Memory & Cognition
2002, 30 (1), 150-157

The influence of referential discourse context on modifier attachment in Dutch

TIMOTHY DESMET, CONSTANTIJN DE BAECKE, and MARC BRYSBAERT
Ghent University, Ghent, Belgium

In an eye-tracking experiment we investigated the influence of referential context on the attachment of a relative clause to two possible hosts (as in "Someone shot the servant of the actress who was on the balcony"). The attachment of the relative clause was disambiguated grammatically at the first word after the onset of the ambiguity in order to investigate immediate effects of discourse. The contexts had been verified in a sentence completion study to make sure that they induced a strong bias toward early or late closure. The results of the reading experiment, however, revealed no significant interaction of referential context with the attachment preference of the relative clause. The only robust and consistent effect we found was a preference for early closure, independent of the preceding context. These data favor accounts positing that referential context does not influence the initial attachment decision, but does play a role in later phases of sentence processing.

An important debate in syntactic processing research concerns the way in which sentence parsing is influenced by the surrounding discourse context. Syntax-first theorists claim that only syntactic principles guide the initial phase of analysis and that referential information is ignored during this phase (Brysbaert & Mitchell, 1996; De Vincenzi & Job, 1995; Ferreira & Clifton, 1986; Frazier, 1987; Frazier & Rayner, 1982; Konieczny & Hemforth, 1996; Mitchell, Cuetos, Corley, & Brysbaert, 1995; Rayner & Pollatsek, 1989). In contrast, discourse-sensitive theorists argue that the initial analysis can be influenced by referential information (Gibson, 1998; Ni, Crain, & Schankweiler, 1996; Spivey-Knowlton & Tanenhaus, 1998) and that discourse factors may even exert their influence before other factors come into play (e.g., Altmann, Garnham, & Dennis, 1992; Altmann & Steedman, 1988; Crain & Steedman, 1985).

The main reason why the issue has remained unresolved is that the empirical evidence presented by both syntax-first and context-sensitive accounts has been mixed and can often be questioned on methodological grounds. For instance, many experiments that demonstrated immediate effects of referential context are subject to the criticism that the effects of context were examined relatively late in the sentence. According to Mitchell, Corley, and Garnham (1992), a delay of even two or three words from the onset of the ambiguity could provide ample opportunity for the thematic controller to implement a discourse-based revision of the initial analysis. Analogously, studies revealing no effects of referential context on sentence parsing have been criticized because the specific contexts constructed by the researchers may not have been sufficiently biasing (Spivey-Knowlton & Tanenhaus, 1994).

Since Crain and Steedman's (1985) original study, most of the research on immediate discourse effects has centered around the complement clause versus relative clause ambiguity (e.g., Altmann, 1988; Altmann, Garnham, & Henstra, 1994; Mitchell & Corley, 1994; Mitchell et al., 1992; van Berkum, Brown, & Hagoort, 1999). In a sentence like "John told the girl that . . .", the word *that* introduces an ambiguity between a sentential complement (e.g., "John told the girl that he was having trouble with his car") and a relative clause (e.g., "John told the girl that had been calling to hang up").

Most recently, van Berkum et al. (1999) registered brain potentials while participants were reading sentences that were disambiguated either as a complement or as a relative clause. The sentences were preceded by a complement-favoring context (only one possible referent for "the girl") or a relative clause-favoring context (two possible referents for "the girl"). The presence of two referents favors a relative clause because such a clause singles out the intended referent. Van Berkum et al. (1999) measured whether participants' brain waves would show a P600/SPS effect when the disambiguating word of the target sentence was encountered (previous research had shown that such an effect occurs when participants are confronted with syntactic processing difficulties). The effect was found only when the continuation of the target sentence violated the expectancies raised by the previous context, as predicted by discourse-sensitive theories of sentence parsing. Unfortunately, van Berkum et al. (1999) presented all possible

T.D. is a research assistant of the Fund for Scientific Research, Flanders (F.W.O.-Vlaanderen). The research was supported in part by Grant G.0250.99 from the research program of the Fund for Scientific Research, Flanders (F.W.O.-Vlaanderen) to C.D.B. We thank two anonymous reviewers for their comments. Correspondence should be addressed to T. Desmet, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium (e-mail: timothy.desmet@rug.ac.be).

—Accepted by previous editorial team

combinations of contexts and target sentences to the same participants without filler items, so that participants may have used situation-specific processing strategies, the more because stimulus presentation rate was rather slow (see Brysbaert & Mitchell, 2000; see van Berkum, Hagoort, & Brown, 2000, for a reply to these criticisms).

