecutive order, but never two consecutive items in the same condition) and was the same for both lists. Each list was presented to an equal number of participants and each participant only saw one list.

Apparatus and Stimulus Specifications. Stimuli were presented on a NEC MultiSync 5FG computer monitor. The maximum number of characters on one line of the screen was 80. Characters appeared in Courier New, size 12. If sentences spanned more than one line, the different lines were separated by a blank line. The sentence then disappeared and the asterisk was shown again. Most sentences (i.e., 73 out of 120) spanned two lines. The disambiguating verb was followed by at least three words on the first line of each sentence. Viewing distance was 85 centimeters, making 1 degree of visual angle equivalent to 4.4 character positions. Both X and Y positions were collected with a sample frequency of 200 Hz and a spatial resolution of .25 degrees using an Amtech ET III infrared pupil reflectance eye tracker (cf. Katz, Mueller, & Helmle, 1987). Only the movements of the right eye were recorded. Head movements were minimized by the use of a bite-bar, combined with a chin and forehead rest.

Procedure. Participants were tested in two separate sessions of approximately 1.5 hours each. Time between sessions ranged from 1 to 10 days. At the start of the first session it was verified that participants had normal vision, and a bite-bar was prepared for each individual participant. Participants were instructed to read the sentences carefully and with normal speed. No comprehension questions had to be answered, as the results of a replication using self-paced reading had shown that adding questions did not affect the general pattern of reading times (see also Hoeks et al., 2002).

One experimental session consisted of two blocks of 83 sentences each. Every sentence was preceded by a screen with an asterisk, indicating the exact location of the beginning of the first word of the following stimulus sentence. Participants were instructed not to blink when reading the sentences but only at the asterisk. When the right-hand button was pushed the asterisk was replaced by a stimulus sentence. Participants were asked to push this same button immediately after they had finished reading the sentence.

Results

The eye-movement data were screened for blinks, track loss, and artefacts caused by the eye tracking apparatus. For analysis purposes, all target sentences were divided into regions of one or more words, as in (12); only the italicized regions were analyzed[1].

12. The thief shot | *the jeweler* | and | *the cop* | *risked* | *his life* | during | *the ensuing fight.*

For every region five dependent measures were calculated (see Rayner, 1998, for a review of measures and nomenclature), which will be defined below: first-pass reading time, forward reading time, first-pass regressions, total time, and regression-path duration (RPD). *First-pass reading time* is defined as the time spent in a region before leaving that region to the right or to the left, provided that the reader enters that region for the first time, and that the region was not skipped on an earlier pass through the sentence. The *forward reading time*[2] is essentially the same as first-pass reading time, but it includes only cases where the reader did not make regressions starting from any point within a prespecified 'critical region' of the sentence (except for within-word regressions); under the current analysis, this regions starts at the determiner of the object NP (i.e., *the jeweler* / *the board*) up to and including the post-disambiguation region. *First-pass regressions* are regressive eye movements originating from a particular region when visiting this region for the first time, provided that that region was not skipped on an earlier pass through the sentence. Regression percentages given are based on the number of times a region was actually fixated in first-pass reading. *RPD* or *regression-path duration* is the time spent in a region in first pass before leaving that region to the right, plus all the time spent in regressing to earlier parts of the sentence. Finally, *total time* is the total time spent in a specific region, so including rereading. In the computation of the measures mentioned above, the duration of the saccades between the fixations that contributed to those measures was included. In other words, ‘time spent’ was taken as a variable, instead of ‘sum of fixation durations’, since it is rather implausible that lexical and supra-lexical processing stops during saccades (cf., Cozijn, 2000; Irwin, 1998; Vonk & Cozijn, 2003; see also Rayner, 1998).

