



On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking

Arnout W. Koornneef^{a,*}, Jos J.A. Van Berkum^{b,c}

^a Utrecht Institute of Linguistics OTS, Utrecht University, Trans 10, 3512 JK Utrecht, The Netherlands

^b Department of Psychology, University of Amsterdam, The Netherlands

^c FC Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands

Received 3 October 2005; revision received 6 December 2005

Available online 31 January 2006

Abstract

In two experiments, we examined the recent claim (Stewart, Pickering, & Sanford, 2000) that verb-based implicit causality information is used during sentence-final clausal integration only. We did so by looking for mid-sentence reading delays caused by pronouns that are inconsistent with the bias of a preceding implicit causality verb (e.g., “*David praised Linda because he...*”). In a self-paced reading task, such pronouns immediately slowed down reading, at the two words immediately following the pronoun. In eye tracking, bias-inconsistent pronouns also immediately perturbed the reading process, as indexed by significant delays in various first pass measures at and shortly after the critical pronoun. Hence, readers can recruit verb-based implicit causality information in the service of comprehension rapidly enough to impact on the interpretation of a pronoun early in the subordinate clause. We take our results to suggest that implicit causality is used proactively, allowing readers to focus on, and perhaps even predict, who or what will be talked about next.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Language comprehension; Interpersonal verbs; Implicit causality; Pronoun resolution; Immediacy; Prediction

Introduction

If you praise somebody, you will typically do so because of *his or her* behavior, not yours. If you apologize to somebody, however, the most likely relevant cause is *your* behavior, not theirs. These simple probabilistic asymmetries express part of our knowledge about

when certain interpersonal transactions are appropriate, and as such help us deal with—and make sense of—the social world in which we live.

Perhaps not surprisingly, we also recruit this knowledge when particular interpersonal transactions are being described in language (Garvey & Caramazza, 1974). When asked to complete a sentence fragment such as “*David praised Linda because...*”, for example, readers or listeners will be inclined to continue the *because*-clause with something about Linda, as in example (1) below.

* Corresponding author. Fax: +31 30 253 6000.

E-mail addresses: arnout.koornneef@let.uu.nl (A.W. Koornneef), berkum@psy.uva.nl (J.J.A. Van Berkum).

(1) David praised Linda because she had done well.

Alternatively, after “*David apologized to Linda because...*”, people tend to continue with something about David. In the above constructions, interpersonal verbs like *praise* and *apologize* thus supply information about whose behavior or state is the more likely immediate cause of the event at hand. Because it is conveyed implicitly as part of the meaning of the verb, this probabilistic cue is usually referred to as *implicit causality*.

Early research on the comprehension of sentences with implicit causality verbs (e.g., Au, 1986; Brown & Fish, 1983; Caramazza, Grober, Garvey, & Yates, 1977; Garvey & Caramazza, 1974; Garvey, Caramazza, & Yates, 1975) has shown that readers make use of this cue to arrive at the correct sentence interpretation. One consistent finding, for instance, is that if the causality information explicitly provided in the remainder of the sentence goes *against* the bias of the verb at hand, as in example (2) below, people need more time to read the entire sentence (Caramazza et al., 1977). Note that, as illustrated in (2), the probabilistic bias of an implicit causality verb can be negated without rendering the sentence ungrammatical or incoherent. The fact that readers slow down on a bias-inconsistent sentence is therefore not the result of an overt anomaly. Instead, it reflects something more subtle about the way we use various sources of information in everyday language comprehension.

(2) David praised Linda because he was very proud.

The early studies on implicit causality unequivocally revealed that people are sensitive to the implicit causality information associated with particular interpersonal verbs. In the field of social cognition, these findings have led to a closer examination of the semantic features by means of which implicit causality verbs represent information about interpersonal exchanges, and of how these features might relate to models of social cognition (e.g., Brown & Fish, 1983; Semin & Marsman, 1994; see Rudolph & Försterling, 1997, for review). For psycholinguists, however, an important unresolved issue is how and when verb-based implicit causality information is brought to bear on the actual processing of an unfolding sentence. Two contrasting accounts have been proposed, which we will denote as *clausal integration* and *immediate focusing*, respectively. According to the clausal integration account (e.g., Garnham, Traxler, Oakhill, & Gernsbacher, 1996; Stewart et al., 2000), verb-based implicit causality information is brought to bear on comprehension relatively late in the unfolding sentence, at or towards the end of the subordinate clause. More specifically, implicit causality effects such as observed by Caramazza et al. are believed to emerge when people have read the main clause as well as the bulk of the sub-

ordinate clause, and subsequently combine the causal information provided by both clauses into a single representation. Such ‘retroactive’ integration would be more difficult if, as in (2), the causality information made explicit in the subordinate clause is inconsistent with the main verb’s implicit causality bias.

Proponents of the immediate focusing account (e.g., Greene & McKoon, 1995; Long & De Ley, 2000; McKoon, Greene, & Ratcliff, 1993) have instead suggested that implicit causality information can be brought to bear on comprehension much more rapidly, at the beginning of the subordinate *because*-clause, and perhaps even already at the verb itself. In this account, implicit causality is assumed to immediately bring one of the persons referred to into focus, at the expense of the other. Because readers and listeners prefer to relate a personal pronoun to the most focused antecedent (see Arnold, 1998, for review), such an immediate modulation of focus would in turn affect the ease with which a subsequent pronoun is processed. In this account, the bias-inconsistent sentence (2) would take more time to read because “*David praised Linda because*” immediately brings Linda into focus, thereby making it more difficult to resolve the subsequent pronoun *he* in (2) than the pronoun *she* in (1). Of course, this early conflict at the pronoun may have additional processing consequences further downstream in the sentence. Thus, the immediate focusing account is not incompatible with an additional late processing delay at or towards the end of the bias-inconsistent subordinate clause. What distinguishes this account from the ‘retroactive’ clausal integration alternative is that only the former allows for much earlier, ‘proactive’ effects of implicit causality, right after the critical verb, its arguments, and a connective like *because*.¹

To keep track of the use of implicit causality information, as a sentence is unfolding, recent studies have exploited a so-called probe (or probe verification) task. In this task, the participant is asked to decide as rapidly as possible whether or not a particular probe word (e.g., “Linda”) has appeared in the sentence fragment presented so far. By presenting the probe at various critical positions within the unfolding sentence, this method can be used to determine at what point(s) in the sentence a particular discourse referent is in focus. Using this probe task, Garnham et al. (1996) obtained support for the clausal integration account: an effect of implicit causality on late probes presented at the end of a sentence, but no such effect on early probes presented right after “*because he*.” However, the results of three other studies with the probe task (Greene & McKoon, 1995; Long & De Ley, 2000; McDonald & MacWhinney,

¹ The terms ‘proactive’ and ‘retroactive’ are taken from Garnham’s (2001) review of implicit causality research.

1995) suggest an earlier use of implicit causality, in line with immediate focusing. Thus, findings obtained with the probe task are mixed, with some support for either account.

The advantage of a probe task is that it more or less directly queries the accessibility of the pronoun's antecedents (e.g., Linda), and thus seems to provide a straightforward measure of focus and any verb-based modulations thereof. However, there is another side to this coin, for such direct querying may also invite readers to memorize the content words of a sentence (see Gordon, Hendrick, & Foster, 2000, for evidence that this can happen), and to invest an unrepresentative amount of attention in assigning antecedents to pronouns. More generally, people do not verify probes when they read or listen to linguistic input. Although neither of these concerns necessarily invalidates the results obtained with a probe task, they do point to the need for additional, less obtrusive measures.

To avoid some of the potential problems associated with the probe task, Stewart et al. (2000) tracked the use of implicit causality information during sentence comprehension by means of a self-paced reading task. In the most relevant experiment (Experiment 4), people read sentences much like our examples (1) and (2) in two large fragments, one up to and including the critical pronoun and the other fragment containing the remainder of the sentence. Although the average reading times on bias-inconsistent first fragments like "David praised Linda because he" were 50 ms larger than reading times on bias-consistent fragments like "David praised Linda because she", this early difference was not significant. However, inconsistency with verb bias did significantly delay reading of the subsequent second fragments, by some 206 ms. These and related reading time findings have led Stewart et al. to strongly embrace the clausal integration account, in which implicit causality only becomes relevant "at the point where the interpretation of the two clauses are integrated into a single representation for the sentence as a whole" (p. 424). In a recent and thorough review of implicit causality research, Garnham (2001) also arrives at the conclusion that verb-based implicit causality information is used towards the end of the sentence, during the 'retroactive' integration of main and subordinate clause.