Another type of ambiguity for which researchers have looked at referential discourse effects is the attachment of a relative clause (RC) to two possible host sites (as in the sentence "Someone shot the servant of the actress who was on the balcony"). Cuetos and Mitchell (1988) reported that English speakers showed a preference for low attachment (i.e., to "actress") in this construction, as predicted by the late-closure principle of Frazier's (1987) garden-path theory. However, Spanish speakers showed a preference for high attachment (i.e., to "the servant"), in contrast with the predictions of Frazier's theory. Subsequent research has confirmed the Spanish NP1-attachment bias for most non-English languages (for reviews, see Mitchell & Brysbaert, 1998; Mitchell, Brysbaert, Grondelaers, & Swanepoel, 2000; see also Carreiras & Clifton, 1999).

Zagar, Pynte, and Rativeau (1997) looked at the influence of referential context on this type of syntactic ambiguity. As in van Berkum et al. (1999), contexts were created that introduced one or several actresses, assuming that in a context of a single actress it is more likely to modify the servant (NP1) than the actress (NP2), whereas the reverse is true in a context of several actresses. However, contrary to the predictions of the discourse-sensitive theories, the reading time data of Zagar et al. did not reveal an influence of referential context: Reading times were consistently higher in the NP2-attachment condition than in the NP1-attachment condition irrespective of context. Discourse context had an effect only on the number of correct answers given to comprehension questions at the end of the sentence, making Zagar et al. conclude that context did influence the final interpretation of the sentence but did not affect the syntactic mechanism of relative clause attachment. On the basis of this study, researchers have excluded referential discourse as a factor contributing to initial modifier attachment in sentence parsing.

Given the discrepancy between van Berkum et al. (1999) and Zagar et al. (1997), it seemed worthwhile to repeat the latter study because it contained two possible methodological problems. The first is that texts were presented without filler materials, which may have drawn the participants' attention to the grammatical structure under investigation. The second is that Zagar et al.'s lack of a context effect could be interpreted as the result of rather weak contexts. Although the effect of context was significant in their sentence completion study, it was very small. When an early-biasing context was presented, 89% of the completions were NP1 attachments. A late-closure-biasing context only reduced this NP1-attachment preference to 85%.

To repeat Zagar et al. (1997), we constructed 30 pairs of sentences that were of the form "NP1 of NP2 relative pronoun VERB PP" (Table 1). The sentences were disambiguated by the verb that agreed in number with either NP1 or NP2. In half of the sentences NP1 was singular, whereas NP2 was plural. The reverse was true in the other half of the sentences. Verb number agreement was chosen as disambiguation because this allowed us to disambiguate the sentence on the basis of grammatical principles. We selected verbs for which the singular and the plural form had the same number of letters in order to control for word length. The NP1-attachment versions and NP2-attachment versions of the sentences differed only in the number of the disambiguating verb and were identical in all other respects. The verb was always the first word after the relative pronoun, so the sentence was disambiguated as early as possible.

Three types of contexts were constructed. The early-closure biasing contexts were made by providing two possible referents for NP1, whereas there was only one referent for NP2. Exactly the reverse pattern was used for the late-closure-biasing contexts. In the neutral contexts, none of the attachment hosts was mentioned (25 cases) or there was only one possible referent for both NPs (5 cases).

COMPLETION STUDY

A completion study was run to examine the power of the different referential contexts to influence the final interpretation of the ambiguous relative clause. On the basis of the consistent NP1 bias reported in Dutch for this type of construction (Brysbaert & Mitchell, 1996; Mitchell et al., 2000), we hypothesized that a majority of the sentences in the neutral condition would be completed by a relative clause referring to NP1. If the biasing contexts were strong enough, this initial NP1 bias should be magnified by the early-closure-biasing contexts, whereas the late-closure-biasing contexts should wipe out the bias or even turn it into a preference for NP2 attachments.

Method

Participants. Thirty undergraduate students from Ghent University participated in the experiment for course credit. All were native Dutch speakers and were naive as to the research question.

Materials and Design. Thirty test sentences were constructed, each sentence preceded by three different contexts (these materials can be found at http://allserv.rug.ac.be/~tidesmet/index.html). All test sentences had a version with an NP1 continuation and a version with an NP2 continuation (Table 1). In the completion experiment only the beginnings of the sentences were presented—that is, the part up to and including the relative pronoun (e.g., "The police interrogate the advisor of the politicians who . . ."). Sentences were preceded by one of the three different contexts: an early-closure-biasing context (two potential referents for NP1), a late-closure-biasing context (two potential referents for NP2), and a neutral context. Each participant saw only one of the possible contexts per sentence (according to a Latin square design). In addition to the target texts, 50 filler texts were included in the stimulus list to divert the participants' attention from the specific grammatical structure under investigation.