For each region a number of analyses was performed. ANOVAs were conducted on the participant means (F1-analysis) and the item means (F2-analysis) for each of the five eye-movement measures. These analyses involved the factor ambiguity (i.e., ambiguous vs. control), which was treated as within-participants and within-items, and the factor thematic fit (i.e., poor fit vs. good fit), which was treated as within-participants, but between-items (see also section *Materials*). The means of all measures are shown in Table 3; F-values and significance levels can be found in Table 4.

(Tables 3 and 4 about here)

Ambiguous NP. There was no significant interaction, nor were there significant main effects at the ambiguous NP *the cop* / *the carpenter* in either of the five dependent measures.

Disambiguating verb. Total times showed a significant main effect of ambiguity at the disambiguating verb *risked* / *scrapes*: ambiguous sentences took longer to read than the unambiguous controls. No other main effects or interactions were found for total times or any other measure.

Post-disambiguation region. At the post-disambiguation region *his life* / *the paint*, a significant interaction between ambiguity and thematic fit was found in forward reading times, RPDs and total times. In the analysis of the forward reading times the interaction was significant by participants, but not by items. Post-hoc comparisons revealed a significant effect of ambiguity (ambiguous slower than controls) of 19 ms in the good fit condition ($F_1(1,25) = 5.54$, $p < .05$; $F_2(1,59) = 4.99$, $p < .05$). In the poor fit condition there was a non-significant 9 ms difference in the opposite direction (*p*-values > .20); no other effects were significant for forward reading times. A similar pattern was present in the post-hoc comparisons involving RPDs and total times: RPDs were 51 ms longer for ambiguous sentences than for controls in the good fit condition ($F_1(1,25) = 18.16$, $p < .001$; $F_2(1,59) = 14.63$, $p < .001$); in the poor fit condition this difference was 5 ms, which was not significant (*F*s < 1). Total times were 33 ms longer for ambiguous as compared to control sentences in the good fit condition ($F_1(1,25) = 16.96$, $p < .001$; $F_2(1,59) = 15.17$, $p < .001$); in the poor fit condition this difference amounted to 11 ms (both *p*-values > .10).

Analyses at the post-disambiguation region also revealed main effects of ambiguity (i.e., ambiguous > control) in all measures except forward reading times and first-pass regressions. Main effects of thematic fit (i.e., good fit > poor fit) were present in first-pass reading times (i.e., in the analysis by participants), regressions (i.e., in the analysis by participants), RPDs, and total times. Please note that in forward reading times, RPDs, and total times main effects were qualified by a significant interaction.

Other regions

1. Object NP of First Clause. No significant interactions of ambiguity and thematic fit were found at the object NP *the jeweler(,) / the board(,)*. The main effect of ambiguity was significant in forward reading times (i.e,. in the analysis by items), regressions, RPDs, and total times (i.e., in the analysis by participants); it did not reach significance in first-pass reading times. This effect of ambiguity reflected the ambiguous condition (without the comma) being easier in terms of shorter reading times and fewer regressions, than the control condition (with the comma). Main effects of thematic fit (i.e., good fit > poor fit) were found in first-pass reading times, forward reading times and total times, though in all of these measures only in the analysis by participants.

2. Sentence-final region (i.e., final three words of a sentence). Forward reading times were not computed because of the high incidence of regressions at the end of the sentence, leaving too few observations on which to base an average. Computation of RPDs did not seem appropriate either, since it was impossible to determine whether readers were making regressions in order to re-read the sentence or whether they were just making a saccade to the screen-position where the asterisk for the next sentence would appear. Only first-pass reading times and regression percentages were computed. No significant effects were found.