If this conclusion is correct, however, it would set the influence of implicit causality information on sentence comprehension apart from that of a wide range of other sources of information that have been shown to affect the comprehension process very rapidly. For example, we know from ERP research (see Kutas & Van Petten, 1994, for review) that sentences such as "He buttered the memo before he took a bite" elicit an N400 effect right at the anomalous noun, which betrays the immedi-

ate use of verb-based semantics. Verb-specific lexical information is also rapidly accessed and used in sentence-medial syntactic ambiguity resolution (Osterhout & Holcomb, 1992; Trueswell, Tanenhaus, & Kello, 1993). Furthermore, and particularly relevant to our present concerns, the semantics of a verb can immediately restrict the domain of reference within a particular context (Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003; Nieuwland & Van Berkum, in press). To the extent that it provides relevant and sufficiently strong (i.e., biasing) information, why would the particular semantic features of a verb that express implicit causality be an exception to the rule? Note that ERP studies on story comprehension suggest that even *wider discourse* is immediately and incrementally brought to bear on the syntactic, semantic, and referential analysis of an unfolding sentence (e.g., St. George, Mannes, & Hoffman, 1994; Van Berkum, Brown, & Hagoort, 1999a; Van Berkum, Zwitserlood, Brown, & Hagoort, 2003; see Van Berkum, 2004; Van Berkum, in preparation, for review). If probabilistic global discourse-level information can affect the analysis of an unfolding sentence at every relevant word coming in, within only a few hundred milliseconds from word onset, why would the impact of a relevant local causality cue need to wait for the sentence to terminate?

A principled delay in the use of implicit causality might make sense in the context of models that (possibly inspired by the thorough feature-space analyses of social psychologists; e.g., Rudolph & Försterling, 1997) grant a special status to the causality-related semantic features of verbs like *praise* and *disappoint*. It can also make sense if one assumes that the processing of subordinate clauses is only *locally* incremental, and that such clauses are related to their main clause context only when fully processed (see Millis & Just, 1994, for exactly this claim). However, if, as we were inclined to believe, implicit causality is 'just another bit of verb meaning' (and subordinate clauses are not initially processed in a semantic vacuum), a principled substantial delay in using this information would be very hard to understand. In the two studies reported below, we therefore set out to test the hypothesis that implicit causality information is only used during sentence-final clausal integration. We did so by looking for mid-sentence reading delays caused by pronouns that are inconsistent with the bias of a preceding implicit causality verb (cf. example 2). In Experiment 1, we kept track of the impact of implicit causality by means of word-by-word self-paced reading. In Experiment 2, we used eye tracking, as participants freely read through the same materials. We predicted that verb-based implicit causality information would become available rapidly enough to bear on the interpretation of a pronoun encountered very early in the sentence.

Experiment 1

As discussed before, the self-paced reading findings of Stewart et al. (2000) were taken to support the clausal integration account. However, in the Stewart et al. experiments the critical sentences were presented in two relatively large fragments, with the first part consisting of the main clause plus the connective and the anaphor (either a pronoun or proper name), and with the second part consisting of the remainder of the subordinate clause. Thus, a sentence like “*David praised Linda because he was very happy with the results*” would be presented as [*David praised Linda because he*] [*was very happy with the results*]. With only two such ‘sample points’ per sentence, the temporal resolution of the Stewart et al. design is rather low. An important consequence is that the observed average reading times are very large, making it difficult to pick up on small reading time effects. Note that in the most relevant experiment (Experiment 4), fragments containing bias-inconsistent pronouns did delay reading by 50 ms. However, the average total reading time on those first fragments exceeded 2000 ms. The variance associated with such large reading times may well have swamped a small pronoun effect.

Furthermore, because of the specific sentence partitioning used by Stewart et al., some of the early processing consequences elicited by an inconsistent pronoun may actually have ended up in reading times at the first few words in the *second* fragment, and as such erroneously have been taken as evidence for the clausal integration account. With only three button presses per trial, the participants in the Stewart et al. self-paced reading experiments will probably have suffered somewhat less from the spill-over of effects caused by a task-induced ‘button-press rhythm.’ However, spill-over may also have a much deeper cause. For instance, ERP research in which a singular gender-marked pronoun *he* had two equally salient male antecedents, or only female antecedents, has revealed that the processing consequences of such referential problems can last until over a second after pronoun onset, i.e., well beyond the average time that people look at such words (Van Berkum, Zwitserlood, Bastiaansen, Brown, & Hagoort, 2004). The implication is that the processing consequences of a bias-inconsistent pronoun in the Stewart et al. experiment may well have delayed reading of the first few words in the second fragment. In all, there is reason to believe that the large fragments used in the Stewart et al. self-paced reading experiments may have caused them to miss small effects on the pronoun itself, and confound true sentence-final clausal integration effects with earlier effects spilled over from the pronoun, making it impossible to determine whether implicit causality information enters the comprehension process early or late.

To solve the above problems and obtain the temporal resolution needed to detect potentially small sentence-medial effects, we used *word-by-word* self-paced reading in our experiment. As illustrated by the critical sentences (3a) and (3b) below, the pronoun was held constant across condition. Instead of contrasting the Dutch equivalents of *he* and *she* (as in examples 1 and 2), we manipulated whether or not the fixed pronoun *he* was consistent with the verb’s bias by swapping the argument position of the man and the woman involved. We avoided the Dutch equivalent of *she* because it is ambiguous between a singular and a plural third-person pronoun. Furthermore, using the same pronoun across conditions allowed us to examine the claim (Garnham, 2001; Oakhill, Garnham, Reynolds, & Wilshire, 1998) that implicit causality effects observed in language comprehension would be artifacts of ‘low-level’ (e.g., word frequency) differences between the respective critical fragments used in bias-consistent and -inconsistent items.

(3a) *verb bias towards the 2nd NP, bias-consistent pronoun*

Linda praised David because he had been able to complete the difficult assignment with very little help.

(3b) *verb bias towards the 2nd NP, bias-inconsistent pronoun*

David praised Linda because he had been able to complete the difficult assignment with her help only.

As can be seen in examples (3a) and (3b), the end of the sentence was always consistent with the interpretation enforced by the pronoun, and a coherent interpretation was therefore always available. However, to ensure that any effects spilling over from a bias-inconsistent pronoun into the reading times of the immediately subsequent words would also not be confounded with the effects of the different sentence continuations that were required to achieve overall coherence, at least five words after the pronoun were held constant across conditions.

The critical issue is whether, relative to a consistent pronoun in sentences like (3a), a bias-inconsistent pronoun in sentences like (3b) causes a delay in self-paced reading at or shortly after the offending pronoun. Because the clausal integration account restricts the use of implicit causality information to clause-integration processes occurring at the end of the subordinate clause, this account does not predict an inconsistency effect at or shortly after the pronoun. According to immediate focusing account, however, fragments like “*X praised Y because*” immediately bring *Y* into focus, which should make it more difficult to resolve the subsequent pronoun *he* in (3b) than in (3a). We know from other research with on-line tasks (e.g., Boland et al., 1998; Arnold, Eisenband, Brown-Schmidt, & Trueswell,

2000; Van Berkum et al., 2004) that gender-dependent problems with pronoun resolution begin to emerge within only a few hundred milliseconds after pronoun onset. The immediate focus account therefore predicts that an inconsistency effect should appear at the pronoun, or, due to spill-over, rapidly thereafter.

Method

Participants

Participants were 24 undergraduate psychology students (21 female, mean age 21, range 18–33 years) who received course credit or money for their participation. In this and the following experiment, participants were native speakers of Dutch, without a diagnosed reading or learning disability.

Materials

Based on translations of English verbs used in previous research, 116 Dutch verbs were selected, of which 57 were expected to have a strong implicit causality bias towards the second noun phrase (henceforth NP2, e.g., the verb *prijzen*, to praise) and 59 a strong bias towards the first noun phrase (NP1, e.g., the verb *teleurstellen*, to disappoint). In a written sentence completion pretest, 45 native speakers of Dutch (23 female, mean age 22, range 18–61 years) were asked to complete the Dutch equivalent of fragments like “*John disappointed Paul because he_*”, and to subsequently denote the intended male antecedent of the pronoun *he* by encircling the name. If the participant marked NP1 the trial was scored as 1, and if he or she marked NP2 the trial was scored as 2. Calculating the mean score for every verb provides information about the direction of the verb’s bias as well as its strength (indexed via the consistency of the bias across participants), with a strong NP1-bias giving a score close to 1, a strong NP2-bias giving a score close to 2, and the absence of a bias giving a score around 1.5. Twenty verbs with a strong NP1-bias and 20 verbs with a strong NP2-bias were selected for Experiment 1 (mean bias NP1-verbs: 1.14, NP2: 1.94; Table 1 contains the relevant verbs and their approximate English translation).