Procedure. The 80 texts were typed on sheets of paper. The participants were instructed to read each text and to complete the last sentence by writing down the first continuation that came to mind. The completion of the entire set of texts took about 45 min.

Table 1
Materials Used in the Present Study

Target Sentences
NP1 disambiguation
De agenten verhoren de adviseur van de politici die spreekt met een zachte stem. (The police interrogate the advisor of the politicians who speaks with a soft voice.)
NP2 disambiguation
De agenten verhoren de adviseur van de politici die spreken met een zachte stem. (The police interrogate the advisor of the politicians who speak with a soft voice.)
Contexts
Neutral
De gerechtelijke politie houdt zich bezig met een politiek schandaal. Enkele hooggeplaatsten worden opgepakt voor ondervraging. Aangezien de meeste ondervraagden wensen mee te werken, lijkt het onderzoek vlot te verlopen. (The judicial police are investigating a political scandal. Some highly placed people are taken in for questioning. Given the fact that most people are cooperating, the investigation runs very smoothly.)
NP1 biasing
De gerechtelijke politie houdt zich bezig met een politiek schandaal. Twee adviseurs die voor politici werken, worden opgepakt wegens fraude. Hoewel de ene adviseur in paniek is, houdt de andere zich rustig. (The judicial police are investigating a political scandal. Two advisors, working for politicians, are charged with fraud. Although one advisor is seized by panic, the other one remains calm.)
NP2 biasing
De gerechtelijke politie houdt zich bezig met een politiek schandaal. Een adviseur die voor enkele politici werkt, wordt opgepakt wegens fraude. Hoewel een groep van politici in paniek is, houden de anderen zich rustig. (The judicial police are investigating a political scandal. An advisor, working for politicians, is charged with fraud. Although some politcians are seized by panic, the other ones remain calm.)

Results

Due to the number difference between NP1 and NP2, all sentences were clearly disambiguated toward an NP1 or NP2 attachment by the continuation provided by the participants (there was only one instance of an ambiguous continuation in a total of 900 observations).

In the neutral context, 60% of the provided continuations contained an NP1 attachment, which confirms the Dutch preference for high attachment. When the same target sentences were preceded by an NP1-biasing context, the preference for high attachment was increased to 83%. Equally important, when an NP2 biasing context was presented before the target sentences, the high attachment preference was turned into a low attachment preference (39% of NP1 continuations). A repeated measures analysis of variance (ANOVA) was run on the percentage of NP1 continuations and revealed that the main effect of context was significant [$F_1(2,58) = 52.56, p < .001$; $F_2(2,58) = 46.90$, $p < .001$]. Planned comparisons also showed that the biasing contexts had a significant influence on the continuations of the target sentences [neutral vs. NP1 context: $F_1(1,29) = 48.56, p < .001$; $F_2(1,29) = 19.78, p < .001$; and neutral vs. NP2 context: $F_1(1,29) = 21.77, p < .001$; $F_2(1,29) = 35.77, p < .001$].

Discussion

The results of the completion study show that the contexts we created were strong enough to influence the final interpretation of the relative clause. In contrast to the contexts of Zagar et al. (1997), the presence of a late-closure-biasing context reversed the attachment preference from a 60% NP1 preference in the neutral condition to an NP2 preference (39% NP1 continuations). In addition, the early-closure-biasing context exaggerated the 60% NP1 preference to a 83% preference.

READING EXPERIMENT

The contexts of the completion study were now used in an on-line reading experiment to see whether they would influence initial sentence parsing, as predicted by discourse-sensitive theories, or whether their effect would be virtually absent, as reported by Zagar et al. (1997). To ensure that we were probing for discourse effects in the initial phase of sentence parsing, the sentences were disambiguated one word after the onset of the ambiguity. Moreover, the disambiguation was achieved grammatically (by number agreement) rather than semantically or pragmatically.

