Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Memory and Language 49 (2003) 353–374
Journal of Memory and Language
www.elsevier.com/locate/jml

Disambiguation preferences and corpus frequencies in noun phrase conjunction

Timothy Desmet[a,*] and Edward Gibson[b]

[a] *Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium*
[b] *Massachusetts Institute of Technology, USA*

Received 22 August 2002; revision received 10 February 2003

Abstract

Gibson and Schütze (1999) showed that on-line disambiguation preferences do not always mirror corpus frequencies. When presented with a syntactic ambiguity involving the conjunction of a noun phrase to three possible attachment sites, participants were faster to read attachments to the first site than attachments to the second one, although the latter were shown to be more frequent in text corpora. In the present study, we investigated whether a particular feature in their items—disambiguation using the pronoun 'one'—could account for this discrepancy. The results of a corpus analysis and two on-line reading experiments showed that the presence of this pronoun is indeed responsible for the high attachment preference in the conjunction ambiguity. We conclude that for this syntactic ambiguity there is no discrepancy between on-line preferences and corpus frequencies. Consequently, there is no need to assume different processes underlying sentence comprehension and sentence production on the basis of the noun phrase conjunction ambiguity.

Keywords: Noun phrase conjunction; Corpus frequency; Syntactic ambiguity resolution; Eye movements

Whereas robust frequency effects have long been discussed in research on lexical access (e.g., Forster & Chambers, 1973; Howes & Solomon, 1951; Preston, 1935; Rubenstein, Garfield, & Millikan, 1970), it is only recently that frequency has started to play a substantial role in sentence processing research (e.g., Jurafsky, 1996; Mitchell, Cuetos, Corley, & Brysbaert, 1995; Rohde, 2002; Seidenberg & MacDonald, 1999; Tabor, Juliano, & Tanenhaus, 1997). For instance, constraint-based accounts hypothesize that one factor that affects people's preferred interpretations of temporarily ambiguous sentences is the relative frequencies of different argument structures for words in the sentence (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Garnsey, 1994; Trueswell, Tanenhaus, & Kello, 1993). For example, in the sentence beginning with "Amanda believed the senator...," the noun phrase (NP) "the senator" is temporarily ambiguous between the direct object of the verb 'believed' (e.g., "Amanda believed the senator during the speech") and the subject of an embedded sentence (e.g., "Amanda believed the senator was lying"). Because the verb 'believed' is most frequently used with an embedded sentence, accounts that state that the sentence comprehension mechanism relies on this kind of frequency-information can explain why people find it easier to read the embedded sentence interpretation than the direct object interpretation following the verb 'believed,' but harder when it follows verbs like 'understood,' which occurs more often with an object NP (see Ford, Bresnan, & Kaplan, 1982; Mitchell & Holmes, 1985, for early investigations of the

* Corresponding author. Fax: +32-09-264-64-96.
E-mail address: Timothy.Desmet@rug.ac.be (T. Desmet).

0749-596X/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0749-596X(03)00025-1

role of argument structure information in sentence parsing).

A theoretical account that relies heavily on the role of frequency information to explain syntactic ambiguity resolution is the tuning hypothesis (Cuetos & Mitchell, 1988; Cuetos, Mitchell, & Corley, 1996; Mitchell, 1994; Mitchell & Brysbaert, 1998; Mitchell & Cuetos, 1991; Mitchell et al., 1995). According to this proposal, the initial analysis of a syntactically ambiguous structure is determined by exposure to similar structures in the past. Supporting this claim, there is evidence that the preferred interpretation of certain temporarily ambiguous structures is the more frequent resolution in text corpora (Corley, 1996; Cuetos et al., 1996; Mitchell & Brysbaert, 1998; Mitchell et al., 1995; cf., Tabor et al., 1997, for a related account).

Exposure-based syntactic ambiguity resolution theories make a straightforward prediction: patterns found in comprehension measures should correlate with patterns found in analyses of text corpora. However, two studies showed that this correlation is not always present (Gibson & Schütze, 1999; Mitchell & Brysbaert, 1998). The study by Mitchell and Brysbaert (1998) concerns the relative clause attachment ambiguity in Dutch examples with the syntactic structure, illustrated in English in (1):

(1) Someone shot [$_{NP1}$ the servant of [$_{NP2}$ the actress]] [$_{RC}$ who was on the balcony].

In this syntactic ambiguity the relative clause (RC) can either be attached to NP1 "the servant" (*high* attachment) or to NP2 "the actress" (*low* attachment). It has been shown that Dutch-speaking readers have a preference to attach the RC high (Brysbaert & Mitchell, 1996; Desmet, De Baecke, & Brysbaert, 2002b; Mitchell, Brysbaert, Grondelaers, & Swanepoel, 2000). In apparent contrast to the comprehension data, Mitchell and Brysbaert (1998) found that low-attaching RCs (469 instances) were more frequent than high-attaching RCs (206 instances) in a Dutch newspaper and magazine article corpus that they analyzed. Mitchell and Brysbaert took this as evidence against the tuning hypothesis. However, in a recent study Desmet, Brysbaert, and De Baecke (2002a) reanalyzed the Mitchell and Brysbaert (1998) corpus and performed an additional corpus analysis in Dutch. They demonstrated the existence of a grain size of analysis at which the corpus data are not in contradiction with the high attachment preference in reading times. The grain size problem relates to the question of what structures to count when performing a corpus analysis (or in terms of the tuning hypothesis: What kinds of structures influence the human sentence processing mechanism?). For instance, with regard to the RC attachment ambiguity in sentence (1), one could count all instances in which a RC can be attached to one of two preceding NPs. However, both finer and coarser grain sizes are possible. One of the finest-grain solutions would be that statistics are collated for the exact words that are used. For instance, of all occasions where a RC follows the words "the servant of the actress," how many times was the attachment resolved high versus low? A coarser-grain solution could be to count all occurrences of two NPs followed by any kind of modifying constituent. Desmet et al. argued that Dutch comprehension studies showing a high-attachment bias mainly used items containing a human NP1. When the corpus searches were also restricted to sentences with a human NP1, high-attaching RCs were more frequent than low-attaching RCs, which refutes the Dutch evidence against the tuning hypothesis presented by Mitchell and Brysbaert (1998).

A second finding is more problematic for frequency-based accounts. In a series of studies, Gibson and Schütze (1996, 1999) and Gibson, Schütze, and Salomon (1996) investigated how people process a syntactic ambiguity involving conjoined noun phrases to three potential attachment sites, as illustrated in (2).

(2) The salesman ignored *a customer* with *a child* with *a dirty face* and...
  (a) a wet diaper. (low conjunction)
  (b) one with a wet diaper. (middle conjunction)
  (c) one with a baby with a wet diaper. (high conjunction)

Both an off-line survey rating the comprehensibility of the different attachments (Gibson et al., 1996) and a reading time study (Gibson & Schütze, 1999) indicated that middle attachments are harder to understand than high attachments. However, in contrast to the prediction of exposure-based accounts of sentence processing, detailed corpus analyses at several grain sizes revealed that middle attachments are significantly more frequent than high attachments (Gibson et al., 1996). As an explanation for this discrepancy between comprehension complexity and corpus frequency, Gibson and Schütze (1999) argued for different processes underlying sentence comprehension and sentence production. They proposed that both comprehension and production are sensitive to locality considerations, but that an additional factor, favoring high attachment, is only involved in sentence comprehension and not in sentence production.

In the present paper we investigated an alternative explanation for the mismatch between the corpus frequencies and the comprehension results. The sentences that were used by Gibson and Schütze (1999) in their online reading study may not have been representative of the syntactic structures that were included in the corpus counts. One characteristic of their items is particularly cumbersome: the use of the pronominal 'one' to disambiguate the conjunction attachment (see 2). It has been shown that when there are several possible antecedents

for a pronoun (or *anaphor*), the first-mentioned antecedent is often the most accessible (Chang, 1980; Corbett & Chang, 1983; Gernsbacher, 1989; Gernsbacher & Hargreaves, 1988; Gordon, Grosz, & Gilliom, 1993). It could be argued that the presence of the 'one'-pronoun also induced a high attachment preference in the reading experiment of Gibson and Schütze (1999).

More directly related to the present ambiguity, Hemforth, Konieczny, and Scheepers (2000) showed that high attachment preferences in German can be changed to low attachment preferences when the pronoun is left out. They investigated sentences as in (3):

(3a) Die Tochter der Lehrerin, die aus Deutschland kam, traf John. [The daughter of the teacher, who came from Germany, met John.]

(3b) Die Tochter der Lehrerin aus Deutschland traf John. [The daughter of the teacher from Germany met John.]

Sentence (3a) is an RC attachment ambiguity and contains the relative pronoun 'who.' Sentence (3b) is a prepositional phrase attachment ambiguity which is very similar in meaning to (3a), but which does not contain an anaphor. Whereas Hemforth et al. observed a high attachment preference with the former, a low attachment preference was observed in the latter. On the basis of these results they proposed the anaphoric resolution hypothesis, according to which the presence of a relative pronoun induces a high attachment preference in ambiguous attachments as in (3) (cf., a more recent proposal by Hemforth & Konieczny, 2002).

The goal of the present studies was to investigate whether the presence of the pronoun 'one' in the Gibson and Schütze (1999) study induced a similar high attachment preference in the conjunction ambiguity and whether this feature could explain the discrepancy between the corpus frequency and the sentence comprehension complexity. First, we present some corpus data to show that there is a different attachment preference in the corpus when the pronoun 'one' is present. Second, we present a self-paced reading experiment repeating the experiment of Gibson and Schütze, both with items containing the pronoun 'one' and with items not containing a pronoun, showing that the presence of the pronoun 'one' also influences on-line reading. Finally, this finding is generalized to Dutch in an eye-tracking experiment.

Corpus analysis

To investigate whether the presence of pronouns makes a difference in the way attachments to three preceding noun phrase attachment sites are made, Gibson and Schütze (1999) analyzed the Brown corpus (Kučera & Francis, 1967) and the Wall Street Journal (WSJ) corpus: two one-million word parsed corpora in the University of Pennsylvania Treebank (Marcus, Santorini, & Marcinkiewicz, 1993). In particular, Gibson and Schütze counted the number of noun phrases conjoined to the first or second of three preceding noun phrase attachment sites, where the conjoined noun phrase contains a pronoun as its head (e.g., 'one,' 'ones') or in a pre-head position (e.g., 'its,' as in "Steele's comment on Swift's change of parties and its effect on their friendship"). This count is presented in Table 1 under the heading 'Pronouns.' In both corpora the pattern of frequencies follows the general trend (see 'All Instances' in Table 1) that was found by Gibson et al. (1996): more middle attachments than high attachments.

Gibson and Schütze concluded that the presence of the pronoun was not a likely explanation for the mismatch between corpus frequency and comprehension complexity. Nonetheless, their counts are extremely small, and it is doubtful that they can give any reliable information. Moreover, these counts comprise all kinds of pronouns: not only the anaphor 'one' is included, but also other pronouns that are in the head or pre-head position of the conjoined noun phrase. However, pronouns which are used for definite reference could differ in their use from 'one' anaphora, which blend properties of definite and indefinite reference (an indefinite NP is one which introduces a new entity into the discourse context, whereas a definite NP refers to an entity that is identifiable, for instance because it has already been mentioned in the discourse context). We therefore looked at how many of their counts contained structures that had the anaphor 'one' or 'ones' in the conjoined noun phrase. As can be seen in Table 1 (see the row labeled 'One/Ones'), the use of the anaphor 'one' in conjoined noun phrases to three possible attachment

Table 1
Corpus frequencies of middle and high attachments in the three site conjunction ambiguity in the Brown and Wall Street Journal corpus

	Brown corpus		WSJ corpus		Total	
	Middle	High	Middle	High	Middle	High
All instances	107	54	88	68	195	122
Pronouns	9	4	7	3	16	7
One/Ones	1	0	0	1	1	1

sites is extremely rare. Both for the Brown and the WSJ corpus we found only one occurrence of the anaphor 'one/ones' within the conjoined noun phrase. Given these overall low numbers, we think that no substantial claims can be based on these specific corpus searches, apart from the observation that the pronoun 'one' is almost never used in conjoined noun phrases with three possible attachment sites.