Discussion

The aim of the present study was to investigate how thematic information is used in resolving the NP- versus S-coordination ambiguity. The interaction between thematic fit and ambiguity that we predicted was found at the post-disambiguation region of temporarily ambiguous S-coordinations, in forward reading times, RPDs and total times. This interaction reflected the presence of processing difficulty in good fit sentences and the absence thereof in the poor fit sentences. Thus, the thematic misfit information was used rapidly to minimize the processing difficulty caused by the NP-coordination preference. It was somewhat surprising to find the critical interaction between thematic fit and ambiguity not at the disambiguating verb (or even earlier) but in the post-disambiguation region, especially given the widely held belief that it is rather uncommon for eye-tracking studies to find effects one or two words downstream from the critical word (but see, e.g., Van Gompel, Pickering, & Traxler, 2001). Thus, the present results suggest that the command to move the eyes from the disambiguating verb to a subsequent region is issued *before* it has become clear that there is a processing problem, or before the processor has decided what to do about it and adjusts its motor program accordingly (see Just, Carpenter, & Woolley, 1982, for a similar argument). [3]

But though thematic information was used rapidly, and processing difficulty was reduced greatly, we also found significant evidence for residual processing difficulty. Most importantly, at the disambiguating verb, total times showed a significant main effect of ambiguity - in the absence of an interaction - indicating that readers incurred processing difficulty in the good fit, but also in the poor fit condition. This same pattern was found for first-pass reading times at the post-disambiguation region, where there was a significant main effect of ambiguity, but no interaction with thematic fit. Thus, the pattern of both total times and first-pass reading times does not support the claim that processing difficulty is completely eliminated in the poor thematic fit sentences. [4]

This state of affairs is most consistent with the results of Clifton et al. (2003), who found that thematic information may reduce, but not completely eliminate garden-path effects. Garden-path theory, the framework adopted by Clifton et al., actually predicts a reanalysis effect at the *ambiguous NP* of poor fit sentences (but not in good fit sentences), at least for versions that include a 'thematic processor', an independent module that examines the plausibility of each decision made by the syntactic processor, and that may propose syntactic alternatives that are more plausible than the structure that is initially chosen (e.g., Rayner, Carlson, & Frazier, 1983). Thus, the preference to coordinate NPs leads to the violation of the main verb's selectional restrictions, and hence, the thematic processor will select S-coordination because it is semantically more plausible. However, there is no sign of a reanalysis effect at the ambiguous NP. Nevertheless, it may be possible that the effect is delayed, and appears one word later at the disambiguating verb. Our results fit somewhat less well with predictions made on the basis of Ferreira and Clifton (1986), who put forward the strong view that thematic information does not affect ambiguity resolution at all. Though this seems to be the case when we look exclusively at first-pass reading times, there is undoubtedly a strong reduction of forward reading times, total times, and RPDs at the post-disambiguation region of poor fit sentences.

The present results seem least compatible with the predictions made by Trueswell et al. (1994), who claimed that thematic information should be able to eliminate garden-path effects without leaving a trace. As we have seen, the factor thematic fit does not, or at least not immediately, outweigh the factors that in the constraint-based framework may be held responsible for the preference for NP-coordination. These are the minimal topic-structure principle (i.e., assume only one topic), and, if the parser is sensitive to coarse-grained frequencies, also the frequency of usage of the connective (biasing strongly towards NP-coordination). It is possible that minimal topic-structure and frequency *together* are simply too strong to be instantly overcome by the manipulation of thematic fit. This could be taken to suggest that the parser uses the strongly biasing coarse-grained frequency data, in which case two factors favoring NP-coordination might team up against the very strong thematic fit manipulation. Alternatively, if the parser were to make use of fine-grained frequencies (thus ruling out frequency as a factor of importance, see corpus study), simple topic-structure must be at least as strong as thematic fit to cause this pattern of data. On a more speculative note, the reason for the other factors temporarily resisting the effect of thematic fit might also lie in the fact that simple topic-structure and frequency information are present *before* the ambiguous constituent is encountered, whereas the thematic fit between ambiguous NP and preceding main verb can only be evaluated *after* the ambiguous noun is actually read. Under this account, the thematic fit information arrives relatively late in the ambiguity resolution process which might then give rise to a delayed use of the thematic fit information. We did not find any significant effect at the ambiguous NP itself, so possible processes of competition must be assumed to be delayed by one word, to become manifest only at the disambiguating verb of poor fit sentences.