The underlying assumption in this mode of disambiguation is that people do pay attention to number agreement. In a number of studies featuring a wide variety of techniques (including sentence matching, naming, ERPs, self-paced reading, and eye tracking), it has indeed been shown that subject–verb number agreement is noticed in comprehension (Freedman & Forster, 1985; Nicol, Forster, & Veres, 1997; Osterhout & Mobley, 1995; Pearlmutter, Garnsey, & Bock, 1999; Sevald & Garnsey, 1995). Very recently, Pearlmutter et al. presented three self-paced reading experiments and concluded that subject–verb agreement is an early component of sentence comprehension. Furthermore, they argued for a comprehension system that processes agreement features at the moment they are encountered.

Method

Participants. The 30 participants were undergraduate students from Ghent University who were paid for their participation. All were native Dutch speakers and naive as to the purpose of this study. None had participated in the completion experiment.

Materials and Design. The referential contexts and target sentences of the completion experiment were used. Because there were two versions of each sentence, the experiment consisted of a 3 (contexts: neutral, early-closure biasing, late-closure biasing) × 2 (attachment: NP1 or NP2 agreement between the number of the verb and the number of one of the nouns) factorial design. Participants saw only one of the six possible versions according to a Latin square design. Each stimulus consisted of five lines of text with a maximum of 80 characters per line (so that 3 characters subtended a visual angle of 1°). The critical part of the target sentence was always displayed in the middle of the fourth line of text. The target sentences were followed by a final, neutral, sentence that was the same for all three contexts. In addition to the 30 experimental texts, the stimulus list contained 144 filler texts of the same length.

To ensure that processing differences as a function of NP1 or NP2 attachment could not be due to the fact that the relative clause was a more plausible modifier of one of the NPs, we asked 30 new participants from the same population to rate the plausibility of both continuations. In particular, we asked them to rate descriptions of situations on a scale from 1 (*totally implausible*) to 7 (*totally plausible*). One situation was described as "an advisor who speaks with a soft voice" (i.e., the NP1 modification of the sentence in Table 1). Another was described as "politicians who speak with a soft voice" (i.e., the NP2 modification of the sentence in Table 1). The 60 test descriptions were mixed in a pseudorandom order with 60 filler utterances, so that participants rated a total of 120 items and never saw both versions of the same sentence in close contiguity. On the basis of this study, we were able to conclude that both RC attachments were highly and equally plausible (NP1 = 5.68; NP2 = 5.72; both F_1 and $F_2 < 1$).[1]

Procedure. Participants were seated at a distance of 75 cm from a 17-in. display and were wearing an SMI Eyelink headband-mounted eye-tracking system. The sampling rate of the Eyelink system is 250 Hz and is based on infrared video-based tracking technology that happens simultaneously for both eyes. Although the system compensates for head position, this compensation is not accurate enough to allow single character resolution. Therefore, we asked the participants to put their chin on a cushion (in order to allow the small jaw movements needed for the "yes" or "no" answers to the occasional comprehension question) and to move as little as possible.

The experiment was divided into two sessions (87 stimulus texts per session), each of which started with a calibration and a validation procedure. In the calibration procedure the participants had to fixate nine calibration points that were presented randomly one at the time in the form of a 9-point grid. When the last calibration target had been presented, 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" on two consecutive calibrations the validation procedure was started. Running a validation immediately after the calibration enabled us to assess the accuracy of the system in predicting gaze position from pupil position. In the validation phase, the targets were presented in the same way as in the calibration procedure. When the participant 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 computed. 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°. When one or two eyes did not pass the test, new calibration and validation procedures were initiated. For 3 participants it was not possible to obtain a good grade of validation after considerable effort. These participants were replaced by 3 other participants.

After the validation procedure was completed, a random permutation of the stimulus texts (different for each participant) was presented in a self-paced manner. At the beginning of a trial the Eyelink system performed a calibration check (with a single fixation point in the center of the screen) and was adjusted in case the check was negative. The five lines of text started at the upper left corner of the screen and appeared every 3.75 cm, so that the critical fourth line of text was presented at the midline of the screen, where measurement accuracy was maximal. Participants were asked to read the sentences for meaning. To encourage them to do so, 26 of the 174 texts (15%) were followed by a question (posed by the experimenter), which had to be answered with "yes" or "no" (feedback was given). In order to prevent participants from explicitly memorizing the texts, the questions were kept relatively easy. None of the questions concerned the research topic (i.e., the attachment of the relative clause or the specific referents in the contexts).

Results

The overall accuracy of the answers to the questions was high (94%), meaning that the participants read the sentences for meaning. The eye-tracking data for the target sentences were analyzed in relation to seven regions that were defined as follows: (1) the beginning of the main clause; (2) NP1; (3) the prepositional phrase, consisting of the preposition "van (*of*)" and NP2; (4) the relative pronoun; (5) the disambiguating verb; (6) the next word; and (7) the remainder of the relative clause. Only the data of one eye (usually the right eye) were analyzed.