In another attempt to figure out what influence 'one' anaphora could have on noun phrase conjunction, we performed an additional corpus analysis. In this search we included all noun phrase conjunctions that could attach to two possible attachment sites. We did this because this ambiguity is much more frequent than the three-site ambiguity and might give us more reliable numbers. The results of this corpus search are presented in the top half of Table 2.

As can be seen, the low attachments (2544 instances) substantially outnumber the high attachments (1196 instances). Most interestingly, when we look at those ambiguities including the pronoun 'one,' this pattern is reversed. Now the high attachments (30 instances) outnumber the low attachments (7 instances), both in the Brown and the WSJ corpus.

In a final attempt to generalize these findings, we performed a search in which we included all ambiguous noun phrase conjunctions; i.e. all structures with more than one possible attachment site (see bottom half of Table 2). This showed that when the pronoun 'one' was present in the conjoined noun phrase, this noun phrase was most often attached to the first mentioned attachment site. In the Brown corpus, 29 out of 34 ambiguous noun phrase conjunctions were attached to the first noun phrase, whereas only 5 were attached to any of the following possible noun phrase attachment sites. In the Wall Street Journal, 13 out of 21 were attached to the first mentioned noun phrase, while 8 were attached to any of the remaining attachment sites. It appears that the general pattern is one in which an ambiguous noun phrase conjunction containing the word 'one' prefers to be attached high (total: 42 out of 55).

These numbers are rather small, but when taken together with the literature on anaphor resolution (Chang, 1980; Corbett & Chang, 1983; Gernsbacher, 1989; Gernsbacher & Hargreaves, 1988; Gordon et al., 1993) and relative clause attachment in German (Hemforth et al., 2000) they suggest that the presence of the anaphor 'one' in the Gibson and Schütze (1999) items induced a high attachment preference. In the following self-paced reading experiment we investigated this possibility more directly.

Experiment 1

In this experiment the disambiguation with the pronoun 'one' was directly compared with a disambiguation that did not contain this pronoun, using the same procedure as Gibson and Schütze (1996). Going through the corpus sentences that were collected by Gibson et al. (1996), we noticed that most conjunction ambiguities were disambiguated by a parallel noun phrase, i.e. a noun phrase that was similar in syntactic structure or semantics or both as the noun phrase it is conjoined with. To illustrate this point, some high attachment (see 4) and middle attachment examples (see 5) from the Brown corpus are provided below.

(4) Examples of high attached conjoined NPs from the Brown corpus
   (a) *about one-fourth* of *the total food calories* in *the American diet* and about one-third of the total nutrients consumed by all livestock and poultry
   (b) *the leader* of *the Israelite exodus* from *Egypt* and the leadership of the Puritan clergy in colonial New England
   (c) *the Government* of *the United States* of *America* and the Government of India

(5) Examples of middle attached conjoined NPs from the Brown corpus
   (a) *the son* of *Mrs. James Baines* of *Los Angeles*, *California*, and Carl E. Howard of Santa Monica, California
   (b) *the distinction* between *the Soviet zone* of *Germany* and the Soviet sector of Berlin

Table 2
Corpus frequencies of low and high attachments in other conjunction ambiguities in the Brown and Wall Street Journal corpus

	Brown corpus		WSJ corpus		Total	
	Low	High	Low	High	Low	High
Two site conjunction ambiguity						
All instances	1508	604	1036	592	2544	1196
One/Ones	2	22	5	8	7	30
All conjunction ambiguities (>1 attachment site)						
One/Ones	5	29	8	13	13	42

(c) *the merger* of *the Public Relations Society* of *America* and the American Public Relations Association

Based on this general pattern, in the items that did not contain a pronoun we disambiguated using a noun phrase that was similar both in meaning and in its syntactic structure to the noun phrase it was conjoined with. An example item is given in (6). In the items that contained the pronoun, we simply replaced the conjoined noun phrase by 'one,' as in (7).

(6) A column about a soccer team from the suburbs and...
   (a) an article about a baseball team from the city were published in the Sunday edition.
   (b) a baseball team from the city was published in the Sunday edition.

(7) A column about a soccer team from the suburbs and...
   (a) one about a baseball team from the city were published in the Sunday edition.
   (b) one from the city was published in the Sunday edition.

This design allowed us to investigate the hypothesis that the high attachment preference in the Gibson and Schütze (1999) study was due to the presence of the pronoun 'one.' According to this hypothesis, we should replicate the high attachment preference with items having the form of (7), but in accordance to the corpus frequencies we should find evidence for a middle attachment preference with items like (6).

Method

Participants. Forty-eight students and other affiliates from MIT and the surrounding community were paid for their participation. All were native English speakers and were naive as to the purposes of the study.

Materials and design. Twenty-four sets of sentences were constructed, each with four conditions, crossing the type of conjunction (either with the pronoun 'one' or with a parallel NP) and the attachment site (either high or middle). An example item was given in (6–7). The general form of the items is shown in (8).

(8) NP1 Prep1 NP2 Prep2 NP3 and
   (a) Parallel—high: NP4 Prep1 NP5 Prep2 NP6 were Verb...
   (b) Parallel—middle: NP5 Prep2 NP6 was Verb...
   (c) One—high: (the) one Prep1 NP5 Prep2 NP6 were Verb...
   (d) One—middle: (the) one Prep2 NP6 was Verb...

Overall, the parallel NPs in the high conjunctions (NP4) had the same frequency as the NPs in the middle conjunctions (NP5). The NPs parallel to NP1 had a mean log frequency of 1.46 versus a mean log frequency of 1.37 for the NPs parallel to NP2 ($t < 1$).

The items were very similar in form to the high- and middle-attachment items that were presented in Gibson and Schütze (1999). However, in both of their experiments, the three attachment sites were placed in object position. In the present study, the syntactic structure containing the three attachment sites was the subject of the sentence. This change to the items was made for two reasons. First, we wanted to avoid a potential ambiguity in which one of the prepositional phrases could be interpreted as a dependent of the matrix verb phrase.

One of the items used by Gibson and Schütze that contains such an ambiguity is presented in (9).

(9) The costume designer drew a sketch of a dress with a zipper and one with a belt to show the director.

Upon encountering the first instance of "with a" it is unclear whether the prepositional phrase will attach to the verb 'drew' (as in "The costume designer drew a sketch of a dress with a pencil") or to the NP "a dress," as in (9).

Second, and more importantly, including the conjoined NP before the VP created the possibility of another disambiguation cue: When following an NP1 attachment the verb was plural and following an NP2 attachment the verb was singular. Several studies have shown that subject–verb agreement is an early component of sentence comprehension (e.g., Nicol, Forster, & Veres, 1997; Osterhout & Mobley, 1995; Pearlmutter, Garnsey, & Bock, 1999; Sevald & Garnsey, 1995).

In addition to the subject–verb agreement cue, two further cues disambiguated the attachment site for the conjoined NP to the high or middle attachment site. First, as in Gibson and Schütze (1999) the preposition following the word 'one' or the parallel NP was the same as the preposition following the high or middle attachment site. For instance, the preposition 'about' following 'an article' in (6a) and following 'one' in (7a) supports the high attachment because the high attachment site is also followed by the preposition 'about.' Second, plausibility information also helped to disambiguate the attachment to the middle or high attachment site. For example, the prepositional phrase 'about a baseball team' in (7a) can attach to the high attachment site 'a column' because it is plausible for a column to be about a soccer team and a baseball team. However, this prepositional phrase cannot attach to the middle site 'a soccer team,' because it makes no sense for a soccer team to be about a baseball team.

Participants were presented with a yes/no question following each sentence to make sure they were reading for comprehension. In order to determine whether the difficulty to obtain the correct interpretation depended on the attachment of the NP conjunction, the comprehension questions for 16 of the 24 experimental

sentences were designed to probe for this interpretation. Specifically, the comprehension question asked whether there were one or two of the objects named by the high or middle attachment site. If high attachment was required, then there were two objects corresponding to NP1 and one corresponding to NP2. If middle attachment was required, then there was only one object corresponding to NP1 and two corresponding to NP2. For instance, after reading sentence (6) or (7) participants could be asked "Did the Sunday edition publish any columns that was about only one team?"

Each participant saw only one of the four possible versions of each test sentence, according to a Latin square. In addition to the 24 experimental items, 168 filler items were presented, including 28 items from another experiment that contained quite complex VP-ellipses. For each participant, the sentences were presented in a pseudorandomized order, in that at least one filler item was shown between two experimental items. All 24 sets of experimental sentences can be found in Appendix A.

Procedure. As in Gibson and Schütze (1999), the participants performed a word-by-word self-paced noncumulative moving-window reading task (Just, Carpenter, & Woolley, 1982) controlled by a Macintosh computer running software developed in the lab. The Macintosh display allowed for up to 100 characters to appear on each line. This way, in all conditions the part where the sentence became ambiguous (the word "and") up to the point where the sentence was completely disambiguated (the main verb of the sentence) could be presented on the same line (the second line). The three attachment sites always appeared on the first line. A trial began with a series of dashes marking the position of the words in the sentences. Participants pressed the space bar to reveal each subsequent word in the sentence and to revert the previous word to dashes. The amount of time the participant spent reading each word was recorded as the time between key-presses. A comprehension question appeared after the final word of each sentence. Participants had to respond 'yes' or 'no' by pressing one of two keyboard keys. When they answered a question incorrectly, the word 'INCORRECT' flashed briefly on the screen. No feedback was given for correct responses. Participants were asked to read the sentences at a natural rate and to be sure that they understood what they read. They were instructed to answer the questions as quickly and accurately as possible and to take wrong answers to the comprehension questions as an indication to read more carefully. Before the main experiment started, two screens of instructions and a short list of eight practice items and questions was presented in order to familiarize the participants with the self-paced reading task. It took approximately 40 min for participants to complete the experiment.

Results

Comprehension question performance. The mean response accuracies in percentages for both middle- and high-attachment conditions as a function of type of disambiguation are presented in Table 3.

Table 3 only includes the answers to the sixteen questions that asked about the appropriate interpretation for the sentence, because these questions can give us an idea how hard or easy it is to obtain the high-or middle attachment interpretation in each of the conditions. A two-factor ANOVA was performed, crossing type of disambiguation and attachment site. None of the effects in the analyses over items reached significance (all $F2$s < 2.20, all p-values $> .15$), probably because of the small number of items that were analyzed. The analysis over participants revealed a significant main effect of attachment site ($F1(1,47) = 10.87,\ p < .01$), such that comprehension questions about the high attachments were answered correctly more often than comprehension questions about the middle attachments (73% versus 61%). There was no main effect of disambiguation type (Fs < 1), but there was an interaction between attachment site and disambiguation type ($F1(1,47) = 4.33,\ p < .05$). This interaction demonstrates that the advantage of high attachments over middle attachments was mainly due to the conditions in which the pronoun 'one' was present (a significant difference of 18%: $F1(1,47) = 14.53,\ p < .001$; $F2(1,15) = 4.14,\ p = .06$), and less so in the cases where the attachment was indicated by a parallel NP (a nonsignificant difference of 6%: $F1(1,47) = 2.32$, $p = .14$; $F2 < 1$).