Following McKoon et al. (1993), we constructed two different three-sentence stories for every verb for the main experiment, using scenarios that were expected to be of interest to the average participant. An example that involves an NP1 verb is given in (4), together with an approximate translation. In the first sentence a situation was sketched in which a man and a woman were introduced by name. For half of the texts, the first-mentioned character was the man, and for the other half, the woman (balanced across the NP1 and NP2 stories). In the second sentence a pronominal (usually *they*) was used to keep both characters in focus to an equal extent. The main clause of the third sentence contained the crit-

Table 1

Implicit causality verbs used in Experiment 1 and 2 with their bias

Verb	Bias
<i>NP1-biased verbs</i>	
fascineren (<i>fascinate</i>)	1.00
excuses aanbieden (<i>apologize</i>)	1.00
bekennen aan (<i>confess</i>)	1.04
teleurstellen (<i>disappoint</i>)	1.04
vervelen (<i>bore</i>)	1.05
hinderen (<i>be in the way</i>)	1.09
oplichten (<i>swindle</i>)	1.13
kwellen (<i>hurt</i>)	1.14
storen (<i>disturb</i>)	1.14
smeken (<i>beg</i>)	1.17
bellen (<i>call</i>)	1.18
ergeren (<i>annoy</i>)	1.18
verbazen (<i>amazed</i>)	1.18
verontrusten (<i>worry</i>)	1.18
winnen van (<i>win</i>)	1.18
misleiden (<i>mislead</i>)	1.22
irriteren (<i>irritate</i>)	1.22
liegen tegen (<i>lie to</i>)	1.22
inspireren (<i>inspire</i>)	1.23
intimideren (<i>intimidate</i>)	1.23
<i>NP2-biased verbs</i>	
minachten (<i>hold in contempt</i>)	2.00
benijden (<i>envy</i>)	2.00
bewonderen (<i>admire</i>)	2.00
vrezen (<i>fear</i>)	2.00
waarderen (<i>appreciate</i>)	2.00
bekritisieren (<i>criticise</i>)	2.00
prijzen (<i>praise</i>)	1.96
complimenteren (<i>compliment on something</i>)	1.96
haten (<i>hate</i>)	1.96
ontslaan (<i>fire</i>)	1.96
respecteren (<i>respect</i>)	1.95
straffen (<i>punish</i>)	1.95
feliciteren (<i>congratulate</i>)	1.95
verantwoordelijk stellen (<i>hold responsible</i>)	1.91
aanklagen (<i>press charges against</i>)	1.91
troosten (<i>comfort</i>)	1.91
verafschuwen (<i>loathe</i>)	1.86
aanbidden (<i>adore</i>)	1.86
houden van (<i>love</i>)	1.86
bedanken (<i>thank</i>)	1.82

ical verb and felicitously repeated the names of the two characters. The subordinate clause contained the critical pronoun *he* and was always adjoined to the main clause by the connective *because*.

(4a) *NP1-biased verb, bias-consistent pronoun*

David en Linda reden allebei behoorlijk hard. Bij een druk kruispunt botsten zij met hun auto’s stevig op elkaar. David bood zijn excuses aan Linda aan omdat hij volgens de getuigen van het ongeluk alle schuld had.

(David and Linda were both driving pretty fast. At a busy intersection they crashed hard into each other. David apologized to Linda because he according to the witnesses was the one to blame.)

(4b) NP1-biased verb, bias-inconsistent pronoun

David en Linda reden allebei behoorlijk hard. Bij een druk kruispunt botsten zij met hun auto's stevig op elkaar. Linda bood haar excuses aan David aan omdat hij volgens de getuigen van het ongeluk geen schuld had.

(David and Linda were both driving pretty fast. At a busy intersection they crashed hard into each other. Linda apologized to David because he according to the witnesses was not the one to blame.)

The critical manipulation was whether or not the pronoun *he* was consistent with the verb's implicit causality bias. For NP1-biased verbs, *he* is consistent if the male character occupies the NP1-position of the main clause, as in (4a), but inconsistent if the female character occupies that position, as in (4b). For NP2-biased verbs, the mapping is reversed, as in example (3) given before. To accommodate for spill-over, at least five words after *he* were held constant across conditions. After these five words the consistent and inconsistent versions diverged, and ended with explicit causal information that made the story coherent as a whole (see Appendix A for all Dutch originals).

To rule out any accidental biases in the constructed materials, incomplete versions of our stories were presented to 20 native speakers of Dutch (18 female, mean age 24, range 19–41 years) in a second paper-and-pencil completion test. The participants (none of which had participated in the sentence pretest) read our stories in the form illustrated by “David and Linda were both driving pretty fast. At a busy intersection they crashed hard into each other. ___ apologized to ___ because ___”, and they were asked to use the names in the story to fill in the blanks of the main clause, and to subsequently finish the story as they liked. There was no preference for a particular order of mention of the male or female character in the main clause of the critical sentence. In stories with an NP1-biased verb, participants used the man-verb-woman order—the order used in the bias-consistent version of these stories in the main experiments—about half of the time (53%). Similarly, in stories with an NP2-biased verb, participants used the woman-verb-man order—the order used in the bias-consistent version of these stories in the main experiments—also about half of the time (51%). This suggests that the specific way in which the two genders had been paired to the two verb arguments in our stories was deemed equally plausible for the consistent and inconsistent versions. In addition, the subsequent completion of subordinate clauses revealed that the wider story context had essentially not changed the bias of our verbs. Across

the completions in which participants had used either *he* or *she* immediately after the connective *because* (95% of all completions), the NP1-verb stories showed a strong bias towards the subject of the main clause (mean bias 1.10) and the NP2-verb stories a strong bias towards the object (mean bias 1.90).

The stimuli were divided into two lists, with each list containing 20 bias-consistent and 20 bias-inconsistent stories, and with only one version of each story in a particular list. To avoid strategic processing, we used as many NP1-verbs as NP2-verbs (10 each) within each condition in each list. Forty stories of an unrelated experiment were included as fillers (see Van Berkum et al., 2005b, Experiment 3). One pseudo-randomization was used for both lists. The original randomization order was used for one half of the participants, the reversed order for the other half. A set of 40 2-choice questions was included to encourage discourse comprehension. Half of the questions followed the experimental items and half followed the filler items. Furthermore, half of the correct answers appeared on the left of the screen and half on the right. The questions never directly probed the referent of the pronoun.

Procedure

The stories were presented on a fast LCD screen (Iiyama TXA 3834 MT) and responses were collected with response boxes integrated in the armrests of the chair. The stories were presented in a standard non-cumulative moving-window self-paced reading paradigm, using a non-proportional Courier 14p font. Subjects read through each story word by word, with each button press disclosing the next word while replacing all other letters in the story by hyphens. As they pressed their way through a story, subjects could see its overall sentential and formatting layout (including punctuation), as well as the position of the currently visible word therein. To prevent ‘edge effects’ in reading times, the critical region, which included the pronoun and five subsequent words, was always separated from the left and right paragraph edges by at least one word. Subjects were asked to process each story for comprehension, and to adapt their speed to this. Participants progressed through a text by pressing a button with the index finger of their dominant hand.

Each session started with a written instruction. The actual self-paced reading experiment consisted of four blocks. Block one was a practice block in which 10 stories were presented, five followed by a question. This familiarized the participant with the procedure and in addition gave the experimenter the opportunity to monitor the reading speed of the participants. Participants who read extremely slowly were encouraged to speed up. The participant had an obligatory one-minute break between the different experimental blocks. Each experimental block started with two practice stories. After

completion of the self-paced reading experiment, participants were administered a Dutch version of the reading span task (Daneman & Carpenter, 1980). A full session was completed within 60 min, with an average time-on-task of 40 min.

Analysis

Prior to all analyses, we evaluated the comprehension question performance of our participants. All participants scored above 75% correct (mean score 94%). Reading times more than 2 standard deviations from both the participant's mean and the item's mean in a particular condition were treated as missing data (2.1%). We report means and statistical analyses for the factor Consistency for the critical pronoun, each of the five words following the pronoun (spill-over region), and each of the four words preceding it (pre-critical region). We only discuss effects that are significant by subjects (F_1) and by items (F_2), and we report the associated $\min F'$ values in tables (Clark, 1973; Raaijmakers, Schrijnemakers, & Gremmen, 1999).

Because implicit causality results are customarily reported for NP1- and NP2-verbs separately, our analyses of variance will also take direction of the bias into account (Verb Bias). However, note that the comparison of inconsistency effects elicited by NP1 verbs and NP2 verbs is confounded with the effects of distance between anaphor and antecedent, of first mention, and of the antecedent's structural position (see Garnham, 2001, for an overview of the relevance of these factors). For this reason, we focus our discussion and figures on the main effects of Consistency.

Results and discussion

As can be seen from the mean reading times (collapsed over Verb Bias) in Fig. 1 and the associated F tests displayed in Table 2, words in the pre-critical region were read equally fast across Consistency condition. However, readers began to slow down right at the bias-inconsistent pronoun, with a significant main effect of Consistency emerging at the first two words after this pronoun. This effect was not significantly modulated by whether the verb was biased towards NP1 or NP2.²

² We also did not find any significant modulations of the inconsistency effect as a function of whether the participant's reading span was above or below the median score for this sample (span = 5.08 vs. 2.67 respectively, $SD = 1.00$ vs. .44). We note that, because our undergraduate participants were all relatively highly skilled readers, the absence of significant reading span effects in our study is not necessarily incompatible with the skill-dependent inconsistency effects reported by Long and De Ley (2000), and otherwise refrain from further discussion.