We first analyzed the cumulative region reading times (CRRT) because it has been shown that this variable has some advantages over first-pass reading times (FPRT) for the analysis of sentence parsing difficulties (e.g., Brysbaert & Mitchell, 1996; Liversedge, Paterson, & Pickering, 1998). The CRRT is defined as the sum of the fixations between the moment when the eyes first cross the front border of the region and the moment when they first cross the back border. This means that regressions originating from a particular region are added to the CRRT of that region (although they are not added to the FPRT). The reason for adding regressions to the reading time of a region is that processing difficulties usually manifest themselves in either of two ways: by prolonged reading of the disam-

Table 2
Mean Cumulative Region Reading Times (CRRTs, in Milliseconds) for Each of the Seven Regions and the Mean Sentence Reading Times (SRTs, in Milliseconds) as a Function of Referential Context and Attachment Type

	CRRT							
	1	2	3	4	5	6	7	SRT
Neutral								
NP1	549	401	507	77	338	78	823	2,773
NP2	539	374	437	82	349	114	951	2,847
NP1 biasing								
NP1	586	366	446	98	297	84	818	2,695
NP2	595	369	442	82	313	116	1,087	3,004
NP2 biasing								
NP1	609	417	428	66	296	126	776	2,718
NP2	622	360	439	60	282	149	943	2,854

biguating region or by rereading the previous ambiguous part of the sentence. Table 2 presents the CRRTs of the seven regions as a function of context (neutral, NP1 biasing, NP2 biasing) and attachment site (NP1 or NP2). The sentence reading time (SRT) is the sum of the CRRTs over the regions.

A 3 × 2 ANOVA with two repeated measures (context and attachment) on the total sentence reading times revealed a significant NP1-attachment bias [$F_1(1,29) = 8.48$, $p < .001$; $F_2(1,29) = 5.35$, $p < .05$]. There was no main effect of the preceding context [$F_1(2,58) < 1$; $F_2(2,58) < 1$], and most importantly there was no interaction between context and attachment site [$F_1(2,58) = 1.16$, $p < .32$; $F_2(2,58) = 1.35$, $p < .27$]. This means that sentences with an NP1 attachment were read faster than sentences with an NP2 attachment, even when they were preceded by a strong NP2-biasing context. A closer look at Table 2 shows that the NP1 bias is almost entirely due to the CRRTs of Regions 6 and 7—that is, the regions following the disambiguating verb. Although the effect starts to show up at Region 6 (a mean NP1 advantage of 30 msec), statistically the NP1 bias becomes reliable at Region 7 only [$F_1(1,29) = 12.99$, $p < .01$; $F_2(1,29) = 9.74$, $p < .01$]. Because Regions 6 and 7 together made up a phrase, we also ran analyses of the sum of the CRRTs of both regions and this separately for the three different contexts. This revealed that the NP1-attachment advantage was significant even in the NP2-biasing context [$F_1(1,29) = 7.11$, $p < .05$; $F_2(1,29) = 4.34$, $p < .05$]. The advantage was also robust for the NP1-biasing context [$F_1(1,29) = 7.49$, $p < .05$; $F_2(1,29) = 8.83$, $p < .01$] and for the neutral context over participants [$F_1(1,29) = 6.01$, $p < .05$], but not quite over stimuli [$F_2(1,29) = 3.27$, $p = .08$]. At Region 5 (the disambiguating verb) there was no effect of RC attachment (both Fs < 1), nor an interaction with context (both Fs < 1).

To examine whether the NP1-attachment advantage originated predominantly from longer first-pass readings or from more regressions, we analyzed both variables separately. FPRT is defined as the sum of the fixations between the first entrance of a region and the first exit, either to the left or to the right, provided that the region has been fixated during first-pass reading. Percentage regression is defined as the number of times the eyes leave a region to the left relative to the number of times this region has been looked at during first-pass reading.

Mean FPRTs as a function of context and attachment site are presented in Table 3. As can be seen, the clear and consistent preference for NP1 attachment disappeared in FPRTs. Separate 3 (context) × 2 (attachment) ANOVAs of Regions 5, 6, and 7 confirmed that there was no significant attachment bias (all Fs < 1), nor was there a significant interaction with context (all Fs < 2.42, all ps $> .10$). In contrast, there was a clear NP1-attachment preference in percentage regressions (Table 4). However, as with the CRRT data, the effect reached significance only in Region 7 [$F_1(1,29) = 12.27$, $p < .01$; $F_2(1,28) = 8.35$, $p < .01$]. There was no significant interaction between attachment site and preceding context in Region 5, 6, or 7 [Region 5: F_1 and $F_2 < 1$; Region 6: $F_1(2,20) = 1.61$, $p > .22$; $F_2 < 1$; Region 7: F_1 and $F_2 < 1$].