As indicated by the accuracy on the comprehension questions, participants seemed to have a hard time understanding some versions of the sentences. We believe it is not likely that participants simply did not read for comprehension. First, response accuracy for the eight remaining questions which did not ask for the conjunction interpretation was relatively high (85% in all four conditions). Second, response accuracy was equally high on the non-experimental filler items (85%). It is more probable that participants were thrown off by the attachment ambiguity. Especially reading the middle attachment conditions of sentences containing a pronoun, accuracy was virtually at chance

Table 3
Mean response accuracy in percentage for middle and high attachments as a function of disambiguation type

Attachment site	Disambiguation type	
	And one	Parallel NP
High	74	72
Middle	56	66
	+18	+6

Table 4
Mean residual reading times in milliseconds (and corresponding raw reading times in milliseconds per word) as a function of disambiguation type and attachment site for each of the six regions of Experiment 1

	Regions					
	1	2	3	4	5	6
Pronoun disambiguation						
High	−20	−19	−57	−47	−47	−13
	(362)	(344)	(326)	(348)	(320)	(391)
Middle	−18	−13	−55	−12	−18	36
	(363)	(351)	(322)	(375)	(350)	(442)
High advantage	2	6	2	35	29	49
Parallel disambiguation						
High	−29	−20	−49	−31	−52	−12
	(352)	(356)	(333)	(365)	(317)	(390)
Middle	−26	−43	−47	−24	−49	0
	(354)	(341)	(331)	(363)	(318)	(407)
High advantage	−3	−23	−2	7	3	12

(56%). Note, however, that in the original Gibson and Schütze (1999) study, participants that were drawn from the same student pool performed even worse in this condition (39 and 41% in their Experiments 1 and 2, respectively). Therefore, as in Gibson and Schütze, we only analyzed the reading time data of those sentences that received a correct answer to the comprehension question.

Reading times. To adjust for differences in word length across conditions as well as overall differences in participants' reading rates, a regression equation predicting reading times from word length was derived for each participant, using all filler and target sentences in the experiment (Ferreira & Clifton, 1986; see Trueswell et al., 1994, for discussion). As in the Gibson and Schütze study, at each word position, the reading time predicted by the participant's regression equation was subtracted from the actual measured raw reading time to obtain a residual reading time.[1] Analyses on the raw reading times showed a very similar pattern as the analyses on the residual reading times, although the statistics were not as strong.

For the analyses of the reading time data, the sentences were separated into six regions, as is illustrated in (10). Region 1 included the three NPs that were the potential attachment sites. Region 2 were the words "and one" in those items that contained the pronoun. In the other items this region consisted of the parallel NP, which was the first disambiguation cue in these items. Region 3 consisted of the following PP, which was the first disambiguation cue in the "and one" items. It was also a further disambiguation cue in the parallel items. Region 4 consisted of the main verb of the sentences, which further disambiguated the attachment ambiguity by the number of the copula ('was' or 'were'). Region 5 consisted of the two words following the main verb, and Region 6 was the rest of the sentence. The mean residual reading times for each of these regions (and the corresponding raw reading times) as a function of disambiguation type (pronoun versus parallel disambiguation) and attachment site (high versus middle) are presented in Table 4.

(10a) Regions of analysis for the high-attached disambiguations

Region 1	A column about a soccer team from the suburbs
Region 2	Pronoun disambiguation: and one Parallel disambiguation: and an article
Region 3	about a baseball team from the city
Region 4	were published
Region 5	in the
Region 6	Sunday edition.

(10b) Regions of analysis for the middle-attached disambiguations

Region 1	A column about a soccer team from the suburbs
Region 2	Pronoun disambiguation: and one Parallel disambiguation: and a baseball team
Region 3	from the city
Region 4	was published
Region 5	in the
Region 6	Sunday edition.

[1] Most of the resulting mean residual reading times came out to be negative. One possible reason is that a considerable number of the filler items contained syntactic structures that were even harder to understand than the conjunction ambiguity. Given a response accuracy of 85% on those questions that did not probe for the intended attachment, it is unlikely that participants read the sentences too fast to properly understand them. Note that even on those questions asking for the correct attachment accuracy was well above chance in three of the four conditions.

The main prediction of the present study was an interaction between disambiguation type and attachment site at or following the disambiguating regions. We expected to replicate the high attachment preference in the "and one" items, but we expected a middle attachment preference in the parallel items. In order to look for this interaction we performed a repeated measures ANOVA with two factors (disambiguation type, containing the levels pronoun disambiguation and parallel disambiguation, and attachment site, containing the levels high and middle attachment) on the residual reading times of the entire sentence excluding Region 1 (NP1–NP3), where the sentences are identical in all conditions. This analysis revealed that there was no main effect of disambiguation type ($F1$ and $F2 < 1.14$): Sentences containing the pronoun disambiguation (on average −31 ms) were read as fast as sentences containing the parallel disambiguation (on average −37 ms). Also the main effect of attachment site was not significant, although the analysis showed a tendency for the high conjunctions to be read faster than the middle conjunctions (−40 ms versus −27 ms; $F1(1,47) = 3.96$, $p = .05$; $F2(1,23) = 3.93$, $p = .06$). As predicted, the analysis revealed an interaction between attachment site and disambiguation type, but only in the analysis over participants ($F1(1,47) = 5.69$, $p < .05$) and not in the analysis over items ($F2(1,23) = 2.54$, $p = .12$). Planned comparisons revealed that in the "and one" disambiguation condition sentences with a high conjunction were read faster than sentences with a middle conjunction (−43 ms versus −18 ms; $F1(1,47) = 7.73$, $p < .01$; $F2(1,23) = 5.04$; $p < .05$), but that there was no difference between high conjunctions (−38 ms) and middle conjunctions (−35 ms) in the parallel disambiguation condition ($F1$ and $F2 < 1$). Next we performed similar two-factor repeated measures ANOVAs on each of these five regions separately in order to find out which regions contributed to this interaction.

Region 2. The ANOVA that was performed on the second region ("and one/NP") revealed a main effect of disambiguation type which was significant in the analysis over participants ($F1(1,47) = 4.71$, $p < .05$) and marginal in the analysis over items ($F2(1,23) = 3.35$, $p = .08$). Participants spent less time reading the string "and NP" than "and one." There was a tendency of a main effect of attachment site (middle attachments faster than high attachments) in the analysis over items ($F2(1,23) = 4.12$, $p = .05$), but this comparison was not significant in the analysis over participants ($F1(1,47) = 1.21$, $p = .28$). However, these main effects were qualified by an interaction between disambiguation type and attachment site, which was significant over participants but not over items ($F1(1,47) = 4.93$, $p < .05$; $F2(1,23) = 1.10$, $p = .31$). Planned comparisons indicated that the parallel NP that was conjoined to the middle site tended to be read faster than the parallel NP that was conjoined high ($F1(1,47) = 5.56$, $p < .05$; $F2(1,23) = 3.23$, $p = .09$), but that in sentences containing the pronoun, high conjunctions did not differ from middle conjunctions ($F1$ and $F2 < 1$).

Region 3. A two-factor ANOVA on Region 3 (the PP) revealed no main effect of disambiguation type ($F1(1,47) = 1.21$, $p = .28$; $F2(1,23) = 2.45$, $p = .13$), no main effect of attachment site ($F1$ and $F2 < 1$) and no interaction ($F1$ and $F2 < 1$). Furthermore, planned comparisons revealed that high conjunctions did not differ from middle conjunctions (all Fs < 1) in either the "and one" or the parallel disambiguation sentences.

Region 4. At Region 4, the main verb, there was also no significant main effect of disambiguation type ($F1$ and $F2 < 1$), no significant main effect of attachment site ($F1(1,47) = 2.10$, $p = .15$; $F2(1,23) = 1.84$, $p = .19$) and no significant interaction ($F1$ and $F2 < 1.56$). Planned comparisons did not reveal a difference between high and middle conjunctions in the parallel disambiguation sentences ($F1$ and $F2 < 1$). However, in the "and one" sentences there was a marginal effect of attachment site in the analysis over participants ($F1(1,47) = 2.94$, $p = .09$), but not in the analysis over items ($F2(1,23) = 1.44$, $p = .24$). Here, high conjunctions were read faster than middle conjunctions.

Region 5. The ANOVA at Region 5, the two words following the main verb, revealed a significant main effect of disambiguation type ($F1(1,47) = 7.19$, $p < .05$; $F2(1,23) = 5.82$, $p < .05$). This region was read faster following the parallel disambiguation compared to the pronoun disambiguation. There was also a tendency of a main effect of attachment site such that the high conjunctions were read faster than middle conjunctions ($F1(1,47) = 3.62$, $p = .06$; $F2(1,23) = 2.89$, $p = .10$). However, there was also an interaction between disambiguation type and attachment site which was significant in the analysis over participants ($F1(1,47) = 5.56$, $p < .05$), but marginal in the analysis over items ($F1(1,23) = 3.55$, $p = .07$). Planned comparisons showed that the advantage of high over middle conjunctions is restricted to the sentences with the pronoun disambiguation ($F1(1,47) = 6.69$, $p < .05$; $F2(1,23) = 3.55$, $p = .07$). There was no difference at all between high and middle conjunctions in sentences containing a parallel NP ($F1$ and $F2 < 1$).

Region 6. At the end of the sentence there was no main effect of disambiguation type ($F1(1,46) = 1.30$, $p = .26$; $F2(1,22) = 1.63$, $p = .21$). There was a marginal main effect of attachment site in the analysis over participants ($F1(1,46) = 3.76$, $p = .06$), but this effect was not significant in the analysis over items ($F2(1,22) = 2.15$, $p = .16$). Although there was no significant interaction ($F1(1,46) = 1.90$, $p = .17$; $F2(1,22) = 1.31$, $p = .27$), planned comparisons showed that the marginally significant advantage of high conjunctions over middle conjunctions is restricted to

the pronoun disambiguation condition ($F1(1,47) = 4.34$, $p < .05$; $F2(1,22) = 2.59$, $p = .12$) and that there was no difference in the parallel disambiguation condition ($F1$ and $F2 < 1$).

Discussion

This self-paced reading experiment demonstrated that the presence or absence of the pronoun 'one' in a noun phrase can influence the preference to conjoin the noun phrase with the first or second of three possible noun phrase attachment sites. When the pronoun was present, we replicated the findings of Gibson and Schütze (1999): conjunctions to the middle attachment site were more difficult to process than conjunctions to the high attachment site. If we look at Table 4, we see that this effect started to show up numerically at the disambiguating verb (a marginally significant high attachment advantage of 35 ms at Region 4) and becomes significant at the next two words (29 ms at Region 5) and the end of the sentence (49 ms at Region 6) However, when we disambiguated the conjunction ambiguity by using parallel noun phrases, the only effect that we obtained was that the conjunctions to the middle attachment site tended to be read faster than conjunctions to the high attachment site (a middle attachment advantage of 23 ms at Region 2), although this effect did not quite reach significance in the items analysis.

These data—together with the corpus counts—provide suggestive evidence contrary to the claims made by Gibson and Schütze (1999). Nonetheless, there are some aspects about the present experiment that obscure the interpretation. Most importantly, the present results were not statistically robust. For instance, even though the interaction between disambiguation type and attachment site was reliable in the analysis over participants, it did not reach standard levels of significance in the analyses over items. Second, while the finding that with parallel disambiguation middle conjoined NPs were read faster than high conjoined NPs could indicate that middle conjunctions are easier to process than high conjunctions, there is an alternative possible source of the middle conjunction preference in this experiment. In particular, it is possible that the processing difference between middle and high conjunctions reflects lexical properties of the NPs, i.e. a difference in word recognition processes instead of a difference in attachment complexity. Because the lexical frequency of the attaching head noun was controlled across the two conditions, lexical frequency cannot account for the observed difference. But it remains possible that there might have been more semantic priming in the middle-attached nouns than in the high-attached nouns. Although we think that this is unlikely, because attaching nouns were all selected to be closely related to the head noun in the NP site that they were designed to conjoin to, we cannot exclude this possibility entirely.