An important point that has not been discussed yet, is the fact that in poor fit sentences NP-coordination may be unlikely for reasons other than thematic fit, namely because of the infelicity of having a coordination of an inanimate and an animate NP, with the inanimate appearing first. Indeed, if we take a look at our corpus study, we find that only one conjoint object NP has this inanimate-animate order and that, in general, mixed animacy is very uncommon. In addition, McDonald, Bock and Kelly (1993) found that when language users have to give acceptability judgments to sentences containing conjoint object NPs of mixed animacy, they generally prefer to have the animate conjunct first. Thus, in the poor fit condition there may be two factors at work arguing against NP-coordination: the selection restriction of the first main verb and the inappropriate order of an inanimate preceding an animate. Interestingly, the parallel ERP-study (Hagoort et al., in prep) included sentences where the first main verb does not select against animate objects, as in, for example, "Jasper *saw* the board and the carpenter scraped ..". Preliminary results revealed a P600 / SPS (i.e., Syntactic Positive Shift, see Hagoort, Brown, & Groothusen, 1993) time-locked to the noun of the ambiguous NP, which reflects the effortful syntactic processing that can be brought about by syntactic, but also by semantic anomalies (see, e.g., Hoeks, Stowe, & Doedens, 2004). This shows that the processor quickly reacts to the imbalance in animacy. In contrast, there were no significant differences between ambiguous and control sentences in the *poor fit* condition (i.e., where the first main verb *did* carry selection restrictions) that was also part of the parallel ERP experiment (see also Footnote 4). Thus, these results suggest that there is a distinct difference between processing an NP that is inappropriate as the object of a verb, and the processing of this same NP when it violates the preferred order of elements in a conjoint NP. For further discussion of why these conditions behave so differently we refer to Hagoort et al. (in prep).

A final point that we want to make concerns the severity of the garden-path effect that was observed here. Estimated processing difficulty seemed to be rather modest: at the disambiguating verb, the difference between the ambiguous and control condition only amounted to 12 ms; at the post-disambiguation region processing difficulty also did not seem huge (first-pass reading times: 13 ms; regressions: 2 percent; total times: 33 ms; RPDs: 51 ms). The results from other studies support this observation of moderate gardenpathing for the NP- versus S-coordination ambiguity (Hagoort et al., in prep; Hoeks et al., 2002; Kaan & Swaab, 2003; and also in two replications using self-paced reading reported by Hoeks, 1999). Thus, the garden-path effect seems to be rather weak, which is quite unexpected under a number of sentence processing accounts. For instance, garden-path theory predicts large effects because of the costly structural revisions after initial minimal attachment. On the other hand, constraint-based models expect large effects because of the strong constraints that are in favor of NP-coordination (i.e,. simple topic-structure and, in some models, coarse-grained frequency bias). Furthermore, models of reanalysis, such as the one proposed by Sturt and Crocker

(Sturt & Crocker, 1996; 1997; Sturt, Pickering, & Crocker, 1999) also predict a large reanalysis effect in the ambiguity at hand because the thematic link between the ambiguous NP and the preceding verb must be severed when the sentence turns out to be S-coordinated. The 'Attach Anyway' model of Fodor and Inoue (1998), however, does predict relatively low reanalysis costs, because after revision the first main verb is not left with an unfilled argument slot for object, as it only has to give up its second, coordinated argument (i.e., the ambiguous NP), while retaining syntactic and thematic links to the 'real' object NP. As the issue of what actually constitutes a small or a large effect can only be effectively addressed by some kind of within-experiment comparison, we must leave it to future research to determine whether there is really a difference in strength of the garden-path effect between temporarily ambiguous S-coordinations and other syntactic ambiguities.