Although we had made sure that the RCs could be attached equally well to NP1 and NP2 (see the Method section), it might still be objected that the on-line NP1 preference was due to the fact that the semantics of the verbs in the RCs were more easily integrated within the discourse situations created by NP1 modifications than within the discourse situations created by NP2 modifications.[2] Applied to the example sentence in Table 1, it might be ob-

Table 3
Mean First-Pass Reading Times (FPRTs, in Milliseconds) of Each of the Seven Regions as a Function of the Referential Context and Attachment Type

	FPRT						
	1	2	3	4	5	6	7
Neutral							
NP1	545	299	404	230	295	220	450
NP2	542	329	401	228	297	214	437
NP1 biasing							
NP1	567	301	364	228	264	225	415
NP2	581	290	376	221	285	221	400
NP2 biasing							
NP1	599	322	361	212	293	216	439
NP2	628	309	365	191	274	256	453

Table 4
Mean Percentages of Regression for Each of the Seven Regions as a Function of Referential Context and Attachment Type

	1	2	3	4	5	6	7
Neutral							
NP1	2	19	12	9	12	8	47
NP2	1	13	5	20	9	27	61
NP1 biasing							
NP1	3	15	10	10	9	21	49
NP2	2	17	11	9	12	26	60
NP2 biasing							
NP1	3	18	9	12	9	22	43
NP2	1	14	11	5	10	28	54

jected that "speaking with a soft voice" is more relevant to somebody who is interrogated (i.e., the advisor) than to the people surrounding this person (i.e., the politicians). To check for this, we ran a second plausibility study in which we asked 28 new participants to rate the plausibility of the situations described by the full sentences. Thus, participants would be asked about the plausibility of the situation depicted in the sentence "The police interrogate the advisor of the politicians who speaks with a soft voice" and the plausibility of the situation depicted in the sentence "The police interrogate the advisor of the politicians who speak with a soft voice." Although it is possible that this type of plausibility ratings is influenced by the syntactic load of the different versions, we tried to limit this impact by explicitly asking the participants to imagine the situation depicted by the sentence and to determine for themselves how likely such a situation was. The results of the rating study (which in all other respects was similar to the rating study described in the Method section) revealed that the NP1 continuations indeed tended to be slightly but significantly more plausible than the NP2 continuations [5.40 vs. 4.97; $F_1(1,27) = 23.13$, $p < .001$; $F_2(1,29) = 10.35$, $p < .01$].[3] Therefore, we decided to run an additional analysis of covariance (ANCOVA) on the combined CRRTs of Regions 6 and 7 to find out whether the overall NP1-attachment preference could by explained by the semantics of the verb in the RC related to the discourse situation evoked by the main clause. As in the original F_2 ANOVA, we took context and attachment as repeated measures, but this time we included the difference between the NP1 and NP2 plausibility ratings as covariate. The ANCOVA returned the same results as the ANOVA: a significant main effect of attachment [$F(1,28) = 5.83$, $p < .05$], no effect of context ($F < 1$), and—most importantly—no interaction between attachment and context ($F < 1$). In addition, none of the interactions with the covariate were significant (all Fs < 1.97). Therefore, we conclude that the NP1-attachment preference in the reading data cannot be accounted for by differences in attachment plausibility.

Discussion

Referential contexts that in the completion study induced a shift of more than 40% in the number of NP1 attachments only slightly (and nonsignificantly) modulated the general advantage of reading sentences with NP1 disambiguation in the reading experiment. Adding the CRRTs of the last three regions together, NP1 disambiguated sentences were read 317 msec faster than were NP2 disambiguated sentences after an NP1-biasing context, but they were also read 176 msec faster after an NP2-biasing context. This replicates a previous study by Zagar et al. (1997) and seems hard to reconcile with strong discourse-sensitive accounts of sentence parsing.