These two issues were addressed in Experiment 2, in which we tried to replicate the previous findings in Dutch using eye-tracking technology. If the interaction between disambiguation type and attachment site originates from a general characteristic of pronoun resolution, we should be able to replicate it in another language. In addition, the reading situation in an eye-tracking experiment is ecologically more valid. Specifically, participants have the opportunity to reread parts of a sentence when they have trouble understanding the sentence. Given the difficulty that is experienced with conjunctions to three attachment sites, this may be an important advantage over the self-paced reading task. Furthermore, eye-tracking could allow us to gain more insight into the nature of the effect. In particular, this method may allow us to discover whether the interaction is an effect that originates from the influence of information that is immediately assimilated or from delayed difficulty in processing. This is especially important given the possible lexical origin of the middle conjunction preference in the parallel disambiguation items. It has been shown that variables related to word recognition, such as word frequency (e.g., Henderson & Ferreira, 1990, 1993; Just & Carpenter, 1980; Rayner, 1977; Rayner & Raney, 1996) and semantic priming effects of words read earlier in the sentence (Carroll & Slowiaczek, 1986; Morris, 1994) are reflected in early measures of eye-movement data, such as gaze duration and first-fixation duration (see Rayner, 1998, for an extensive overview). If lexical priming is responsible for the middle conjunction preference, then this should show up in early measures of the eye-movement data on the specific NPs that were used. However, if we find evidence for an advantage of middle conjunctions over high conjunctions later in the sentence or on eye-movement measures such as percentage of regressions, it is less likely that such an explanation could be responsible for the interaction that was found in Experiment 1.

Experiment 2

Method

Participants. A total of 32 undergraduate students from Ghent University participated for course credit. All were native Dutch speakers and were unaware of the aim of the research. All participants had normal, uncorrected vision or wore contact lenses.

Materials and design. The 24 sets of sentences were Dutch translations of the English sentences used in Experiment 1. However, we made some changes to the items in order to make them somewhat easier to

process. First, the three potential NP attachment sites were no longer in subject position, but were placed in object position. This way, participants had read the subject and the main verb of the sentence before they encountered the ambiguous part of the sentence. This also means that number agreement of the main verb was no longer an extra disambiguation cue. However, in the original Gibson and Schütze (1999) study this cue was also absent and they still found the high attachment advantage. Second, the potential NP attachment sites were only made up of a determiner and a noun. In the English items, a considerable number of NPs contained an adjective as well. Leaving out the adjective made the ambiguous and disambiguating components of the sentences a little shorter and easier to read. Apart from these two changes, the items were similar in construction and content as the English sentences. Sentences (11) and (12) illustrate the Dutch sentences that were constructed on the basis of sentences (6) and (7), used in Experiment 1.

(11) De krant publiceerde een column over een voetbalploeg uit de randgemeente en...
  (a) een artikel over een basketbalploeg uit de hoofdstad omdat het sportseizoen er opnieuw aankomt.
  (b) een basketbalploeg uit de hoofdstad omdat het sportseizoen er opnieuw aankomt.
  [The newspaper published a column about a soccer team from the suburbs and (an article about) a basketball team from the capital because the sports season is about to start.]

(12) De krant publiceerde een column over een voetbalploeg uit de randgemeente en...
  (a) één over een basketbalploeg uit de hoofdstad omdat het sportseizoen er opnieuw aankomt.
  (b) één uit de hoofdstad omdat het sportseizoen er opnieuw aankomt.
  [The newspaper published a column about a soccer team from the suburbs and one (about a basketball team) from the capital because the sports season is about to start.]

As in Experiment 1, the parallel NPs that were conjoined high had the same frequency as the parallel NPs that were conjoined to the middle NP (1og frequency of 1.22 versus 1.36, respectively; $t < 1$). Participants saw only one of the four possible versions according to a Latin square design. In addition to the 24 experimental sentences, the stimulus list contained 96 filler sentences of the same length, including items from other unrelated experiments. All 24 sets of Dutch experimental sentences (and their English translations) can be found in Appendix B.

Procedure. A SMI Eyelink headband-mounted eye-tracking system was used to record participants' eye movements. The Eyelink system samples both the horizontal and vertical signal every 4 ms and is based on an infrared video-based tracking technology that happens simultaneously for both eyes. Participants were seated at a distance of 75 cm from a 17-inch display. Although the Eyelink system compensates for head position, this compensation is not accurate enough to allow single character resolution. Therefore, participants were asked to put their head on a height-adjustable chin rest and to move as little as possible. A practice session was included to allow participants to become familiar with the eye-tracking equipment and the experimental procedure.

Both the practice session and the experimental session started with a calibration and validation procedure. In the calibration procedure the participants were asked to fixate nine calibration points that were presented randomly one at the time in the form of a 9-point grid. The calibration was evaluated by a built-in routine and each eye's calibration was graded "good," "poor," or "failed." Only when the calibration of both eyes was graded "good" the validation procedure was started. The validation procedure assesses the accuracy of the system in predicting gaze position from pupil position. In the validation phase, nine target points are presented in the same way as in the calibration procedure. When the participants fixated these, the calibration values were used to estimate the gaze position of the participant, and the error (i.e., the difference between the target position and the computed gaze position) was calculated. As in the calibration procedure, each eye was graded separately and was accepted only when the maximal distance between the target position and the computed gaze position did not exceed 0.5° for each of the nine target points. It was not possible to obtain a good grade of validation for 4 participants after considerable effort. These participants were replaced by 4 new participants.

After the calibration and validation procedure was completed, the sentences were presented in a different random order for each participant. Each trial started with a calibration check (with a single fixation point in the center of the screen) and was adjusted in case the check was negative. Participants were asked to read each sentence as it was presented, and push a button when they had finished. The sentences were presented as in Experiment 1; i.e. all ambiguous and disambiguating parts of the sentence appeared on the same line. To encourage participants to read the sentences for meaning, 30 of the 120 sentences (25%) were followed by a yes-no question, posed by the experimenter. The experimenter told them whether they had answered the question correctly or not. We did not ask a question after each sentence, because wearing the eye-tracker becomes uncomfortable after a while, so we tried to keep the experimental session as short as possible. Moreover, the main goal of asking questions was to make sure the participants were reading the sentences for comprehen-

sion. Therefore, the questions were comparable to those questions of Experiment 1 that were not about the syntactic ambiguity that was the subject of the present investigation. For instance, in relation to the example sentence (11–12), participants could be asked whether it was true that the sports season was near its end. The experiment started with a practice session consisting of 8 practice sentences of which two were followed by a question. The entire experiment took about 40 min.

Results

The comprehension questions were answered correctly in 96% of the experimental trials and 90% of the filler trials, which means that participants were reading the sentences for comprehension. The fact that response accuracy was considerably higher than in Experiment 1 can be explained by the fact that we made the sentences easier to understand (see Materials and Design), by the fact that the comprehension questions in the present experiment were not about the conjunction ambiguity, but about other parts of the sentence (see Procedure section), and by the fact that the eye-tracking procedure allowed for rereading, whereas the self-paced reading experiment did not.

The reading time data were analyzed using regions that—at least for the ambiguous and disambiguating parts of the sentences—were as similar as possible to the ones used in the self-paced reading experiment. The target sentences were divided into eight regions, illustrated in (13). Region 1 included the subject and the main verb of the sentence (e.g., "The newspaper published"). Regions 2, 3, and 4 consisted of the first (NP1), the second (NP2), and the third attachment site (NP3), respectively. Region 5 was identical to Region 2 in the self-paced reading experiment. It consisted of the words "en één [and one]" in those items that contained the pronoun and of the parallel NP in the other items. Region 6 was identical to Region 3 in the self-paced reading experiment and consisted of the following PP. Region 7 included the following two words (Region 5 of the self-paced reading study) and Region 8 was the rest of the sentence (Region 6 of the self-paced reading study).

(13a) Regions of analysis for the high-attached disambiguations

Region 1	De krant publiceerde
Region 2	een column
Region 3	over een voetbalploeg
Region 4	uit de randgemeente
Region 5	Pronoun disambiguation: en één Parallel disambiguation: en een artikel
Region 6	over een basketbalploeg uit de hoofdstad
Region 7	omdat het
Region 8	sportseizoen er opnieuw aankomt.

(13b) Regions of analysis for the middle-attached disambiguations

Region 1	De krant publiceerde
Region 2	een column
Region 3	over een voetbalploeg
Region 4	uit de randgemeente
Region 5	Pronoun disambiguation: en één Parallel disambiguation: en een basketbalploeg
Region 6	uit de hoofdstad
Region 7	omdat het
Region 8	sportseizoen er opnieuw aankomt.

As in Experiment 1, analyses were performed on residual reading times (see Trueswell et al., 1994, for a discussion of using residual reading times instead of the milliseconds per character adjustment for eye-movement data). Analyses on the raw reading times resulted in numerically and statistically similar patterns.

Total reading times. We first analyzed Total Reading Times (TRT) in order to see how much time participants spent reading each region. The TRT is defined as the sum of all fixations within a specific region. Table 5 presents the mean residual and raw TRTs of the eight regions as a function of attachment site (high versus middle attachment) for the disambiguations containing a pronoun and the parallel disambiguations. As in Experiment 1, we expected to find an interaction between disambiguation type and attachment site. In order to look whether this interaction was significant, we performed similar two-factor repeated-measures analyses of variance on each region as in the self-paced reading experiment.

At Region 1, which made up the beginning of the sentence, there was no main effect of disambiguation type ($F1(1,31) = 1.15,\ p = .29; F2(1,23) = 1.15,\ p = .29$), no main effect of attachment site ($F1$ and $F2 < 1$), and no significant interaction between both variables (both $F < 1.03$). Planned comparisons showed that there were no reading time differences between high and middle attachments, neither in the pronoun or the parallel disambiguation condition (all Fs < 1).

The next three regions, Region 2, 3, and 4, contained the three attachment NPs. Although we did not predict any effects on these regions, it could not be ruled out that people looked back to the attachment sites more in those conditions that they found difficult to understand, resulting in higher TRTs on these attachment sites. This was not the case, however. There was no main effect of disambiguation type (Region 2: $F1(1,31) = 1.41$, $p = .23$; $F2 < 1$; Region 3: $F1(1,31) = 1.24,\ p = .27$; $F2 < 1$; Region 4: $F1$ and $F2 < 1$), nor was there a main effect of attachment site (Region 2: $F1 < 1$; $F2(1,23) = 1.36,\ p = .26$; Region 3: $F1$ and $F2 < 1$; Region 4: $F1(1,31) = 1.66,\ p = .21$; $F2(1,23) = 1.31,\ p = .26$), nor was there an interaction between disambiguation

Table 5
Mean residual TRTs (and corresponding raw TRTs) in milliseconds as a function of disambiguation type and attachment site for each of the eight regions of Experiment 2

	Regions							
	1	2	3	4	5	6	7	8
Pronoun disambiguation								
High	−24	−33	69	127	−18	83	−126	23
	(799)	(451)	(689)	(735)	(282)	(1238)	(316)	(1219)
Middle	−53	−27	28	130	29	98	−16	17
	(773)	(456)	(649)	(741)	(329)	(740)	(427)	(1212)
High advantage	−29	6	−41	3	47	15	110	−6
Parallel disambiguation								
High	−77	−46	15	147	−9	−12	−116	−23
	(749)	(437)	(639)	(757)	(583)	(1138)	(323)	(1176)
Middle	−55	−55	34	76	−28	−47	−76	−36
	(773)	(427)	(658)	(687)	(532)	(589)	(361)	(1167)
High advantage	22	−9	19	−71	−19	−35	40	−13

type and attachment site (Region 2: $F1$ and $F2 < 1$; Region 3: $F1 < 1$; $F2(1,23) = 1.57$, $p = .22$; Region 4: $F1(1,31) = 1.68$, $p = .20$; $F2(1,23) = 1.68$, $p = .21$). Nevertheless, planned comparisons showed that there was a middle attachment advantage in the parallel disambiguation condition at Region 4, which was significant in the analysis over participants ($F1(1,31) = 5.84$, $p < .05$) and marginal in the analysis over items ($F2(1,23) = 3.58$, $p = .07$).