In sum, we have seen that readers prefer NP-coordination over S-coordination, which causes them to incur processing problems in temporarily ambiguous S-coordinations. Thematic information going against the conjoint NP preference rapidly reduces the garden-path effect, but cannot prevent significant residual processing difficulty. We argued that this pattern of results is most consistent with the predictions of Clifton et al. (2003), to some extent also with Ferreira & Clifton (1986), but least consistent with Trueswell et al. (1994). As we suggested above, garden-path theory may need to assume a one-word delay to account for the fact that the garden-path was not present at the ambiguous NP, where the thematic misfit should have become apparent to the thematic processor. Constraint-based models need to assume that the minimal topic-structure principle, possibly in concordance with frequency, is simply too strong to be immediately overcome by the strong thematic misfit constraint, or that, at least in some instances, thematic information may not be available early enough to stop the processor from going a small step up the garden-path.

References

Altmann, G. T. M. (1994). Regression-contingent analyses of eye movements during sentence processing: Reply to Rayner and Sereno. *Memory and Cognition*, *22*, 286-290.

Altmann, G. T. M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language*, *31*, 685-712.

Altmann, G. T. M., van Nice, K. Y., Garnham, A., & Henstra, J. A. (1994). Late closure in context. *Journal of Memory and Language*, *38*, 459-484.

Brysbaert, M., & Mitchell, D. C. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. *The Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology, 49,* 664-695.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22,* 77-126.

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

Cozijn, R. (2000). *Integration and inference in understanding causal sentences.* Doctoral dissertation, Tilburg University, Tilburg, The Netherlands.

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

Desmet, T., Brysbaert, M., & De Baecke, C. (2002). The correspondence between sentence production and corpus frequencies in modifier attachment. *Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology, 55,* 879-896.

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

Fodor, J. D., & Inoue, A. (1998). Attach Anyway. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 101-141). Dordrecht: Kluwer.

Frazier, L. (1987a). Syntactic processing: evidence from Dutch. *Natural Language and Linguistic Theory*, *5*, 519-559.

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

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

Frazier, L., & Clifton, C., Jr. (1997). Construal: Overview, motivation, and some new evidence. *Journal of Psycholinguistic Research*, *26*, 277-295.

Gibbs, R. W., & Moise, J. F. (1997). Pragmatics in understanding what is said. *Cognition*, *62*, 51-74.

Gibson, E., & Schütze, C. T. (1999). Disambiguation preferences in noun phrase conjunction do not mirror corpus frequency. *Journal of Memory and Language, 40,* 263-279.

Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8,* 439-484.

Hendriks, P. (2004). Coherence Relations, Ellipsis, and Contrastive Topics. *Journal of Semantics, 21,*133-153.

Hoeks, J. C. J. (1999). *The processing of coordination: semantic and pragmatic constraints on ambiguity resolution.* Doctoral Dissertation. University of Nijmegen, Nijmegen, The Netherlands.

Hoeks, J. C. J., Stowe, L. A., & Doedens, L. H. (2004). Seeing words in context: the interaction of lexical and sentence level information during reading. *Cognitive Brain Research, 19,* 59-73.

Hoeks, J. C. J., Vonk, W., & Schriefers, H. (2002). Processing coordinated structures in context: the effect of topic-structure on ambiguity resolution. *Journal of Memory and Language, 46,* 99-119.

Irwin, D.E. (1998). Lexical processing during saccadic eye movements. *Cognitive Psychology*, *36*, 1-27.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 98,* 122–149.

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

Kaan, E., & Swaab, T. Y. (2003). Electrophysiological evidence for serial sentence processing: a comparison between non-preferred and ungrammatical continuations. *Cognitive Brain Research, 17,* 621-635.

Katz, B., Mueller, K., & Helmle, H. (1987). Binocular eye movement recording with CCD arrays. *Neuro-ophtalmology*, *7*, 81-91.

Kehler, A. (2002). *Coherence, Reference, and the Theory of Grammar*. Stanford, CA: CSLI Publications.