As a matter of fact, the present pattern of results very closely resembles that observed by Mitchell et al. (1992), which provided support for syntax-first models. Mitchell et al. examined the complement/relative clause ambiguity discussed in the introduction. Like us, they observed slightly less garden-pathing for the relative clause structures relative to the complement clause structures in a two-referent context relative to a one-referent context (which they attributed to integration effects after the initial syntactic parsing stage). More importantly, they were able to demonstrate an initial preference for the complement clause continuation, even in a two-referent context, that supposedly favored the relative clause interpretation. Similarly, we observed a significant NP1-attachment advantage in a referential context that favored an NP2 attachment.

GENERAL DISCUSSION

In this study the effect of referential context was investigated on the processing of relative clauses that can be attached to two possible noun phrases ("Someone shot **the servant** of **the actress** *who was on the balcony*"). Several explanations have been proposed for the parsing of this particular sentence construction. Some of the proposed accounts, such as the construal theory (Frazier & Clifton, 1996) and the referential account (Hemforth, Konieczny, & Scheepers, 2000; Konieczny & Hemforth, 1996), explain the attachment in terms of detailed grammatical devices. Other proposals assume that the syntactic parser tunes to variations in the language to which it is exposed (e.g., Brysbaert & Mitchell, 1996; Gibson, Pearlmutter, Canseco-Gonzalez, & Hickok, 1996; MacDonald, Pearlmutter, & Seidenberg, 1994; Mitchell et al., 1995; see Mitchell & Brysbaert, 1998, for a more thorough review of the different proposals). However, thus far none of the theoretical frameworks has explained the processing of relative clause attachment in terms of referential discourse factors. Zagar et al. (1997) provided some evidence for this state of affairs.

In view of recent claims about early discourse influence on syntactic processing of a very similar structure (van Berkum et al., 1999), we set out to more thoroughly test the power of the referential discourse context to affect the initial processing of relative clause attachment. In the contexts we used, one of the noun phrases did not have a unique referent, which is a strong trigger for modifier attachment, at least as far as the ultimate interpretation is concerned (e.g., Altmann et al., 1994; Altmann & Steedman, 1988;

Crain & Steedman, 1985). Our completion experiment confirmed this effect: The preferred continuation participants gave to equivalents of the lead sentence "Someone shot the servant of the actress who ..." was strongly determined by the preceding discourse (i.e., whether either the servant or the actress lacked a unique referent). The NP1-attachment preference, previously found for Dutch in the absence of referential discourse and replicated here, could be reversed into an NP2-attachment preference if the second noun phrase had more than one possible referent. However, more crucial to the issue of on-line discourse effects on sentence parsing, the reading times were only slightly influenced by the number of referents to the first or the second noun phrase. In all constructions, there was a significant NP1-attachment preference independent of the preceding context. These data confirm the conclusion of Zagar et al. (1997).

At present, it is not clear why our findings differ from those of van Berkum et al. (1999) and other studies on the effect of referential context (e.g., Altmann et al. 1992; Altmann et al., 1994). One explanation could be that this discrepancy is due to the type of syntactic ambiguity under scrutiny. For some reason, the referential context might affect the complement/relative clause ambiguity more than the ambiguity in relative clause attachment. Spivey-Knowlton and Tanenhaus (1994) suggested that it is harder to find referential context effects in highly skewed syntactic ambiguities than in more balanced ones. If a given ambiguity has one alternative that is much less available than the other, contextual influence will need a lot of time to override the dominant alternative. If the ambiguity is more balanced, however, contextual information will be able to increase the availability of either alternative quite rapidly. So, maybe the relative clause attachment ambiguity is more biased than the complement/relative clause ambiguity, resulting in context effects on the latter but not on the former. However, in the same paper, Spivey-Knowlton and Tanenhaus (1994) stated that the complement/relative clause ambiguity had an unusually high degree of asymmetry in initial availability of the two syntactic alternatives. Moreover, evidence is accumulating that for relative clause attachment ambiguity there is no strongly preferred structural analysis; rather, the attachment can be easily influenced by nonstructural constraints (e.g., Desmet, Brysbaert, & De Baecke, in press; Frazier & Clifton, 1997; Hemforth et al., 2000). If anything, Spivey-Knowlton and Tanenhaus's (1994) account would predict exactly the opposite pattern: stronger context effects in the relative clause ambiguity than in the complement/relative clause ambiguity.

An alternative explanation could focus on the methodological differences between the present and former studies. For instance, the fact that van Berkum et al. (1999) did not include filler items in their materials and presented all conditions of each item to every participant may have induced the use of situation-specific coping strategies (see Brysbaert & Mitchell, 2000, for a more detailed discussion). Moreover, the eye-tracking experiment differed from most of the previous studies in that (1) the ambiguation was resolved at the earliest possible point (the first word following the ambiguous relative pronoun), (2) the disambiguation was based on a grammatical feature (number agreement), and (3) the texts were presented as a whole and not sliced in pieces as in self-paced reading or ERP studies.