As in Experiment 1, we found a significant interaction between disambiguation type and attachment site at Region 5, the region in which the conjunction and the pronoun/NP were presented (which corresponds to Region 2 in Experiment 1). In the present experiment this interaction was significant both in the analysis over participants ($F1(1,31) = 4.60$, $p < .05$) and the analysis over items ($F2(1,23) = 4.76$, $p < .05$). There was no main effect of disambiguation type ($F1(1,31) = 1.52$, $p = .23$; $F2(1,23) = 2.41$, $p = .13$) or attachment site ($F1$ and $F2 < 1$). Planned comparisons revealed that the significant interaction was due to the fact that residual reading times for the high and middle attachment did not differ in the parallel disambiguation condition ($F1 < 1$; $F2(1,23) = 1.76$, $p = .20$), but differed significantly in the pronoun disambiguation condition ($F1(1,31) = 9.59$, $p < .01$; $F2(1,23) = 8.31$, $p < .01$).

At Region 6, the disambiguating PP, the numerical pattern of results was similar to the one in the previous region: high attachments were read faster than middle attachments in the pronoun disambiguation condition, whereas high attachments were read slower than middle attachments in the parallel disambiguation condition. However, as in Experiment 1 (at the corresponding Region 3), the interaction was not significant ($F1$ and $F2 < 1$). Also, the main effect of attachment site was not significant ($F1$ and $F2 < 1$): both the high attachment advantage of 15 ms in the pronoun disambiguation condition and the middle attachment advantage of 35 ms in the parallel disambiguation condition were not significant (Fs < 1). Unlike in Experiment 1, there was a significant main effect of disambiguation type ($F1(1,31) = 11.04$, $p < .01$; $F2(1,23) = 12.42$, $p < .01$), showing that participants spent more time reading the disambiguating PP following a pronoun than when it followed a full NP. This could reflect the extra processing that is needed to establish the referent of the pronoun. Unlike in the parallel disambiguation conditions, participants have to figure out what 'één [one]' stands for. In example item (11–12), for instance, 'één [one]' means another column (in the high attachment condition) or another soccer team (in the middle attachment condition). Alternatively, as we concluded in the discussion of the corpus data, the use of a pronoun in the three-site conjunction ambiguity is extremely rare, which could explain why people have more difficulty in reading the pronoun disambiguation compared to a full NP disambiguation.

Region 7 consisted of the two words following the disambiguating PP. Based on the results of Experiment 1, we expected an interaction to show up here. Indeed, there was an interaction between disambiguation type and attachment site and in this experiment it was significant both in the $F1$ and $F2$ analysis ($F1(1,31) = 5.42$, $p < .05$; $F2(1,23) = 5.56$, $p < .05$). There was no main effect of disambiguation type ($F1(1,31) = 2.64$, $p = .11$; $F2(1,23) = 1.84$, $p = .19$), but there was a main effect of attachment site ($F1(1,31) = 20.23$, $p < .001$; $F2(1,23) = 19.87$, $p < .001$). Planned comparisons showed that the interaction was due to the fact that high attachments were read significantly faster than middle attachments in the pronoun disambiguation condition ($F1(1,31) = 22.72$, $p < .001$; $F2(1,23) =$

33.28, $p < .001$), whereas reading times of high and middle attachments did not differ in the parallel disambiguation condition ($F1(1,31) = 3.06$, $p = .10$; $F2(1,23) = 2.23$, $p = .15$).

At Region 8, the end of the sentence, there was no significant interaction ($F1$ and $F2 < 1$) and no main effect of attachment site ($F1$ and $F2 < 1$). There was a trend towards a main effect of disambiguation type, but it was not quite significant ($F1(1,31) = 3.05$, $p = .09$; $F2(1,23) = 3.53$, $p = .07$). As in Region 6, participants seemed to spend more time reading the end of the sentence when the sentence contained a pronoun compared to when it contained a full NP. Planned comparisons showed that high and middle attachments did not differ in either of the disambiguation conditions (all Fs < 1).

The pattern that emerged when calculating TRTs seems to closely follow the pattern found in the self-paced reading experiment: high attachments were substantially easier to read than middle attachments at Regions 5 and 7, but only when there was a pronoun in the conjoined NP. When the sentence was disambiguated by a parallel NP, the only significant effect was a middle advantage at Region 4. However, the interactions in total reading times could be due either to differences in first-pass reading times or in the amount of regressive eye-movements made by the participants. We therefore analyzed both variables. First-Pass Reading Time (FPRT) was defined as the sum of fixations between the first entrance of a region and the first exit, either to the left or the right, provided that the region has been fixated during first-pass reading. Table 6 shows the mean residual and raw FPRT for each of the eight regions as a function of disambiguation type and attachment site. Percentage of regressions is defined as the number of trials in which the eyes leave a region to the left, relative to the number of trials this region has been looked at during first-pass reading. Mean percentages of regressions for the eight regions as a function of disambiguation type and attachment site are presented in Table 7.

Table 6
Mean residual FPRTs (and corresponding raw FPRTs) in milliseconds as a function of disambiguation type and attachment site for each of the eight regions of Experiment 2

	Regions							
	1	2	3	4	5	6	7	8
Pronoun disambiguation								
High	34	−28	66	98	−52	−169	−79	131
	(621)	(312)	(506)	(530)	(154)	(659)	(230)	(991)
Middle	3	−37	5	103	−66	−121	−39	−102
	(592)	(304)	(444)	(536)	(143)	(333)	(269)	(752)
High advantage	−31	−9	−61	5	−14	48	40	−233
Parallel disambiguation								
High	5	−35	16	124	44	19	−72	98
	(596)	(305)	(459)	(558)	(463)	(845)	(235)	(964)
Middle	20	−45	38	64	43	−9	−61	4
	(610)	(295)	(481)	(496)	(439)	(442)	(248)	(871)
High advantage	15	−10	22	−60	−1	−28	11	−94

Table 7
Mean percentage of regressions as a function of disambiguation type and attachment site for each of the eight regions of Experiment 2

	Regions							
	1	2	3	4	5	6	7	8
Pronoun disambiguation								
High	0	10	7	23	1	41	9	54
Middle	0	12	10	18	1	52	9	67
High advantage	0	2	3	−5	0	11	0	13
Parallel disambiguation								
High	0	8	6	19	7	13	10	59
Middle	0	15	8	23	4	8	10	49
High advantage	0	7	2	4	−3	−5	0	−10

First-pass reading times. An ANOVA on the data for Region 1 revealed no main effect of disambiguation type ($F1$ and $F2 < 1$), no main effect of attachment site ($F1$ and $F2 < 1$), and no interaction between both variables ($F1(1,31) = 3.07,\ p = .09;\ F2(1,23) = 1.47,\ p = .24$). The same was true for Region 2 (disambiguation type: $F1$ and $F2 < 1$; attachment site: $F1 < 1$; $F2(1,23) = 1.26,\ p = .27$; interaction: $F1$ and $F2 < 1$). Planned comparisons showed that high and middle attachments did not differ in any of the conditions of Region 1 and 2 (all Fs < 1.11, all $p > .24$).

At Region 3, there were also no main effects of disambiguation type (F1 and $F2 < 1$) and attachment site ($F1(1,31) = 1.13,\ p = .30; F2(1,23) = 2.23,\ p = .15$). However, at this region the analyses revealed a significant interaction between disambiguation type and attachment site ($F1(1,31) = 7.16,\ p < .05; F2(1,23) = 5.98,\ p < .05$). Planned comparisons showed that the NP2 was read faster in middle attachment sentences than in high attachment sentences, but only in the pronoun condition ($F1(1,31) = 7.08,\ p < .05; F2(1,23) = 6.50,\ p < .05$) and not in the parallel disambiguation condition ($F1 < 1; F2(1,23) = 1.14,\ p = .30$).

At Region 4 there was no main effect of disambiguation type ($F1$ and $F2 < 1$), no main effect of attachment site ($F1(1,31) = 2.62,\ p = .12; F2(1,23) = 2.98,\ p = .10$), and no interaction ($F1(1,31) = 2.22,\ p = .15$; $F2(1,23) = 2.48,\ p = .13$). However, planned comparisons showed that the significant middle attachment advantage in the parallel disambiguation that we found in the analysis on the TRTs originated from FPRTs ($F1(1,31) = 5.07,\ p < .05;\ F2(1,23) = 6.32,\ p < .05$). In the pronoun disambiguation there was no difference between high and middle attachments ($F1$ and $F2 < 1$). The significant effects in the FPRTs of Regions 3 and 4 are hard to explain theoretically because the sentences were completely identical in all of the conditions up until Region 4 and FPRTs can not have been influenced by later parts of the sentence.

Recall that we found a significant interaction in total reading times during Region 5. The present analysis on this region revealed that this interaction in TRTs was not due to an interaction in first-pass reading times ($F1$ and $F2 < 1$). Also there was no main effect of attachment site in this region ($F1$ and $F2$). However, the analysis revealed a main effect of disambiguation type ($F1(1,31) = 25.95,\ p < .001;\ F2(1,23) = 47.78,\ p < .001$), such that participants were faster to read the pronoun "een [one]" in the pronoun disambiguation conditions than the full NPs that were used in the parallel conditions.

In Region 6, on the disambiguating PP, there was also no significant interaction between disambiguation type and attachment site in the FPRTs ($F1(1,31) = 2.74,\ p = .11;\ F2(1,23) = 1.40,\ p = .25$), nor a significant main effect of attachment site ($F1$ and $F2 < 1$). As in the analysis on total reading times, there was a significant main effect of disambiguation type ($F1(1,31) = 17.76,\ p < .001;\ F2(1,23) = 31.79,\ p < .001$), which showed that participants found it harder to read the PP following the pronoun compared to when it followed a full NP.

On Region 7, there was no main effect of disambiguation type ($F1$ and $F2 < 1$), but there was a main effect of attachment site ($F1(1,31) = 4.01,\ p = .05;\ F2(1,23) = 4.39,\ p < .05$), which showed that high attachments were read faster than middle attachments. As was the case for Region 5, the interaction we found on Region 7 on the total reading times was not due to a significant interaction in FPRTs ($F1(1,31) = 1.93,\ p = .17;\ F2(1,23) = 1.64,\ p = .21$). Nevertheless, planned comparisons revealed that the high attachment advantage was significant in the pronoun disambiguation condition ($F1(1,31) = 5.85,\ p < .05; F2(1,23) = 6.06,\ p < .05$) and was totally absent in the parallel disambiguation condition ($F1$ and $F2 < 1$), which could explain some part of the interaction in the TRTs at this region.