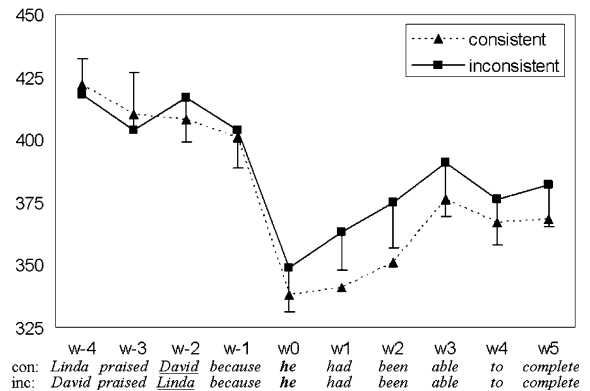


Fig. 1. Mean reading times (in ms) for the consistent and inconsistent condition in Experiment 1 (error bars indicate one side of a 95% confidence interval based on the MSE -value; see Masson and Loftus, 2003).

This pattern of results suggests that implicit causality information becomes available rapidly enough to have an impact on the interpretation of a pronoun occurring in mid-sentence. The immediate focusing account, where implicit causality information immediately makes the bias-consistent referent more accessible than the bias-inconsistent referent, just before the referring pronoun comes along, explicitly predicted such results. The clausal integration account, however, assumes that implicit causality is used *only* after the bulk of the subordinate clause has been read and the interpretations of the two clauses are integrated into a single interpretation for the sentence as a whole, an assumption that is difficult to reconcile with our findings.

Because the inconsistency effect in our experiment did not reach significance on the critical pronoun itself, and, presumably due to spill-over, emerged as a reliable effect on the two subsequent words only, our findings in fact provide a straightforward explanation for why Stewart et al. (2000) did not find any evidence for an early effect of implicit causality. Recall that Stewart et al. presented the sentences in two large fragments, with the split right after the critical pronoun. If, like the readers in our study, the readers in the Stewart et al. experiments only really slowed down right *after* the bias-inconsistent pronoun, this early effect would be visible in reading times to the second fragment only, and would as such mistakenly be taken as support for late clausal integration. Thus, the present word-by-word reading time results confirm our earlier hypothesis about Stewart et al.'s early null result. More generally, our results testify to the importance of tracking the use of a potentially relevant cue to sentence comprehension with sufficient temporal resolution.

Although our findings are in line with the immediate focusing account, we need to address two possible concerns over the non-cumulative word-by-word self-paced

Table 2
Mean reading times and analysis of variance results for Experiment 1

	Wordposition									
	–4	–3	–2	–1	0	1	2	3	4	5
	e.g.:									
Consistent:	<i>Linda</i>	<i>praised</i>	<i>David</i>	<i>because</i>	<i>he</i>	<i>had</i>	<i>been</i>	<i>able</i>	<i>to</i>	<i>complete</i>
Inconsistent:	<i>David</i>	<i>praised</i>	<i>Linda</i>	<i>because</i>	<i>he</i>	<i>had</i>	<i>been</i>	<i>able</i>	<i>to</i>	<i>complete</i>
Reading time (in ms)										
<i>Total</i>										
Consistent	422	410	408	401	338	341	351	376	367	368
Inconsistent	418	404	417	404	349	363	375	391	376	382
<i>NP1-bias</i>										
Consistent	420	403	409	403	347	350	340	388	373	371
Inconsistent	411	409	425	424	357	382	371	401	387	378
<i>NP2-bias</i>										
Consistent	425	417	408	400	329	333	362	364	361	365
Inconsistent	424	399	410	384	340	343	379	380	364	386
<i>F1 (Consistency, df = 1, 23)</i>										
<i>F</i>	<1	<1	1.065	<1	1.452	9.107	7.477	1.993	1.015	3.187
<i>MSE</i>	1127	2992	1755	1290	1772	1241	1815	2561	1774	1540
<i>p</i>	.526	.623	.313	.720	.245	.006*	.012*	.171	.324	.087
<i>F2 (Consistency, df = 1, 38)</i>										
<i>F</i>	<1	<1	1.362	<1	1.733	9.379	9.643	1.966	1.386	3.423
<i>MSE</i>	1944	1099	1137	2269	1096	955	1239	2625	894	1257
<i>p</i>	.685	.519	.250	.848	.196	.004*	.004*	.169	.246	.072
<i>minF' (Consistency)</i>										
<i>F</i>	<1	<1	<1	<1	<1	4.620	4.211	<1	<1	1.650
<i>df</i>	1, 59	1, 48	1, 53	1, 55	1, 55	1, 57	1, 53	1, 58	1, 52	1, 56
<i>p</i>	.731	.694	.443	.866	.378	.036*	.045*	.324	.478	.204
<i>F1 (Consistency × Verb Bias, df = 1, 23)</i>										
<i>F</i>	<1	1.235	<1	6.161	<1	2.181	1.326	<1	<1	1.244
<i>MSE</i>	2025	2854	1208	1278	858	1289	985	2850	1151	1104
<i>p</i>	.668	.278	.378	.021*	.973	.153	.261	.941	.425	.276
<i>F2 (Consistency × Verb Bias, df = 1, 38)</i>										
<i>F</i>	<1	3.096	<1	3.086	<1	2.789	1.026	<1	<1	<1
<i>MSE</i>	1944	1099	1137	2269	1096	955	1239	2625	894	1257
<i>p</i>	.701	.087	.477	.087	.982	.103	.317	.989	.382	.352
<i>minF' (Consistency × Verb Bias)</i>										
<i>F</i>	<1	<1	<1	2.056	<1	1.223	<1	<1	<1	<1
<i>df</i>	1, 60	1, 41	1, 61	1, 61	1, 57	1, 53	1, 60	1, 39	1, 55	1, 61
<i>p</i>	.773	.353	.577	.157	.982	.274	.450	.992	.552	.474

* $p < .05$.

reading method that we used. First, although a word-by-word reading paradigm has a higher temporal resolution than the clause-by-clause paradigm of Stewart et al., the task is of course somewhat unnatural, as people normally do not press a button after reading each individual word (but see Mitchell, 2004, for an eloquent defense of the self-paced reading task). Furthermore, and perhaps more important in the context at hand, our use of a non-cumulative moving-window paradigm made it

impossible for participants to regress to earlier parts of the sentence. It could be argued that readers will adapt to this situation by resorting to a more incremental processing strategy, in which they more immediately use the information afforded by each word—such as implicit causality—than they would do in unconstrained reading. To address these two potential concerns over our self-paced reading study, we repeated the experiment with the eye tracking methodology.

Experiment 2

Method

Participants

Participants were 24 members from the Utrecht University community (23 female, mean age 21, range 18–34 years) who received money for their participation.

Materials

The materials were the same as in Experiment 1, with two small exceptions. One is that in order to raise fixation probability, the two words following the critical pronoun *he* were changed, if necessary, such that they had a minimal length of five characters. Furthermore, because of demands associated with the eye tracker and the variable length of words within the five-word spill-over region, the one or two last words of this region (i.e., words 9 and 10 in the results table) were sometimes displayed on the next line. The same fillers, randomizations, and comprehension questions were used.

Procedure

A head-mounted *SMI* eye tracker was used to monitor eye movements. The tracker had an angular resolution of 20 s of arc, and monitored both eye's gaze locations every 4 ms. The stories were presented as a whole on a CRT (Nokia Multigraph 446xpro) screen. Before presentation, a fixation mark appeared on screen at the position of the first word of the first sentence. Participants were instructed to fixate this mark before they made a story visible by pressing a button. After reading a story the participants again pressed this button to progress. The comprehension questions that followed 40 pseudo-randomly determined trials were answered using two buttons on the same response box. Each session started with a written instruction, after which we mounted and calibrated the eye tracker. During the calibration procedure the participants had to fixate a random sequence of dots at various locations on screen. Upon successful calibration the experiment started with five practice trials, two followed by a question. Before the experimental trials were presented the eye tracker was recalibrated. This procedure was repeated three times throughout the experiment. Each session ended with the administration of a Dutch reading span task, and was completed within 50 min.³

³ Due to very low variability, the results of the reading span task did not allow for a sensible median split, and we therefore refrained from analyzing the results of Experiment 2 as a function of reading span group.

Analysis

All participants scored above 85% correct (mean score 94%) on the comprehension questions. Prior to all analyses, we removed 5.6% of the trials because major tracker losses and eye blinks made it impossible to determine the course of fixations at or directly around the critical pronoun. Furthermore, if a fixation was shorter than 80 ms and within one character space of the previous or next fixation, it was assimilated to this fixation. All remaining fixations shorter than 80 ms, as well as fixations longer than 1200 ms or containing blinks, were excluded (8.4%). Because short words like pronouns receive very few fixations, we extended the region for the pronoun by six characters to the left, if the pronoun was skipped during first pass reading. This *leftward-shifting* procedure is based on evidence that readers are able to obtain lexical information from words beginning six characters from a particular fixation, and on evidence that the perceptual span is asymmetric to the right of a fixation (see Rayner & Sereno (1994) for more discussion). In the current experiment, the procedure increased the probability for fixating the critical pronoun from 42 to 86%. If after this procedure, the pronoun was still not fixated, the data point for the trial was treated as missing data (as was done for any other word that was not fixated during first pass reading).