Konieczny, L., Hemforth, B., Scheepers, C., & Strube, G. (1997). The role of lexical heads in parsing: Evidence from German. *Language and Cognitive Processes*, *12*, 307-348.

Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents*. Cambridge, MA: University Press.

Liversedge, S. P., Paterson, K. B., & Underwood, G. (1997). *Exploring the effects of quantifiers on parsing*. Poster presented at the Tenth CUNY Conference on Human Sentence Processing, Sta. Monica, CA.

McDonald, J. L., Bock K., & Kelly, M. H. (1993). Word order and world order: semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology, 25,* 188-230.

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

McRae, K., Feretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes, 12,* 137-176.

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

McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 134-157.

Mitchell, D. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (non-lexical) statistical records. *Journal of Psycholinguistic Research, 24,* 469-487.

Mithun, M. (1988). The grammaticization of coordination. In J. Haiman & S. A. Thompson (Eds.), *Clause combining in grammar and discourse* (pp. 331-360). Amsterdam: Benjamins.

Murray, W. (2000). Sentence processing: Issues and measures. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 649-664). Amsterdam: Elsevier.

Ni, W., Crain, S., & Shankweiler, D. (1996). Sidestepping garden paths: Assessing the contributions of syntax, semantics, and plausibility in resolving ambiguities. *Language and Cognitive Processes*, *11*, 283-334.

Pickering, M. J., & Traxler, M. J. (1998). Plausibility and recovery from garden paths: an eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 940-961.

Pickering, M. J., Traxler, M. J., & Crocker, M. W. (2000). Ambiguity resolution in sentence processing: evidence against frequency-based accounts. *Journal of Memory and Language, 43,* 447-475.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *3*, 372-422.

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.

Rayner, K., & Clifton, C., Jr. (2002). Language comprehension. In D. Medin (Vol. Ed.), *Stevens Handbook of experimental psychology: Vol. 2. Memory and cognitive processes* (3rd ed., pp. 261-316). New York: Wiley.

Rayner, K. & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14,* 191-201.

Rayner, K., & Sereno, S. A. (1994a). Regressive eye movements and sentence parsing: On the use of regression-contingent analyses. *Memory and Cognition*, *22*, 281-285.

Rayner, K., & Sereno, S. A. (1994b). Regression-contingent analyses: A reply to Altmann. *Memory and Cognition*, *22*, 291-292.

Stowe, L. A. (1989). Thematic structures in sentence comprehension. In G.N. Carlson & M.K. Tanenhaus (Eds.), *Linguistic structure in language processing.* Dordrecht: Kluwer.

Sturt, P., & Crocker, M. W. (1996). Monotonic syntactic processing: a cross-linguistic study of attachment and reanalysis, *Language and Cognitive Processes, 11,* 449-494.

Sturt, P., & Crocker, M. W. (1997). Thematic monotonicity. *Journal of Psycholinguistic Research, 26,* 297-322.

Sturt, P., Pickering, M. J., & Crocker, M. W. (1999). Structural change and reanalysis difficulty in language comprehension. *Journal of Memory and Language, 40,* 136-150.

Tanenhaus, M. K., Carlson, G. N., & Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes, 4*, 211-234.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. L. Miller & P. D. Eimas (Eds.), *Speech, language and communication* (pp. 217-262). San Diego, CA: Academic Press.

Traxler, M. J., Pickering, M. J., & Clifton, C., Jr. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language, 39,* 558-592.

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.

Van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language, 45,* 225-258.

Vonk, W., & Cozijn, R. (2003). On the treatment of saccades and regressions in eye movement measures of reading time. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: cognitive and applied aspects of eye movement research* (pp. 291-312). Oxford: Elsevier.