At Region 8 there was no main effect of disambiguation type ($F1(1,31) = 1.34,\ p = .25;\ F2(1,23) = 1.52,\ p = .23$), but a significant main effect of attachment site ($F1(1,31) = 26.03,\ p < .001;\ F2(1,23) = 16.25,\ p < .001$). There was also an interaction between disambiguation type and attachment site, which was significant in the analysis over items ($F2(1,23) = 6.59,\ p < .05$) and marginal in the analysis over participants ($F1(1,31) = 3.12,\ p = .09$). Interestingly, the interaction was opposite to the one we found on total reading times: first-pass reading times were shorter for the middle attachment sentences than for high attachment sentences in the pronoun disambiguation condition ($F1(1,31) = 26.76,\ p < .001;\ F2(1,23) = 27.76,\ p < .001$) and they did not differ in the parallel disambiguation condition ($F1(1,31) = 2.51,\ p = .12;\ F2(1,23) = 2.82,\ p = .11$). One explanation for this effect could be that in the middle attachment sentences containing a pronoun, participants quickly decided they would have to reread parts of the sentence in order to fully understand the meaning of the sentence. Together with the fact that the interactions in TRTs could not be explained by first-pass reading times, this gave us an extra reason to calculate percentage of regressions.

Regressions. Again, at Region 2, 3, and 4 we did not expect to find any effects because the experimental sentences were identical in all conditions. At Region 2, there was no main effect of disambiguation type ($F1$ and $F2 < 1$) and no significant interaction ($F1$ and $F2 < 1.2$). There was a marginal main effect of attachment site ($F1(1,31) = 2.57,\ p = .12;\ F2(1,23) = 5.69,\ p < .05$), but this effect was not very robust: Planned comparisons showed that the high and middle attachments did not differ in the pronoun disambigua-

tion condition ($F1$ and $F2 < 1$) nor in the parallel disambiguation condition ($F1(1,31) = 2.65,\ p = .11$; $F2(1,23) = 3.98,\ p = .06$).

At Regions 3 and 4, there were no significant main effects of disambiguation type (Region 3: $F1$ and $F2 < 1$; Region 4: $F1$ and $F2 < 1$), attachment site (Region 3: $F1(1,31) = 1.24,\ p = .27$; $F2(1,23) = 1.90$, $p = .18$; Region 4: $F1$ and $F2 < 1$) and no significant interaction (Region 3: $F1$ and $F2 < 1$; Region 4: $F1(1,31) = 3.37,\ p = .08$; $F2(1,23) = 3.20,\ p = .09$). Planned comparisons also did not reveal any significant differences between high and middle attachments (all Fs < 1.90, all $p > .18$).

At Region 5, there was no main effect of disambiguation type ($F1(1,31) = 1.71,\ p = .20$; $F2 < 1$), no main effect of attachment site ($F1$ and $F2 < 1$), and no significant interaction ($F1 < 1$; $F2(1,23) = 3.09,\ p = .09$). High and middle attachments did not differ in the pronoun disambiguation condition ($F1$ and $F2 < 1.2$) and the parallel disambiguation condition ($F1(1,31) = 1.23,\ p = .28$; $F2(1,23) = 2.36,\ p = .14$).

At Region 6, the disambiguating PP, analyses revealed a significant interaction between disambiguation type and attachment site ($F1(1,31) = 9.40,\ p < .01$; $F2(1,23) = 8.51,\ p < .01$) and a significant main effect of disambiguation type ($F1(1,31) = 81.20,\ p < .001$; $F2(1,23) = 133.87,\ p < .001$). There was no main effect of attachment site (F1 and $F2 < 1.34$). Planned comparisons revealed that the interaction was due to the fact that participants made more regressions when reading the middle attachment compared to the high attachment in the pronoun disambiguation condition ($F1(1,31) = 8.36,\ p < .01$; $F2(1,23) = 5.23,\ p < .05$), but not in the parallel disambiguation condition, where the numerical pattern was reversed but not significantly so ($F1(1,31) = 1.83,\ p = .19$; $F2(1,23) = 2.39$, $p = .14$).

At Region 7, the two words following the PP, there were no significant main effects or interaction (all Fs < 1). Also planned comparisons revealed no effects (all Fs < 1).

Finally at Region 8, the end of the sentence, there was an interaction between disambiguation type and attachment site ($F1(1,31) = 11.87,\ p < .01$; $F2(1,23) = 11.75,\ p < .01$). There was a tendency towards a main effect of disambiguation type ($F1(1,31) = 3.10,\ p = .09$; $F2(1,23) = 3.41,\ p = .08$) but no effect of attachment site ($F1$ and $F2 < 1$). Again, planned comparisons showed that participants made more regressions when the sentence contained a middle attachment compared to a high attachment in the pronoun disambiguation condition ($F1(1,31) = 5.50,\ p < .05$; $F2(1,23) = 8.62,\ p < .01$). The reverse pattern in the parallel disambiguation condition reached significance at the end of the sentence ($F1(1,31) = 4.49,\ p < .05$; $F2(1,23) = 4.25,\ p = .05$).

Discussion

The results of the self-paced reading study presented in Experiment 1 were replicated in this eye-movement experiment. It was shown again that the presence of a pronoun influences the noun phrase conjunction ambiguity. The advantage of high noun phrase conjunctions over middle noun phrase conjunctions that was demonstrated by Gibson and Schütze (1999) was shown to be completely dependent on the presence of the pronoun 'one.' When the pronoun is replaced by a parallel NP, the only significant effect that could be found was an advantage of processing middle conjunctions over high conjunctions. Since the interaction induced by the pronoun was replicated using different participants and different items, it is safe to state that the effects found in Experiment 1 are reliable. Moreover, the fact that the interaction was found in another language and could be generalized to a more natural reading situation, seems to indicate that the interaction is due to a general characteristic of processing the pronoun 'one.'

Using the eye-tracking procedure allowed us to gain a more detailed insight in the nature of the interaction. The pattern found in the self-paced reading experiment was most closely mirrored in the total reading times. The interaction in TRTs was significant at the same regions as in Experiment 1: the region containing "and one/NP" (Region 2 in Experiment 1 and Region 5 in Experiment 2) and the two words following full disambiguation (Region 5 in Experiment 1 and Region 7 in Experiment 2). Further analyses revealed that the interaction reflected in the total reading times was mainly due to interactions in the percentages of regression, rather than to an interaction in the first-pass reading times (although part of the interaction could arguably be explained by somewhat shorter first-pass reading times in high pronoun attachments in the region following the PP). Middle attachments led to more regressions from the disambiguating PP and from the end of the sentence than high attachments in the pronoun disambiguation conditions, but exactly the reversed pattern was found when the sentences were disambiguated with a full NP.

According to some researchers, first-pass reading time reflects the influence of information that is rapidly assimilated by the reader, whereas regressive eye movements reflect a processing barrier that cannot immediately be overcome (Crain, Ni, Shankweiler, Conway, & Braze, 1996; Ehrlich & Rayner, 1983; Ferreira & Henderson, 1990; Frazier & Rayner, 1982; Kennedy, 1983). Whereas some researchers believe regressive eye movements to reflect syntactic re-analysis processes (e.g., Ferreira & Henderson, 1990; Frazier & Rayner, 1982), others use the range of eye-movement measures to investigate the time course of different categories of constraints. Boland and Blodgett (2001), for instance, presented participants with sentences containing noun/

verb homographs (e.g., "duck") preceded by a context that was biased towards the noun or verb interpretation. Although the study revealed both effects of lexical bias and discourse congruency, these two types of constraint yielded quite different eye-movement patterns. While the lexical bias effect was mainly reflected in the initial fixations on the homograph, discourse effects were observed in the percentage of regressions several words following the homograph, and consequently in second-pass reading times. According to Boland and Blodgett, their data challenge syntactic processing models in which all available constraints influence parsing immediately and simultaneously. The present results could be argued to corroborate this suggestion. Our data seem to suggest that at least one form of discourse-based processing, namely pronoun resolution—in which the human sentence processing mechanism has to identify the appropriate antecedent for the pronoun within the discourse context—does not come into play in the initial phases of sentence processing.

The greater variety of dependent reading measures in the eye-tracking experiment compared to the self-paced reading experiment also provides us with a better basis to evaluate the alternative lexical explanation of the middle conjunction preference, that was addressed in the discussion of Experiment 1. In the self-paced reading experiment we found that the NP that was parallel to the middle attachment site was read faster than the NP that was parallel to the high attachment site. Since the NPs are different we could not completely exclude the possibility that this difference reflects lexical properties of the NPs, even though we argued this to be unlikely, given that both NPs were equally frequent and that we selected NPs that were semantically highly related in both conditions. However, such an alternative lexical explanation is even less likely for the effect we found in Experiment 2, where people made significantly more regressive eye movements at the end of high-attachment sentences than at the end of middle-attachment sentences. Since the effect did not emerge on the NPs themselves, but appeared several words downstream and only in percentage of regressions, this effect cannot be explained by differences in lexical access processes on the parallel NPs in the different conditions.

General discussion

In 1999, Gibson and Schütze published a study that has been widely taken as evidence against exposure-based accounts of sentence processing. They argued that for at least one syntactic ambiguity—the conjunction of noun phrases to three potential noun phrase sites—the human sentence parsing mechanism does not use corpus frequencies in arriving at its preference. Therefore they state that the decision principles of sentence comprehension and sentence production must be partially distinct. The data presented in the present study show that the high attachment preference in sentence comprehension found in Gibson and Schütze was completely due to the use of the pronoun 'one' in their items. The only significant effects found in sentences without a pronoun reflected an advantage of middle attachments over high attachments. These data cast serious doubt on Gibson and Schütze's hypothesis that different processes guide preferences in production and comprehension for this kind of ambiguity.

The data reported here are as predicted by two different kinds of accounts. First, the data are most clearly explained by Hemforth et al.'s (2000) anaphoric binding hypothesis, according to which the presence of a pronoun in the ambiguous region induces a high attachment preference, because the parser prefers to coindex pronouns with linguistic elements which belong to the main assertion of a sentence. Hence, coindexations with NPs that are highest in the syntactic structure are preferred. This is probably the most robust finding of the present study. The pattern in which a high attachment preference was induced by the presence of the pronoun was observed both in sentence production (corpus analysis) and sentence comprehension (both on-line reading experiments). Moreover the effect could be generalized both over methods (self-paced reading and eye movements) and languages (English and Dutch).

Second, these data can be argued to be consistent with experience-based accounts of sentence processing, such as the tuning hypothesis (e.g., Mitchell et al., 1995), constraint-based models (e.g., McRae, Spivey-Knowlton, & Tanenhaus, 1998), connectionist models (e.g., Rohde, 2002), or probabilistic models (e.g., Jurafsky, 1996). These theories predict—explicitly or implicitly—that the disambiguation patterns observed in corpora should be reflected in on-line sentence comprehension. The present data are consistent with these theories because they show that there is a grain size at which corpus frequencies of the NP conjunction ambiguity correspond with on-line comprehension measures. In addition, these data constrain experience-based sentence processing accounts. For instance, consider the tuning hypothesis of Mitchell et al. (1995). In order for this hypothesis to remain viable, it must be reformulated slightly, such that pronominal counts must be tabulated, in addition to tabulations of purely structural levels of information (see Desmet et al., 2002a, for a similar conclusion). Similarly, constraint-based models that make predictions concerning the resolution of syntactic ambiguities containing a pronoun—such as the relative clause attachment ambiguity or the noun phrase conjunction ambiguity with the pronoun 'one'—must include a constraint (1) that is sensitive to pronominal binding and (2) whose strength can be calculated on the basis of corpus counts. Together with the study of Desmet et al. (2002a),

the present paper has refuted the most important evidence showing differential patterns of corpus frequencies and sentence comprehension, namely the studies concerning the RC attachment ambiguity in Dutch by Mitchell and Brysbaert (1998) and the three-site conjunction ambiguity in English by Gibson and Schütze (1999).