Further research will have to clarify whether context effects are still present for other ambiguities when the effects are submitted to the same methodological standards as in the present study. Only then will it be worthwhile to start theorizing about why the influence of referential context may vary with different ambiguities.

REFERENCES

Altmann, G. T. M. (1988). Ambiguity, parsing strategies, and computational models. *Language & Cognitive Processes*, **3**, 73-97.

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

Altmann, G. T. M., Garnham, A., & Henstra, J.-A. (1994). Effects of syntax in human sentence parsing: Evidence against a structure-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 209-216.

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

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

Brysbaert, M., & Mitchell, D. C. (2000). The failure to use gender information in parsing: A comment on van Berkum, Brown, and Hagoort (1999). *Journal of Psycholinguistic Research*, **29**, 453-466.

Carreiras, M., & Clifton, C., Jr. (1999). Another word on parsing relative clauses: Eye-tracking evidence from Spanish and English. *Memory & Cognition*, **27**, 826-833.

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

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.

Desmet, T., Brysbaert, M., & De Baecke, C. (in press). The correspondence between sentence production and corpus frequencies in modifier attachment. *Quarterly Journal of Experimental Psychology*.

De Vincenzi, M., & Job, R. (1995). An investigation of late closure: The role of syntax, thematic structure, and pragmatics in initial and final interpretation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1303-1321.

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

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559-585). Hillsdale, NJ: Erlbaum.

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

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

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.

Freedman, S. E., & Forster, K. I. (1985). The psychological status of overgenerated sentences. *Cognition*, **19**, 101-131.

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

Gibson, E., Pearlmutter, N. J., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, **59**, 23-59.

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: Cambridge University Press.

KONIECZNY, L., & HEMFORTH, B. (1996, June). *A visibility-based model of human sentence processing*. Paper presented at the NIAS workshop on Computational Models of Human Syntactic Processing, Wassenaar, The Netherlands.

LIVERSEDGE, S. P., PATERSON, K. B., & PICKERING, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 55-75). Oxford: Elsevier.

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

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: 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., & CORLEY, M. M. B. (1994). Immediate biases in parsing—Discourse effects or experimental artifacts. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 217-222.

MITCHELL, D. C., CORLEY, M. M. B., & GARNHAM, A. (1992). Effects of context in human sentence parsing: Evidence against a discourse-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 69-88.

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.

NI, W., CRAIN, S., & SCHANKWEILER, D. (1996). Sidestepping garden paths: Assessing the contribution of syntax, semantics and plausibility in resolving ambiguities. *Language & Cognitive Processes*, **11**, 282-334.

NICOL, J. L., FORSTER, K. I., & VERES, C. (1997). Subject–verb agreement processes in comprehension. *Journal of Memory & Language*, **36**, 569-587.

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

PEARLMUTTER, N. J., GARNSEY, S. M., & BOCK, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory & Language*, **41**, 427-456.

RAYNER, K., & POLLATSEK, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice-Hall.

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.

SPIVEY-KNOWLTON, M. J., & TANENHAUS, M. K. (1994). Referential context and syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 415-439). Hillsdale, NJ: Erlbaum.

SPIVEY-KNOWLTON, M. J., & TANENHAUS, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1521-1543.

VAN BERKUM, J. J. A., BROWN, C. M., & HAGOORT, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory & Language*, **41**, 147-182.

VAN BERKUM, J. J. A., HAGOORT, P., & BROWN, C. M. (2000). The use of referential context and grammatical gender in parsing: A reply to Brysbaert and Mitchell (2000). *Journal of Psycholinguistic Research*, **29**, 467-481.

ZAGAR, D., PYNTE, J., & RATIVEAU, S. (1997). Evidence for early closure attachment on first-pass reading times in French. *Quarterly Journal of Experimental Psychology*, **50A**, 421-438.

NOTES

1. For comparison purposes, the plausibilities of the continuations of the example sentence in Table 1 were 5.3 for the NP1 modification and 4.8 for the NP2 modification.

2. We thank an anonymous reviewer for pointing out this possibility.

3. The figures for the example sentence in Table 1 were 6.0 for the NP1 continuation and 4.6 for the NP2 continuation.

(Manuscript received October 6, 2000;
revision accepted for publication August 15, 2001.)