Author Note

This study was partly funded by grant # 015.001.103 from the Netherlands Organization for Scientific Research (NWO) awarded to Dr. Petra Hendriks for the project 'Optimization in Coordination'. We would like to thank Marc Brysbaert, Chuck Clifton, Don Mitchell, Martin Pickering, Keith Rayner, Laurie Stowe, and Roger van Gompel for very helpful comments on earlier versions of this paper. We are also grateful to Ingeborg Prinsen and Judith Pijnacker for their help with the corpus study. The complete set of Dutch experimental materials can be obtained electronically via http://www.let.rug.nl/~hoeks/QJEP2005_mat.pdf

Address correspondence and reprint requests to John Hoeks, CLCG / BCN NeuroImaging Center, University of Groningen, PO Box 716, 9700 AS Groningen, The Netherlands. Phone: +31 (0)50 3637443; Fax: +31 (0)50 3636855. Electronic mail may be sent to: j.c.j.hoeks@rug.nl

Footnotes

1. Since the connective *en* ('and') was skipped more than 85 % of the time, we performed additional analyses in which the connective is added to either the preceding region (object NP), or the following region (ambiguous NP). The results of both of these analyses did not differ from the results that will be reported below.

2. This measure can be seen as an extension of Altmann's notion of regression-contingent analysis of eye-movement data (Altmann, 1994; Altmann, Garnham, & Dennis, 1992; Altmann, van Nice, Garnham, & Henstra, 1998; Ni, Crain, & Shankweiler, 1996; Vonk & Cozijn, 2003; see also: Rayner, 1998; Rayner & Sereno, 1994a; 1994b). Altmann showed that uncorrected first-pass reading times can underestimate the processing difficulty in a specific region when there is a considerable number of regressions. The duration of fixations immediately preceding a regression are often relatively short, thereby reducing mean first-pass reading times on regions that are actually problematic for the reader. The forward reading time measures reported here are based on approximately 62 % of all observations. See Vonk and Cozijn (2003) for further discussion of the forward reading time measure.

3. Hoeks et al. (2002) speculated that this kind of reading behavior may depend on individual reading styles, as well as on task-related characteristics. For instance, it may vary between groups of participants whether, on average, processing problems are solved by making regressions, by increasing the duration or the number of fixations, or by moving rapidly towards the end of the sentence in the hope that the problem will solve itself in due time. For instance, having a large proportion of participants taking the third route may lead to a greater probability of finding effects downstream from the critical word. Task-related characteristics may also be involved. One can imagine that reading 'strategies' may be quite different depending on whether

participants are reading sentences in context, which typically requires integration with previous and subsequent pieces of text as compared to when they read sentences in isolation, where no such integration is required.

4. Preliminary analyses of the parallel ERP study reported in Hagoort et al. (in prep), using exactly the same materials as we used here, did not reveal significant evidence for processing difficulty in poor fit sentences (as compared to unambiguous controls) either at the noun of the ambiguous NP or at the disambiguating verb. The reason for this difference between eye tracking and ERP registration is not clear, but it does indicate that the residual processing difficulty is indeed rather small.

Table 1 *Frequency Counts of NP-Coordinations Functioning as Grammatical Objects (n=61) as a Function of Definiteness and Animacy*

	An-An	An-In	In-An	In-In	Total
Def-Def	5	0	0	12	17
Def-Indef	0	0	0	2	2
Indef-Def	0	1	1	1	3
Indef-Indef	3	0	0	36	39
	8	1	1	51	61

Note. Def = definite; Indef=indefinite; An=animate; In=inanimate.

Table 2 *Sample Materials of Eye-Tracking Experiment, With Literal English Translations*

Condition	Example Sentence
Good Fit	
Ambiguous	De dief beschoot de juwelier en de agent riskeerde zijn leven tijdens het daaropvolgende gevecht. 'The thief shot the jeweler and the cop risked his life during the ensuing fight.'
Control	De dief beschoot de juwelier, en de agent riskeerde zijn leven tijdens het daaropvolgende gevecht. 'The thief shot the jeweler, and the cop risked his life during the ensuing fight.'
Poor Fit	
Ambiguous	Jasper schuurt de plank en de timmerman krabt de verf van de deuren. 'Jasper sands the board and the carpenter scrapes the paint from the doors.'
Control	Jasper schuurt de plank, en de timmerman krabt de verf van de deuren. 'Jasper sands the board, and the carpenter scrapes the paint from the doors.'