In sum, the present study supports the suggestion that pronominal resolution processes are important to the way the human sentence processing mechanism resolves syntactic ambiguities. Taking this factor into account, we were able to refute an important piece of evidence for making a distinction between sentence comprehension preferences and corpus frequencies in specific, and sentence comprehension and sentence production in general. We can conclude that for this syntactic ambiguity there is no discrepancy between online preferences and corpus frequencies. Consequently, there is no need to assume different processes underlying sentence comprehension and sentence production on the basis of the noun phrase conjunction ambiguity.

Acknowledgments

Timothy Desmet is a Research Assistant of the Fund for Scientific Research—Flanders (Belgium) (F.W.O.—Vlaanderen). A major part of the work presented in this paper was carried out during a stay of Timothy Desmet as a Visiting Student at the Department of Brain and Cognitive Sciences at MIT, which was supported by the Fund for Scientific Research—Flanders (Belgium) (F.W.O.—Vlaanderen). We thank Don Mitchell, Marc Brysbaert, and two anonymous reviewers for their comments.

Appendix A

This appendix contains all 24 sets of English sentences used in Experiment 1. The first version of each item (a) contains the noun phrases that are conjoined to the first of the three possible NP sites (high conjunctions). The second version of each item (b) contains the noun phrases that are conjoined to the second of the three possible NP sites (middle conjunctions). In the pronoun disambiguation, the pronoun "one" was used. In the parallel NP disambiguation, the full NP between parentheses was used.

1a. A column about a soccer team from the suburbs and (one/an article) about a baseball team from the city were published in the Sunday edition.

1b. A column about a soccer team from the suburbs and (one/a baseball team) from the city was published in the Sunday edition.

2a. A joke about a bald man with a hair piece and (one/an anecdote) about a fat woman with a funny hat were told by the clown but nobody laughed.

2b. A joke about a bald man with a hair piece and (one/a fat woman) with a funny hat was told by the clown but nobody laughed.

3a. The book about a heroic soldier in the American Revolution and (the one/the novel) about a famous lieutenant in the Civil War were discussed by the history teacher.

3b. The book about a heroic soldier in the American Revolution and (one/a famous lieutenant) in the Civil War was discussed by the history teacher.

4a. A paper by an influential scientist from the midwest and (one/an article) by a prominent professor from the northeast were summarized by the student.

4b. A paper by an influential scientist from the midwest and (one/a prominent professor) from the northeast was summarized by the student.

5a. A sketch of a red dress with a long zipper and (one/a drawing) of a black coat with a leather belt were copied by the costume designer.

5b. A sketch of a red dress with a long zipper and (one/a black coat) with a leather belt was copied by the costume designer.

6a. The suspect of the grand theft in the museum and (the one/the victim) of the armed robbery in the gallery were interviewed by the police.

6b. The suspect of the grand theft in the museum and (the one/the armed robbery) in the gallery was interviewed by the police.

7a. A comedy about a poor jester with a large pimple and (one/a tragedy) about a rich king with a hunchback were written by Shakespeare.

7b. A comedy about a poor jester with a large pimple and (one/a rich king) with a hunchback was written by Shakespeare.

8a. The paperback with a red stain on the front cover and (the one/the textbook) with a small tear on the first page were removed by the shop owner.

8b. The paperback with a red stain on the front cover and (one/a small tear) on the first page was removed by the shop owner.

9a. A party for the majority candidate for the presidency and (one/a dinner) for the independent candidate for the senate were held in New York City.

9b. A party for the majority candidate for the presidency and (the one/the independent candidate) for the senate was held in New York City.

10a. The legend about the flying horse with the golden saddle and (the one/the fable) about the talking dog with the silver leash were based on stories from east Asia.

10b. The legend about the flying horse with the golden saddle and (the one/the talking dog) with the silver leash was based on stories from east Asia.

11a. A contestant for the $10,000 award for the best fiction and (one/a candidate) for the $5,000 prize for the best non-fiction were congratulated by a member of the jury.

11b. A contestant for the $10,000 award for the best fiction and (the one/the $5,000 prize) for the best non-fiction was congratulated by a member of the jury.

12a. The volume with the main conclusion of the Russian researcher and (the one/the issue) with the basic assumption of the American scientist were sold immediately by the publisher.

12b. The volume with the main conclusion of the Russian researcher and (the one/the basic assumption) of the American scientist was sold immediately by the publisher.

13a. A story about a small house with a pond and (one/a tale) about a huge mansion with a garden were printed in a kids' magazine.

13b. A story about a small house with a pond and (one/a huge mansion) with a garden was printed in a kids' magazine.

14a. A movie about a young woman with breast cancer and (one/a documentary) about a retired widower with a skin disease were listed in the TV guide.

14b. A movie about a young woman with breast cancer and (one/a middle-aged widower) with a skin disease was listed in the TV guide.

15a. A tiger with a white spot on the forehead and (one/a lion) with a black stripe on the back were filmed by a biologist.

15b. A tiger with a white spot on the forehead and (one/a black stripe) on the back was filmed by a biologist.

16a. A photo of a little child on the beach and (one/a picture) of an old man in the ocean were hanging on my aunt's wall.

16b. A photo of a little child on the beach and (one/an old man) in the ocean was hanging on my aunt's wall.

17a. The note about a meeting with a supplier and (the one/the message) about an appointment with a customer were given to the supervisor.

17b. The note about a meeting with a supplier and (one/an appointment) with a customer was given to the supervisor.

18a. An inspection by a private detective from the company and (one/an inquiry) by a chief investigator from the syndicate were ordered by the judge.

18b. An inspection by a private detective from the company and (one/a chief investigator) from the syndicate was ordered by the judge.

19a. The problem with the national election in the late fifties and (the one/the difficulty) with the state election in the early sixties were clearly underestimated by the senate.

19b. The problem with the national election in the late fifties and (the one/the state election) in the early sixties was clearly underestimated by the senate.

20a. A test for the graduate students in chemistry and (one/an exam) for the undergraduates in physics were given in the big auditorium.

20b. A test for the graduate students in chemistry and (the ones/the undergraduates) in physics was given in the big auditorium.

21a. An operation by the voluntary platoon in the south and (one/an action) by the professional squadron in the north were commanded by the old general.

21b. An operation by the voluntary platoon in the south and (the one/the professional squadron) in the north was commanded by the old general.

22a. A discount on the green truck in the showroom and (one/a reduction) on the red convertible on the lot were offered for this weekend only.

22b. A discount on the green truck in the showroom and (the one/the red convertible) on the lot was offered for this weekend only.

23a. The scholarship for a bright student with a poor family and (the one/the contribution) for a black child with a bad background were the main achievements of the charity organization.

23b. The scholarship for a bright student with a poor family and (one/a black child) with a bad background was the main achievement of the charity organization.

24a. A reduction in the property tax for laborers and (one/a cutback) in the interest rate for investors were promised by the president.

24b. A reduction in the property tax for laborers and (the one/the interest rate) for investors was promised by the president.

Appendix B

This appendix contains all 24 sets of Dutch sentences used in Experiment 2. The first version of each item (a) contains the noun phrases that are conjoined to the first of the three possible NP sites (high conjunctions). The second version of each item (b) contains the noun phrases that are conjoined to the second of the three possible NP sites (middle conjunctions). In the pronoun disambiguation, the pronoun "één" was used. In the parallel NP disambiguation, the full NP between parentheses was used.

1a. De krant publiceerde een column over een voetbalploeg uit de randgemeente en (één/een artikel) over een basketbalploeg uit de hoofdstad omdat het sportseizoen er opnieuw aankomt.

1b. De krant publiceerde een column over een voetbalploeg uit de randgemeente en (één/een basketbalploeg) uit de hoofdstad omdat het sportseizoen er opnieuw aankomt.

[1. The newspaper published a column about a soccer team from the suburbs and an article about a basketball team from the capital because the sports season is about to start.]

2a. De clown vertelde een grap over een man met een pet en (één/een anekdote) over een vrouw met een hoed maar niemand kon erom lachen.

2b. De clown vertelde een grap over een man met een pet en (één/een vrouw) met een hoed maar niemand kon erom lachen.

[2. The clown told a joke about a man with a cap and an anecdote about a woman with a hat but nobody laughed about it.]

3a. De leraar besprak een boek over een soldaat in een revolutie en (één/een roman) over een luitenant in een burgeroorlog zodat de leerlingen begonnen weg te dromen.

3b. De leraar besprak een boek over een soldaat in een revolutie en (één/een luitenant) in een burgeroorlog zodat de leerlingen begonnen weg te dromen.

[3. The teacher discussed a book about a soldier in a revolution and a novel about a lieutenant in a civil war so the students started to dream.]

4a. De student bestudeerde een paper van een wetenschapper uit het westen en (één/een artikel) van een professor uit het oosten als voorbereiding op het examen.

4b. De student bestudeerde een paper van een wetenschapper uit het westen en (één/een professor) uit het oosten als voorbereiding op het examen.

[4. The student studied a paper of a scientist from the West and an article of a professor from the East in preparation of his exam.]

5a. De kledingontwerper kopieerde een schets van een jurk met een rits en (één/een tekening) van een jas met een riem vooraleer hij naar huis vertrok.

5b. De kledingontwerper kopieerde een schets van een jurk met een rits en (één/een jas) met een riem vooraleer hij naar huis vertrok.

[5. The fashion designer copied a sketch of a dress with a zipper and a drawing of a coat with a belt before he went home.]

6a. De politie ondervroeg een verdachte van een diefstal in het museum en (één/een slachtoffer) van een overval in de kunstgalerij maar het gesprek leverde geen bruikbare tips op.

6b. De politie ondervroeg een verdachte van een diefstal in het museum en (één/een overval) in de kunstgalerij maar het gesprek leverde geen bruikbare tips op.

[6. The police interrogated a suspect of a theft in the museum and a victim of a robbery in the art gallery but the conversation did not lead to any useful clues.]

7a. De toneelschrijver bedacht een komedie over een nar met een wrat en (één/een tragedie) over een koning met een bochel maar de critici vonden er niks aan.

7b. De toneelschrijver bedacht een komedie over een nar met een wrat en (één/een koning) met een bochel maar de critici vonden er niks aan.

[7. The playwright thought of a comedy about a jester with a wart and a tragedy about a king with a hunchback but the critics didn't like it much.]

8a. De winkelbediende verwijderde een paperback met een vlek op de kaft en (één/een tekstboek) met een scheur op de eerste pagina omdat een klant geklaagd had.

8b. De winkelbediende verwijderde een paperback met een vlek op de kaft en (één/een scheur) op de eerste pagina omdat een klant geklaagd had.

[8. The shop assistant removed a paperback with a stain on the cover and a text book with a tear in the front page because a customer complained about it.]

9a. De stad organiseerde een feest voor een kandidaat voor het presidentschap en (één/een diner) voor een kandidaat voor de senaatsverkiezingen maar de zaal was veel te klein voor de vele genodigden.

9b. De stad organiseerde een feest voor een kandidaat voor het presidentschap en (één/een een kandidaat) voor de senaatsverkiezingen maar de zaal was veel te klein voor de vele genodigden.

[9. The city organized a party for a candidate for the presidency and a dinner for a candidate for the senate but the hall was too small for all the invited guests.]

10a. De tekst was gebaseerd op een legende over een paard met een gouden zadel en (één/een fabel) over een hond met een zilveren leiband maar daar was bijna niets meer van te merken.

10b. De tekst was gebaseerd op een legende over een paard met een gouden zadel en (één /een hond) met een zilveren leiband maar daar was bijna niets meer van te merken.

[The text was based on a legend about a horse with a gold saddle and a fable about a dog with a silver collar but almost nobody noticed anything about that.]