As in Experiment 1, we will report mean reading times and statistical analysis for the factor Consistency (and its interaction with Verb Bias) for four words in the pre-critical region, the critical pronoun, and five words in the spill-over region. We will discuss three different *first-pass* eye movement measures. *First-Fixation Duration* reflects the duration of the very first fixation on a word. *First-Gaze Duration* is the total reading time of a word before the reader either moves on, or looks back in the text. Finally, *Regression Path Durations* are the sum of fixation and saccade durations from the time when the reader encounters a word, to the time when the reader enters the region after this word. This means that if the reader looks back after reading a word (a regression), the regression path time includes all fixation and saccade durations of this regression. All measures are presumed to be sensitive for processes that occur relatively early during comprehension.⁴

⁴ First-gaze durations consist of first fixation durations plus any additional fixations on the word that directly follow the first fixation. Similarly, the regression path duration incorporates the first gaze duration, as well as the duration of the regression (if present) that directly follows after fixating a particular word. As a result, the different measures incrementally allow processes that occur somewhat later to enter the measure.

Results and discussion

Figs. 2–4 display the mean first fixation duration, first-gaze reading time, and regression path time as a function of Consistency and word position, collapsed across Verb Bias. Tables 3–5 report the associated statistics, as well as additional data for the Consistency by Verb Bias interaction. In the pre-critical region, none of the three measures revealed a significant main effect of Consistency. However, relative to their bias-consistent counterparts, pronouns that were inconsistent with the implicit causality bias of the preceding verb reliably perturbed the reading process at or shortly after the pronoun. At the pronoun itself, the inconsistency effect emerged most clearly in the regression path duration. Furthermore, three words after the critical pronoun, significant inconsistency effects emerged both in first fixation and first gaze duration.

As in Experiment 1, there were no reliable interactions between Consistency and Verb Bias at or directly following the critical pronoun. At the second word in

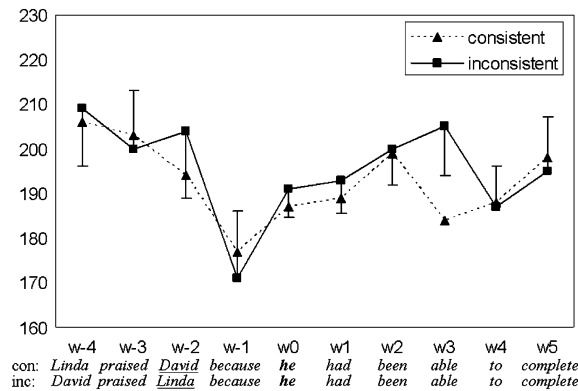


Fig. 2. Mean first fixation durations (in ms) for the consistent and inconsistent condition in Experiment 2.

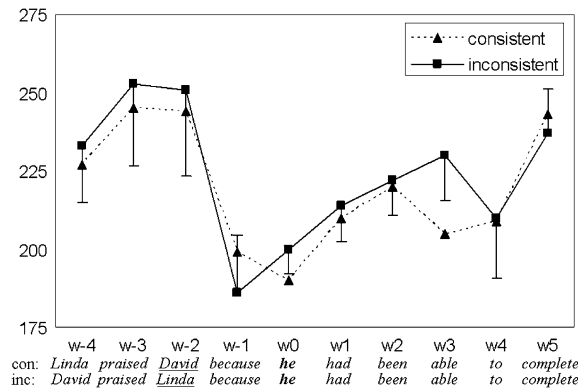


Fig. 3. Mean first-gaze durations (in ms) for the consistent and inconsistent condition in Experiment 2.

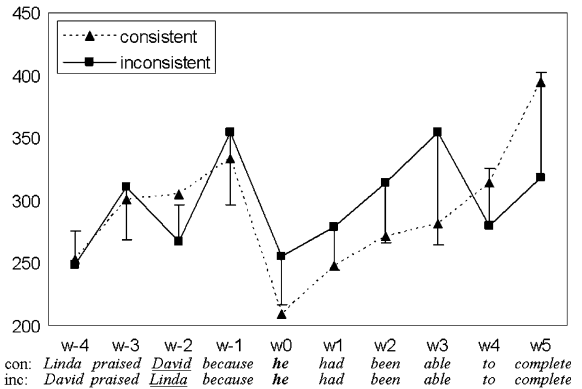


Fig. 4. Mean regression path durations (in ms) for the consistent and inconsistent condition in Experiment 2.

the pre-critical region, we did however obtain a significant Consistency by Verb Bias interaction, in first-gaze and regression path durations. In both measures the reading times on this word in NP1 stories were longer in the consistent versions than in the inconsistent ones, whereas the opposite was observed for NP2 stories. We are inclined to interpret this as a chance effect, possibly related to the fact that, in contrast to the critical region, different words are involved across conditions.

As in Experiment 1, we obtained effects of implicit causality right after—and in this case even at—the critical pronoun, only two words into the subordinate clause. This shows that the early effects of implicit causality on sentence processing obtained in Experiment 1 cannot be attributed to particular strategies adopted in the self-paced reading task. Moreover, these eye tracking findings again show that implicit causality becomes available very early in the comprehension process, as predicted by the immediate focusing account. Regardless of whether readers make their way through a story in self-paced or unconstrained reading, the implicit causality information afforded by biased interpersonal verbs like *praise* or *disappoint* is brought to bear on sentence comprehension rapidly enough to have an impact on the interpretation of a referring pronoun occurring in mid-sentence, only two words into the subordinate clause.

General discussion

In two comprehension experiments, we examined the recent claim (e.g., Garnham, 2001; Stewart et al., 2000) that verb-based implicit causality information is only used during sentence-final clausal integration. We did so by looking for mid-sentence reading delays caused by pronouns that are inconsistent with the bias of a preceding implicit causality verb (e.g., “*David praised Linda because he...*”). In Experiment 1, we kept track of the

Table 3
Mean first fixation durations and analysis of variance results for Experiment 2

	Wordposition									
	–4	–3	–2	–1	0	1	2	3	4	5
	e.g.:									
Consistent:	<i>Linda</i>	<i>praised</i>	<i>David</i>	<i>because</i>	<i>he</i>	<i>had</i>	<i>been</i>	<i>able</i>	<i>to</i>	<i>complete</i>
Inconsistent:	<i>David</i>	<i>praised</i>	<i>Linda</i>	<i>because</i>	<i>he</i>	<i>had</i>	<i>been</i>	<i>able</i>	<i>to</i>	<i>complete</i>
First fixation duration (in ms)										
<i>Total</i>										
Consistent	206	203	194	177	187	189	199	184	188	198
Inconsistent	209	200	204	171	191	193	200	205	187	195
<i>NP1-bias</i>										
Consistent	203	201	187	178	193	189	203	174	188	191
Inconsistent	199	192	206	166	191	194	199	205	185	190
<i>NP2-bias</i>										
Consistent	209	204	200	175	180	189	194	194	188	204
Inconsistent	219	208	202	176	191	191	200	204	188	200
<i>F1 (Consistency)</i>										
<i>F</i>	<1	<1	1.931	<1	1.754	<1	<1	14.268	<1	<1
<i>df</i>	1, 23	1, 23	1, 21	1, 20	1, 23	1, 23	1, 23	1, 23	1, 23	1, 23
<i>MSE</i>	937	961	1291	1281	222	299	357	684	480	828
<i>p</i>	.633	.730	.179	.506	.198	.329	.761	.001*	.762	.662
<i>F2 (Consistency)</i>										
<i>F</i>	<1	<1	<1	1.033	<1	<1	<1	4.913	<1	<1
<i>df</i>	1, 38	1, 38	1, 36	1, 38	1, 38	1, 38	1, 38	1, 37	1, 37	1, 36
<i>MSE</i>	952	1755	843	834	573	344	484	1909	1408	927
<i>p</i>	.791	.568	.487	.316	.334	.515	.780	.033*	.880	.325
<i>minF' (Consistency)</i>										
<i>F</i>	<1	<1	<1	<1	<1	<1	<1	3.654	<1	<1
<i>df</i>	1, 56	1, 40	1, 51	1, 38	1, 61	1, 60	1, 59	1, 56	1, 52	1, 32
<i>p</i>	.816	.767	.533	.576	.434	.585	.836	.061	.892	.688
<i>F1 (Consistency × Verb Bias)</i>										
<i>F</i>	1.411	<1	<1	1.207	3.651	<1	1.098	2.663	<1	<1
<i>df</i>	1, 23	1, 23	1, 21	1, 20	1, 23	1, 23	1, 23	1, 23	1, 23	1, 23
<i>MSE</i>	918	1331	1870	695	275	345	497	929	883	765
<i>p</i>	.247	.415	.401	.285	.069	.661	.306	.116	.726	.790
<i>F2 (Consistency × Verb Bias)</i>										
<i>F</i>	<1	<1	<1	2.895	1.527	<1	<1	<1	<1	<1
<i>df</i>	1, 38	1, 38	1, 36	1, 38	1, 38	1, 38	1, 38	1, 37	1, 37	1, 36
<i>MSE</i>	952	1755	843	834	573	344	484	1909	1408	927
<i>p</i>	.471	.601	.370	.097	.224	.676	.562	.357	.570	.363
<i>minF' (Consistency × Verb Bias)</i>										
<i>F</i>	<1	<1	<1	<1	1.076	<1	<1	<1	<1	<1
<i>df</i>	1, 58	1, 59	1, 51	1, 37	1, 59	1, 59	1, 56	1, 56	1, 40	1, 27
<i>p</i>	.538	.657	.536	.362	.304	.761	.612	.421	.764	.797

* $p < .05$.

impact of implicit causality by means of word-by-word self-paced reading. In Experiment 2, we used eye tracking, as participants freely read through the same materials.