Table 3 *Means of Five Eye-Tracking Measures as a Function of Region, Thematic Fit and Ambiguity*

		Condition			
		Good Thematic Fit		Poor Thematic Fit	
		ambiguous	control	ambiguous	control
Measure	Region				
First-Pass Reading Time (ms)	Object NP	354	354	327	345
	Ambiguous NP	374	387	384	389
	Disambiguating Verb	285	281	288	283
	Post-Disambiguation Region	340	327	316	310
	Final Region	937	954	912	909
Forward Reading Time (ms)	Object NP	351	356	321	347
	Ambiguous NP	373	377	384	379
	Disambiguating Verb	287	284	292	285
	Post-Disambiguation Region	335	316	310	319
Regressions	Object NP	9	17	11	16
	Ambiguous NP	7	7	6	6
	Disambiguating Verb	6	6	7	5
	Post-disambiguation Region	10	8	6	7
	Final Region	41	37	36	36

Regression-	Object NP	406	446	384	428
Path Duration	Ambiguous NP	413	419	411	418
(ms)	Disambiguating Verb	318	317	322	309
	Post-Disambiguation Region	414	363	347	342
Total Time	Object NP	389	394	362	382
(ms)	Ambiguous NP	413	416	417	413
	Disambiguating Verb	321	309	316	303
	Post-Disambiguation Region	381	348	335	324

Table 4 *Main of Ambiguity and Thematic Fit and Their Interaction for all Eye-Tracking Measures.*

	Main Effects and Interaction					
	Ambiguity		Thematic Fit		Interaction	
	F_1	F_2	F_1	F_2	F_1	F_2
First-Pass Reading Time						
Object NP	2.34	2.60	**8.55**	2.65	3.71	1.94
Ambiguous NP	1.01	2.83	1.05	<1	<1	<1
Disambiguating Verb	1.09	1.26	<1	<1	<1	<1
Post-Disambiguation Region	**4.25**	**6.54**	**11.90**	2.58	<1	<1
Final Region	<1	<1	**4.87**	1.50	<1	<1
Forward Reading Time						
Object NP	2.84	**7.41**	**9.35**	1.59	2.85	<1
Ambiguous NP	<1	<1	<1	<1	<1	<1
Disambiguating Verb	<1	1.27	<1	<1	<1	<1
Post-Disambiguation Region	<1	1.95	1.48	<1	**5.45**	1.70
Regressions						
Object NP	**12.00**	**23.24**	<1	<1	<1	<1
Ambiguous NP	<1	<1	1.57	1.35	<1	<1
Disambiguating Verb	<1	1.57	<1	<1	<1	<1
Post-Disambiguation Region	<1	<1	**7.65**	2.39	2.87	2.32

Final Region	2.06	1.62	2.90	2.02	1.28	1.62
Regression-path Duration						
Object NP	**26.53**	**24.85**	3.42	1.30	<1	<1
Ambiguous NP	<1	<1	<1	<1	<1	<1
Disambiguating Verb	<1	2.35	<1	<1	1.11	<1
Post-Disambiguation Region	**8.52**	**11.79**	**18.90**	**7.71**	**20.19**	**7.02**
Total Time						
Object NP	3.54	**4.42**	**6.06**	1.92	1.24	1.09
Ambiguous NP	<1	<1	<1	<1	<1	<1
Disambiguating Verb	**5.52**	**7.16**	2.46	<1	<1	<1
Post-Disambiguation Region	**10.25**	**16.94**	**26.50**	**7.37**	**5.47**	**4.69**

Note. **Bold** = $p < .05$; Degrees of freedom are (1,25) for the F_1-analyses, and (1,118) for the F_2-analyses.