11 a. De jury roemde een kanshebber op een bekroning voor het beste gedicht en (één/een kandidaat) voor een prijs voor de beste roman voordat de winnaar bekend gemaakt werd.

11b. De jury roemde een kanshebber op een bekroning voor het beste gedicht en (één/een prijs) voor de beste roman voordat de winnaar bekend gemaakt werd.

[11. The jury praised a candidate for an award for the best poem and a candidate for a prize for the best novel before the winner was announced.]

12a. De uitgever verkocht een publicatie met een theorie van een onderzoeker en (één/een uitgave) met een stelling van een wetenschapper omdat het publiek daarvoor tegenwoordig erg groot is.

12b. De uitgever verkocht een publicatie met een theorie van een onderzoeker en (één/een stelling) van een wetenschapper omdat het publiek daarvoor tegenwoordig erg groot is.

[12. The publisher sold a publication with a theory of a researcher and an issue with a proposition of a scientist because there were a lot of people interested in that kind of publications.]

13a. In het tijdschrift verscheen een sprookje over een huis met een vijver en (één/een verzinsel) over een villa met een tuin omdat de rest van de gepubliceerde teksten nogal zwaarwichtig was.

13b. In het tijdschrift verscheen een sprookje over een huis met een vijver en (één/een villa) met een tuin omdat de rest van de gepubliceerde teksten nogal zwaarwichtig was.

[13. In the magazine there appeared a fairy tale about a house with a pond and an invention about a villa with a garden because the rest of the published texts were rather weightily.]

14a. De televisiegids vermeldde een film over een vrouw met borstkanker en (één/een documentaire) over een weduwe met een huidziekte maar het uur was verkeerd aangeduid.

14b. De televisiegids vermeldde een film over een vrouw met borstkanker en (één/een weduwe) met een huidziekte maar het uur was verkeerd aangeduid.

[14. The TV guide announced a movie about a woman with breast cancer and a documentary about a widow with a skin disease but the indicated time was wrong.]

15 a. De bioloog filmde een tijger met een vlek op het voorhoofd en (één/een leeuw) met een streep op de rug tijdens een exotische safari in Afrika.

15b. De bioloog filmde een tijger met een vlek op het voorhoofd en (één/een streep) op de rug tijdens een exotische safari in Afrika.

[15. The biologist filmed a tiger with a spot on the forehead and a lion with a stripe on the back during an exotic safari in Africa.]

16a. Aan de muur hing een foto van een kind op het strand en (één/één afbeelding) van een man in de zee omdat de muur anders zo kaal is.

16b. Aan de muur hing een foto van een kind op het strand en (één/een man) in de zee omdat de muur anders zo kaal is.

[16. On the wall there was a picture of a child on the beach and a print of a man in the sea because otherwise the wall was too empty.]

17a. De secretaresse overhandigde een bericht over een vergadering met een leverancier en (één/een boodschap) over een afspraak met een klant omdat de chef nogal vergeetachtig is.

17b. De secretaresse overhandigde een bericht over een vergadering met een leverancier en (één/een afspraak) met een klant omdat de chef nogal vergeetachtig is.

[17. The secretary handed a message about a meeting with a supplier and a message about an appointment with a customer because the boss was rather forgetful.]

18a. De rechter eiste een inspectie door een detective van het bedrijf en (één/een onderzoek) door een speurder van het gerecht vooraleer een oordeel te willen vellen.

18b. De rechter eiste een inspectie door een detective van het bedrijf en (één/een speurder) van het gerecht vooraleer een oordeel te willen vellen.

[18. The judge demanded an inspection by a detective of the company and an investigation by an investigator of the court before he wanted to pass a sentence.]

19a. De minister voorkwam een probleem tijdens een verkiezing in juni en (één/een relletje) tijdens een campagne in november zodat zijn populariteit steeg.

19b. De minister voorkwam een probleem tijdens een verkiezing in juni en (één/een campagne) in november zodat zijn populariteit steeg.

[19. The Minister prevented a problem during an election in June and a riot during a campaign in november which made him more popular.]

20a. De leraar verbeterde een toets voor een student uit het zesde jaar en (één/een examen) voor een leerling uit het eerste jaar omdat beiden ziek waren geweest tijdens de eerste zittijd.

20b. De leraar verbeterde een toets voor een student uit het zesde jaar en (één/een leerling) uit het eerste jaar omdat beiden ziek waren geweest tijdens de eerste zittijd.

[20. The teacher corrected a test of a student from sixth grade and an exam of a pupil from first grade because both had been sick before.]

21a. De generaal leidde een opertatie door een eskadron in het zuiden en (één/een actie) door een peloton in het noorden terwijl hij in rechtstreeks contact stond met de president.

21b. De generaal leidde een opertatie door een eskadron in het zuiden en (één/een peloton) in het noorden terwijl hij in rechtstreeks contact stond met de president.

[21. The general directed an operation by a squadron in the south and an action by a platoon in the north while he had a direct line with the president.]

22a. De autodealer gaf een korting op een truck in de toonzaal en (één/een reductie) op een cabriolet op de parking omdat de verkoop de laatste tijd nogal tegenviel.

22b. De autodealer gaf een korting op een truck in de toonzaal en (één/een cabriolet) op de parking omdat de verkoop de laatste tijd nogal tegenviel.

[22. The car salesman offered a discount on a truck in the showroom and a reduction on a convertible in the parking lot because sales were a bit disappointing lately.]

23 a. De organisatie schonk een beurs voor een student uit een arm gezin en (één/een bijdrage) voor een kind uit een slechte buurt omdat een inzameling onverwacht veel geld had opgebracht.

23b. De organisatie schonk een beurs voor een student uit een arm gezin en (één/een kind) uit een slechte buurt omdat een inzameling onverwacht veel geld had opgebracht.

[23. The organization gave a scholarship to a student from a poor family and a contribution to a child from a bad neighborhood because a collection had raised more money than was expected.]

24a. De minister beloofde een vermindering van een belasting voor werknemers en (één/een verlaging) van een rentevoet voor investeerders omdat het overschot op de begroting groter was dan verwacht.

24b. De minister beloofde een vermindering van een belasting voor werknemers en (één/een rentevoet) voor investeerders omdat het overschot op de begroting groter was dan verwacht.

[24. The Minister promised a decrease of a tax for employers and a reduction of an interest rate for investors because the national budget had an unexpected surplus.]

References

Boland, J. E., & Blodgett, A. (2001). Understanding the constraints on syntactic generation: Lexical bias and discourse congruency effects on eye movements. *Journal of Memory and Language, 45*, 391–411.

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

Carroll, P. J., & Slowiaczek, M. L. (1986). Constraints on semantic priming in reading: A fixation time analysis. *Memory & Cognition, 14*, 509–522.

Chang, F. R. (1980). Active memory processes in visual sentence comprehension: Clause effects and pronominal reference. *Memory & Cognition, 8*, 58–64.

Corbett, A. T., & Chang, F. R. (1983). Pronoun disambiguation—accessing potential antecedents. *Memory & Cognition, 11*, 283–294.

Corley, M. M. B. (1996). *The role of statistics in human sentence processing*. Unpublished PhD thesis, University of Exeter.

Crain, S., Ni, W., Shankweiler, D. P., Conway, L., & Braze, D. (1996). Meaning, memory, and modularity. *Proceedings of the NELS26 Sentence Processing Workshop, MIT Occasional Papers in Linguistics* (Vol. 9, pp. 27–44). Cambridge, MA: MIT Working Papers in Linguistics.

Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition, 30*, 73–105.

Cuetos, F., Mitchell, D. C., & Corley, M. M. B. (1996). Parsing in different languages. In M. Carreiras, J. E. Garcia-Albea, & N. Sabastian-Galles (Eds.), *Language processing in Spanish* (pp. 145–187). Hillsdale, NJ: Erlbaum.

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

Desmet, T., De Baecke, C., & Brysbaert, M. (2002b). The influence of referential discourse context on modifier attachment in Dutch. *Memory & Cognition, 30*, 150–157.

Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior, 22*, 75–87.

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*, 555–568.

Ford, M., Bresnan, J., & Kaplan, R. (1982). A competence-based theory of syntactic closure. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 727–796). Cambridge, MA: MIT Press.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*, 627–635.

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

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.

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

Gernsbacher, M. A., & Hargreaves, D. J. (1988). Accessing sentence participants—the advantage of 1st mention. *Journal of Memory and Language, 27*, 699–717.

Gibson, E., & Schütze, C. T. (1996, March). The relationship between the frequency and the perceived complexity of conjunction attachments: On-line evidence. Poster presented at the 9th Annual CUNY Conference on Human Sentence Processing, New York, USA.

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.

Gibson, E., Schütze, C. T., & Salomon, A. (1996). The relationship between the frequency and the complexity of linguistic structure. *Journal of Psycholinguistic Research, 25*, 59–92.

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

Hemforth, B., & Konieczny, L. (2002, March). Where pronouns and relative clauses differ: Information structure and binding preferences. Paper presented at the 15th Annual CUNY Conference on Human Sentence Processing, New York, USA.

Hemforth, B., Konieczny, L., & Scheepers, C. (2000). Syntactic attachment and anaphor resolution: The two sides of relative clause attachment. In M. Crocker, M. Pickering, & C. Clifton (Eds.), *Architectures and mechanisms for language processing* (pp. 259–281). Cambridge, UK: Cambridge University Press.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 417–429.

Henderson, J. M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47*, 201–221.

Howes, D. H., & Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology, 41*, 401–410.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science, 20*, 137–194.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 3*, 228–238.

Kennedy, A. (1983). On looking into space. In K. Rayner (Ed.), *Eye Movements in Reading: Perceptual and Language Processes* (pp. 237–251). New York: Academic Press.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.

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

Marcus, M. P., Santorini, B., & Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Pernn Treebank. *Computational Linguistics, 19*, 313–330.

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.

Mitchell, D. C. (1994). Sentence parsing. In M. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 375–409). New York: Academic Press.

Mitchell, D. C., & Brysbaert, M. (1998). Challenges to recent theories of crosslinguistic variation in parsing: Evidence from Dutch. In D. Hillert (Ed.), *Syntax and semantics: A crosslinguistic perspective* (pp. 313–335). San Diego, CA: Academic Press.

Mitchell, D. C., Brysbaert, M., Grondelaers, S., & Swanepoel, P. (2000). Modifier attachment in Dutch: Testing aspects of Construal Theory. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 493–516). Oxford: Elsevier.

Mitchell, D. C., & Cuetos, F. (1991). The origins of parsing strategies. In C. Smith (Ed.), *Current issues in natural language processing*. University of Texas, Austin: Centre for Cognitive Science.

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 (nonlexical) statistical records. *Journal of Psycholinguistic Research, 24*, 469–488.

Mitchell, D. C., & Holmes, V. (1985). The role of specific information about the verb in parsing sentences with local structural ambiguity. *Journal of Memory and Language, 24*, 542–559.

Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 92–103.

Nicol, J. L., Forster, K. I., & Veres, C. (1997). Subject–verb agreement processes in comprehension. *Journal of Memory and Language, 36*, 569–587.

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34*, 739–773.

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language, 41*, 427–456.

Preston, K. A. (1935). The speed of word recognition and its relation to reading ability. *Journal of General Psychology, 13*, 199–203.

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition, 4*, 443–448.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review, 3*, 238–244.

Rohde, D. L. T. (2002). A connectionist model of sentence comprehension and production. Unpublished PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9*, 487–494.

Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science, 23*, 569–588.

Sevald, C. A., & Garnsey, S. M. (1995). Safe syntax: Encapsulation of number-marking information in sentence comprehension. Paper presented at the Eighth Annual CUNY Conference on Human Sentence Processing, Tucson.

Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes, 12*, 211–272.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *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 garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 528–553.