In both experiments, we obtained evidence for such pronoun-induced delays. In self-paced reading (Experiment 1), pronouns that were inconsistent with the

implicit causality bias of the verb reliably slowed down reading right after the pronoun. In unconstrained reading (Experiment 2) such pronouns also immediately perturbed the reading process, as indexed by significant delays, in various first pass measures, at and shortly after the critical pronoun. The eye tracking results con-

Table 4
Mean first-gaze durations and analysis of variance results for Experiment 2

	Wordposition									
	-4	-3	-2	-1	0	1	2	3	4	5
e.g.:										
Consistent:	<i>Linda</i>	<i>praised</i>	<i>David</i>	<i>because</i>	<i>he</i>	<i>had</i>	<i>been</i>	<i>able</i>	<i>to</i>	<i>complete</i>
Inconsistent:	<i>David</i>	<i>praised</i>	<i>Linda</i>	<i>because</i>	<i>he</i>	<i>had</i>	<i>been</i>	<i>able</i>	<i>to</i>	<i>complete</i>
First-gaze duration (in ms)										
<i>Total</i>										
Consistent	227	245	244	199	190	210	220	205	209	243
Inconsistent	233	253	251	186	200	214	222	230	210	237
<i>NP1-bias</i>										
Consistent	227	252	229	206	195	211	225	194	208	250
Inconsistent	224	230	232	187	202	209	225	224	201	239
<i>NP2-bias</i>										
Consistent	227	237	259	192	185	208	215	216	209	236
Inconsistent	241	275	269	184	197	218	219	236	219	234
<i>F1 (Consistency)</i>										
<i>F</i>	<1	<1	<1	1.968	5.705	<1	<1	12.814	<1	<1
<i>df</i>	1, 23	1, 23	1, 21	1, 20	1, 23	1, 23	1, 23	1, 23	1, 23	1, 23
<i>MSE</i>	1835	3890	4306	1924	351	747	691	1136	2103	1147
<i>p</i>	.520	.524	.678	.176	.026*	.468	.696	.002*	.861	.344
<i>F2 (Consistency)</i>										
<i>F</i>	<1	<1	<1	1.387	2.096	<1	<1	6.054	1.170	3.094
<i>df</i>	1, 38	1, 38	1, 36	1, 38	1, 38	1, 38	1, 38	1, 37	1, 37	1, 36
<i>MSE</i>	2636	3663	1134	1455	1221	1014	1558	2601	2228	1285
<i>p</i>	.590	.807	.987	.246	.156	.494	.821	.019*	.286	.087
<i>minF' (Consistency)</i>										
<i>F</i>	<1	<1	<1	<1	1.532	<1	<1	4.111	<1	<1
<i>df</i>	1, 61	1, 48	1, 36	1, 57	1, 58	1, 59	1, 57	1, 59	1, 24	1, 37
<i>p</i>	.678	.820	.992	.371	.221	.616	.844	.047*	.863	.402
<i>F1 (Consistency × Verb Bias)</i>										
<i>F</i>	2.074	7.244	<1	<1	<1	1.122	<1	<1	<1	<1
<i>df</i>	1, 23	1, 23	1, 21	1, 20	1, 23	1, 23	1, 23	1, 23	1, 23	1, 23
<i>MSE</i>	842	3000	401	1091	491	730	809	1904	2014	1036
<i>p</i>	.163	.013*	.405	.472	.540	.300	.713	.596	.344	.477
<i>F2 (Consistency × Verb Bias)</i>										
<i>F</i>	<1	4.064	2.261	<1	<1	1.007	<1	<1	<1	<1
<i>df</i>	1, 38	1, 38	1, 36	1, 38	1, 38	1, 38	1, 38	1, 37	1, 37	1, 36
<i>MSE</i>	2636	3663	1134	1455	1221	1014	1558	2601	2228	1285
<i>p</i>	.372	.051	.141	.472	.665	.322	.976	.858	.866	.839
<i>minF' (Consistency × Verb Bias)</i>										
<i>F</i>	<1	2.603	<1	<1	<1	<1	<1	<1	<1	<1
<i>df</i>	1, 59	1, 61	1, 35	1, 53	1, 60	1, 59	1, 39	1, 45	1, 39	1, 39
<i>p</i>	.447	.112	.464	.608	.722	.469	.974	.866	.868	.893

* $p < .05$.

firm that the early use of implicit causality information in Experiment 1 is not a consequence of unnatural comprehension strategies induced by the word-by-word self-paced reading task. Furthermore, because we used identical critical words across bias-consistent and -inconsistent conditions, our results also unequivocally rule

out the possibility [suggested by Oakhill et al. (1998) and Garnham (2001)] that the implicit causality effects observed in language comprehension are artifacts of 'low-level' differences between bias-consistent and -inconsistent critical fragments. Hence, across experiments and methodologies, we have evidence that readers can

Table 5
Mean regression path durations and analysis of variance results for Experiment 2

	Wordposition									
	–4	–3	–2	–1	0	1	2	3	4	5
e.g.:										
Consistent:	<i>Linda</i>	<i>praised</i>	<i>David</i>	<i>because</i>	<i>he</i>	<i>had</i>	<i>been</i>	<i>able</i>	<i>to</i>	<i>complete</i>
Inconsistent:	<i>David</i>	<i>praised</i>	<i>Linda</i>	<i>because</i>	<i>he</i>	<i>had</i>	<i>been</i>	<i>able</i>	<i>to</i>	<i>complete</i>
Regression path duration (in ms)										
<i>Total</i>										
Consistent	253	301	305	333	209	248	271	281	314	395
Inconsistent	249	311	267	354	255	279	314	354	280	318
<i>NP1-bias</i>										
Consistent	249	317	299	357	212	250	276	264	313	466
Inconsistent	240	266	248	329	282	289	311	290	275	306
<i>NP2-bias</i>										
Consistent	256	284	310	308	205	245	265	297	314	323
Inconsistent	258	355	276	378	227	269	316	417	284	330
<i>F1 (Consistency)</i>										
<i>F</i>	<1	<1	8.438	<1	6.144	3.604	3.531	2.858	2.375	3.473
<i>df</i>	1, 23	1, 23	1, 21	1, 20	1, 23	1, 23	1, 23	1, 23	1, 23	1, 23
<i>MSE</i>	3967	9864	4765	18503	8215	6577	12735	44429	11820	40105
<i>p</i>	.809	.644	.008*	.485	.021*	.070	.073	.104	.137	.075
<i>F2 (Consistency)</i>										
<i>F</i>	<1	<1	3.946	<1	4.772	4.609	3.229	1.946	<1	1.920
<i>df</i>	1, 38	1, 38	1, 36	1, 38	1, 38	1, 38	1, 38	1, 37	1, 37	1, 36
<i>MSE</i>	4828	10685	3918	10356	11862	3857	7750	33990	11102	42178
<i>p</i>	.904	.913	.055	.519	.035*	.038*	.080	.171	.598	.174
<i>minF' (Consistency)</i>										
<i>F</i>	<1	<1	2.688	<1	2.685	2.022	1.686	1.157	<1	1.236
<i>df</i>	1, 54	1, 42	1, 56	1, 55	1, 60	1, 53	1, 58	1, 60	1, 45	1, 59
<i>p</i>	.913	.917	.107	.633	.107	.161	.199	.286	.618	.271
<i>F1 (Consistency × Verb Bias)</i>										
<i>F</i>	<1	5.933	<1	3.042	1.645	<1	<1	1.294	<1	6.781
<i>df</i>	1, 23	1, 23	1, 21	1, 20	1, 23	1, 23	1, 23	1, 23	1, 23	1, 23
<i>MSE</i>	3480	15134	4286	16788	8472	4156	8524	41890	14453	24463
<i>p</i>	.635	.023*	.540	.096	.212	.603	.663	.267	.867	.016*
<i>F2 (Consistency × Verb Bias)</i>										
<i>F</i>	<1	5.628	<1	<1	1.082	<1	<1	2.482	<1	2.479
<i>df</i>	1, 38	1, 38	1, 36	1, 38	1, 38	1, 38	1, 38	1, 37	1, 37	1, 36
<i>MSE</i>	4828	10685	3918	10356	11862	3857	7750	33990	11102	42178
<i>p</i>	.497	.023*	.838	.445	.305	.882	.939	.124	.682	.124
<i>minF' (Consistency × Verb Bias)</i>										
<i>F</i>	<1	2.888	<1	<1	<1	<1	<1	<1	<1	1.815
<i>df</i>	1, 45	1, 58	1, 43	1, 51	1, 61	1, 44	1, 40	1, 46	1, 31	1, 56
<i>p</i>	.695	.095	.845	.517	.422	.887	.940	.361	.876	.183

* $p < .05$.

very rapidly recruit verb-based implicit causality information in the service of comprehension, rapidly enough to impact on the interpretation of a pronoun early the subordinate clause.

In the immediate focusing account (Greene & Mc-Koon, 1995; Long & De Ley, 2000; McKoon

et al., 1993), a sentence-initial fragment like “*David praised Linda because*” is assumed to immediately bring Linda into focus at the expense of David. Because readers and listeners prefer to relate a personal pronoun to the most focused antecedent (see Arnold, 1998, for review), this account correctly predicted the

early processing problem that we observed, a delay in reading at and immediately after the subsequent pronoun *he*. No such early effect, however, was predicted by the clausal integration account (Garnham, 2001; Garnham et al., 1996; Stewart et al., 2000), at least not in its most common formulation, in which implicit causality becomes relevant in sentence-final ‘retroactive’ clausal integration only. We turn to a more detailed discussion of the possible mechanisms behind our effects shortly hereafter, and will then also examine the viability of a more refined clausal integration account.

In our self-paced reading experiment, statistically reliable effects of implicit causality first emerged on the two words that followed the critical pronoun. As discussed before, this provides a straightforward explanation for why Stewart et al. may have failed to detect an early implicit causality effect in their two-fragment self-paced reading study. As such, our findings also testify to the importance of tracking the use of a potentially relevant cue to sentence comprehension with sufficient temporal resolution, so that immediate small effects, as well as their immediate spill-over, can be accurately detected. Note that in our eye tracking experiment, bias-inconsistent pronouns had most of their impact on the reading times *after* the pronoun as well. This spill-over phenomenon is entirely consistent with the fact that, in an EEG experiment with (Dutch) spoken sentences like “*Anna shot at Linda as he jumped over the fence*”, the processing consequences of the referentially failing pronoun *he* can be seen to extend in the EEG for at least a second after pronoun onset (Van Berkum et al., 2004; see also Osterhout & Mobley, 1995). Furthermore, because in the present study we had equated the pronoun as well as five words that followed the pronoun across consistent and inconsistent versions of each item, we can be confident that the spill-over delays observed here were elicited by whether the pronoun confirmed or disconfirmed the implicit causality afforded by the verb.

Although the interaction of Consistency and Verb Bias should be interpreted with care, we note that whether the verb was biased towards NP1 (e.g., *disappoint*) or NP2 (e.g., *praise*) did not reliably modulate the size of the inconsistency effect in our data. As such, our results are different from those of Long and De Ley (2000), who observed a significant inconsistency effect for NP2-biased verbs only. In fact, if anything, our data tend to go the opposite way, with numerically larger delays after NP1-verbs than after NP2-verbs at all relevant word positions in Experiment 1 and Experiment 2. Again, we emphasize that the comparison of inconsistency effects elicited by NP1 verbs and NP2 verbs is confounded with the effects of distance between anaphor and antecedent, of first mention, and of the antecedent’s structural position. Because of the specific purpose of our study, we made no attempt to control for any of

these potential confounds, and optimized the design on other dimensions. Follow-up work will have to examine why interactions with verb bias can come and go, and to what extent the above-mentioned confounding factors play a role.

Our findings allow us to reject the hypothesis that verb-based implicit causality information is only used during late, clause-final integration. However, they also bear on two other suggested limitations on the use of implicit causality in comprehension. One is that people might only rapidly exploit this cue as part of some unnatural processing strategy, elicited by experiments in which implicit causality always provides reliable information (Garnham, 2001). Because we obtained an early effect of verb bias in experiments with as many bias-consistent as bias-inconsistent continuations, our results clearly do not support this suggestion. Our findings also refute the hypothesis (Garnham et al., 1996) that verb-based implicit causality might only be relevant to comprehension when the referent of a pronoun cannot be disambiguated by means of morphosyntactic (e.g., gender) cues. In our materials, pronoun gender unambiguously rules out one of the two candidate antecedents. However, in spite of the reliable presence of a strong gender cue, readers apparently cannot help but use the probabilistic implicit causality cue as well.

How does implicit causality affect comprehension?

The latter observation raises an interesting question, for why would people use a relatively weak probabilistic cue if they have a much stronger and highly reliable cue at their disposal? So far, we have discussed the results in the context of two extant theoretical options: immediate focusing and clausal integration. However, although our findings disconfirm the clausal integration hypothesis, this does not by itself compel us to adopt immediate focusing as the correct account. In the remainder, we will examine several possible underlying mechanisms, of which one actually involves a refined version of clausal integration.

The integration account that our findings allow us to dispense with is one in which implicit causality only becomes relevant during *late* ‘retroactive’ clausal integration, at or close to the end of the *because*-clause. This theoretical option is most clearly adopted by Stewart et al. (2000), who were prepared to take implicit causality effects at the pronoun as evidence for immediate focussing and against clausal integration. In his review, Garnham (2001) appears to adopt a similar view. In addition, Long and De Ley (2000) have noted that sentence-final clausal integration of implicit causality information is explicitly predicted by a more general model of multi-clause sentence comprehension, the *Connective Integration Model* (Millis & Just, 1994). According to this model, when readers encounter a connective like *because*, they set the representation of the first clause

aside in working memory until they have independently processed the complete second clause. Crucially, only upon having completed the latter would the information provided by the first clause be related to that provided by the second. The clause-medial implicit causality effects that we observed reveal that such relations are computed much earlier, and thereby show that this specific claim of the Connective Integration Model cannot be correct.

However, our findings are not necessarily inconsistent with an *incremental* clausal integration account, in which the information made available by the subordinate clause is ‘retroactively’ related to the interpretation of the main clause *on a word-by-word basis*. One such account was proposed McDonald and MacWhinney (1995), who in a probe verification task obtained very early implicit causality effects in comprehension, right after the pronoun. McDonald and MacWhinney took their findings as evidence that the implicit causality information provided by the main clause verb can be used as soon as a referring pronoun in the subordinate clause needs to be resolved. Along the same lines, Garnham et al. (1996) also proposed a limited role for implicit causality information during pronoun-induced anaphoric processing. Note, however, that Garnham and colleagues assumed that implicit causality is only used to resolve a morphosyntactically ambiguous pronoun, as in “*Betty punished Diane because she didn’t do the dishes.*” In gender-disambiguated cases such as “*Betty punished Roger because he didn’t do the dishes,*” implicit causality was deemed to be irrelevant, this because “a coherent representation can be set up without considering which participant in an event of punishing is *usually* the cause” (p. 538). To explain our results within an incremental clausal integration framework, the latter assumption would need to be abandoned.

As discussed before, our findings were predicted by the immediate focusing account of how implicit causality affects sentence comprehension (Garnham et al., 1996; Greene & McKoon, 1995; McKoon et al., 1993). According to the immediate focusing model, the implicit causality afforded by something like “*David praised Linda because...*” changes the relative availability of the two discourse entities involved, independent of any potentially upcoming pronouns. In fact, in most characterizations of immediate focusing (Garnham et al., 1996; McKoon et al., 1993; but see Long & De Ley, 2000), the consequences for later pronoun resolution are deeply coincidental: implicit causality happens to highlight certain discourse entities at the expense of others as the system incrementally updates its model of the discourse, and later pronouns happen to be sensitive to the availability of candidate referents (see Gerrig & McKoon, 1998 for a similar view).

Immediate focusing has been characterized as ‘proactive’ (Garnham, 2001) because in this mechanism,

and in contrast to late (as well as the incremental) clausal integration, implicit causality has an effect on the discourse model before the assumed moment of ‘retroactive’ integration. However, our early effects of implicit causality are also compatible with a much more ‘forward-looking’ mechanism. In “*David praised Linda because...*,” the implicit causality associated with *praise* supplies information about whose behavior or state is the more likely immediate cause of the event at hand. As such, and particularly when combined with a connective such as *because* (Ehrlich, 1980; McKoon et al., 1993), it can also support specific *expectations or predictions* about how the unfolding utterance and wider discourse might continue. In particular, comprehenders might anticipate that the *because*-clause is going to provide information about Linda, perhaps even actually refer to Linda. On the basis of research with goal- and source-oriented verbs of transfer (*to receive, to send*), Arnold (2001) has recently proposed that readers and listeners continuously estimate the likelihood that a referent will be continued in the upcoming discourse. In addition, recent experiments with head-mounted eye tracking and event-related brain potentials suggest that people can use the semantics afforded by verbs like *ate* or *told* to anticipate specific upcoming verb arguments, such as those that refer to edible things (Altmann & Kamide, 1999; Kamide et al., 2003) or animate entities (Nieuwland & Van Berkum, in press). We see no deep reason why the semantics of *receive, ate, and told* would support such anticipation whereas the equally constraining semantics of a verb like *praise* would not. In our study, predictions about who will soon be talked about and/or referred to would be disconfirmed by the bias-inconsistent pronoun, and might as such underlie the observed reading time delays.

Exploring the potential of anticipatory mechanisms somewhat further, we note that a fragment like “*David praised Linda because...*” might in fact also lead people to predict specific upcoming *words*. Research with gender-inflected languages has recently shown that listeners and readers can use sentential as well as discourse context to rapidly make such lexical predictions, including, for nouns, their lexically stored gender (Otten & Van Berkum, 2004; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005a; Wicha, Moreno, & Kutas, 2004). In view of these findings, it is not inconceivable that a fragment like “*David praised Linda because*” allowed our readers to not only anticipate upcoming reference to Linda, but to actually also anticipate the specific pronoun *she*. To the extent that they do so, *he* would disconfirm the lexical prediction and cause readers to slow down.

We are currently conducting ERP research that may help us decide whether early implicit causality effects arise because of incremental clausal integration, immedi-

ate focusing, or truly ‘proactive’ anticipation.⁵ Pending relevant data, we reemphasize the puzzle posed by any incremental clausal integration account (cf. Garnham et al., 1996): why use a cue about who *might* be talked about when a definitive cue about who *is* talked about is available? Garnham and colleagues resolved the puzzle by assuming that implicit causality would only impact on the resolution of a morphosyntactically ambiguous pronoun, an assumption that our findings now allow us to reject. If incremental clausal integration is modeled in terms of simultaneous constraint satisfaction (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus & Trueswell, 1995; Trueswell & Tanenhaus, 1994), one might resolve the puzzle by assuming that the latter cue, although formally overruled, is strong enough to delay the system in converging onto a stable representation of the sentence up to that point. Another way out might be to assume that when a pronoun is encountered, implicit causality is actually used to fixate its referent *before* the system considers the gender of the pronoun.⁶

Note, however, that whereas within an incremental clausal integration account, some additional assumptions are needed to make sense of implicit causality effects at gender-disambiguated pronouns, the puzzle is solved for free if anticipatory or immediate focusing mechanisms are at work. The reason is that in both mechanisms, implicit causality information is used *before* the pronoun comes along, i.e., before more definitive evidence on causality has become available.

⁵ The difference between implicit causality exerting its influence via immediate focusing or via the anticipation of what or whom will be talked about next is a subtle one, because focus and anticipation might be deeply related (see Long & De Ley, 2000, for exactly this position). However, whereas immediate focusing can *in principle* be viewed as a ‘blind’ incremental mechanism that simply modulates the availability of certain discourse entities without looking beyond the linguistic input processed so far, the anticipation of a specific upcoming topic or referent is *by definition* a predictive, forward-looking mechanism. It is for this reason that we distinguish these two types of explanations for the impact of implicit causality on language comprehension. At the same time, we sympathize with an account in which the highlighting of particular discourse entities is inextricably intertwined with—and perhaps in some sense even equivalent to—expectations about what or whom will be talked about next (e.g., Arnold, 2001). After all, what’s the point of focusing on something that has just been completely dealt with?

⁶ The idea that a ‘non-local’ probabilistic semantic constraint can sometimes take precedence over a local and hard morphosyntactic constraint may strike one as odd. However, recent evidence from ERPs suggests that when fixing the proper reading of a Dutch relative pronoun, weak discourse-referential factors can also momentarily prevail over a local syntactic gender constraint (Van Berkum et al., 1999a, Van Berkum, Brown, & Hagoort, 1999b; see also Van Gompel & Liversedge, 2003).

How special is implicit causality?

Possibly inspired by the various principled classification schemes that have been proposed for implicit causality verbs (see Rudolph & Försterling, 1997, for review), psycholinguists are sometimes inclined to grant a special status to the causality related semantic features of verbs like *praise* and *disappoint*. In terms of on-line processing, however, we currently see no reason to assume that the semantic constraints provided by implicit causality verbs are brought to bear on interpretation in a ‘special way,’ different from how other (e.g., discourse-level) semantic constraints can impact on comprehension as it unfolds. Elsewhere (Van Berkum et al., 1999a, 1999b, 2003; Van Berkum et al., 2005b; Nieuwland & Van Berkum, in press), we have provided ERP evidence that sentence-internal semantic factors (e.g., local verb-argument animacy relations) and discourse-level ‘pragmatic’ factors (e.g., global theme, cartoon-like genre, voice-inferred speaker characteristics) can simultaneously constrain interpretation, and do so in functionally equivalent ways. In view of the rapid impact of implicit causality reported here, and in the absence of compelling data to the contrary, we propose that implicit causality is just another bit of meaning, which, in interaction with other such bits jointly determines the interpretation of an unfolding utterance and any predictions that can be made from there (see Arnold, 2001, for a very similar proposal). The fact that in this particular case the meaning unlocked by language happens to be about interpersonal exchanges, although critical if one studies social interaction, seems of little relevance if one’s interest is in basic mechanisms of incremental language interpretation.

If implicit causality is just another semantic cue in, say, a graded, constraint-based processing architecture (cf. MacDonald et al., 1994; Tanenhaus & Trueswell, 1995), this suggests a natural answer to several other questions about implicit causality in language comprehension. One question that has received some attention in the literature (e.g., Ehrlich, 1980; McKoon et al., 1993; Stevenson, Crawley, & Kleinman, 1994) is whether the impact of implicit causality verbs critically depends on a subsequent causal connective, such as *because*. Under the model that we are exploring, there would be no principled “yes” or “no” to this question, because interpretation is dynamically determined by a wide variety of semantic constraints, as well as by the sequencing and relative timing with which these constraints become available within the unfolding utterance. This means that whereas a relatively weakly biased implicit causality verb might need an additional *because* (or other causal connective) to have a measurable impact, more strongly biased verbs may well be able to single-handedly modulate the comprehender’s focus and/or expectations. A graded constraint-based perspective on implicit causality

may also ultimately help us understand why the impact of implicit causality varies across studies in terms of timing and effect size. Of course, whether these indeed are the right answers remains to be established in detailed experiments that parametrically vary the strength of a verb-based implicit causality cue, orthogonal to various other constraints. For example, to the extent that verb-based implicit causality is ‘just another cue constraining interpretation,’ suitable wider discourse context should be able to neutralize, and perhaps even reverse the direction of the bias (cf. Arnold, 2001), as a function of respective cue strengths.

In all, is there anything special about the semantic characteristics of verbs like *praise* and *apologize* that we denote as implicit causality? From a language processing perspective, we do not believe there is. Our findings show that people can exploit the semantic information encoded in implicit causality verbs very rapidly, whenever this information is relevant to ongoing comprehension. As such, implicit causality is on a par with a wide range of other sources of information that have been shown to affect the comprehension process very rapidly. We know that the meaning of verbs like *ate* and *told* allows listeners to anticipate specific upcoming verb arguments (e.g., Altmann & Kamide, 1999; Nieuwland & Van Berkum, in press). We also know that sentential as well as wider discourse context can support the prediction of specific upcoming words (Otten & Van Berkum, 2004; Van Berkum et al., 2005a; Wicha et al., 2004). In view of these phenomena, what would be so special about, say, anticipating a specific upcoming referent or pronoun after having read “*David praised Linda because...*”? After all, why would our expectations of where the speaker or writer might go next be informed by what can be buttered, eaten or talked to, but not by why people do the things they do?

Acknowledgments

We thank Natalia Waaijer, Iris Mulders, Frank Wijnen, Eric Reuland, Rein Cozijn, Pim Mak, Jesse Jansen, Bert Molenkamp, Gün Semin, Marcus Spaan, and Geert-Jan Mertens for their involvement in this research.

Appendix A

A.1. NPI-biased verbs

Consistent. Dirk en Maartje waren afgelopen zaterdag samen gaan winkelen, maar het werd niet echt een gezellige dag. Dirk verveelde Maartje omdat hij al vanaf het begin van de dag ontzettend liep te zeuren.

Inconsistent. Dirk en Maartje waren afgelopen zaterdag samen gaan winkelen, maar het werd niet echt een gezellige dag. Maartje verveelde Dirk omdat hij al vanaf het begin van de dag eigenlijk geen zin had.

Consistent. Thea en Paul reden allebei behoorlijk hard. Bij een druk kruispunt botsten zij met hun auto’s stevig op elkaar. Paul bood zijn excuses aan Thea aan omdat hij volgens de getuigen van het ongeluk alle schuld had.

Inconsistent. Thea en Paul reden allebei behoorlijk hard. Bij een druk kruispunt botsten zij met hun auto’s stevig op elkaar. Thea bood haar excuses aan Paul aan omdat hij volgens de getuigen van het ongeluk geen schuld had.

Consistent. Niels en Chantal hadden een klassieke knipperlichtrelatie. Intussen was het zo vaak aan en uit geweest, dat ze er allebei niets meer van snapten. Niels belde Chantal omdat hij nu toch eindelijk wel eens duidelijkheid wilde hebben.

Inconsistent. Niels en Chantal hadden een klassieke knipperlichtrelatie. Intussen was het zo vaak aan en uit geweest, dat ze er allebei niets meer van snapten. Chantal belde Niels omdat hij nu toch eindelijk wel eens duidelijkheid mocht geven.

Consistent. Meestal zijn de boswandelingen die Laura en Hans maken erg leuk, maar nu konden zij elkaar wel schieten. Hans ergerde Laura omdat hij bij elk plantje of beestje zo lang bleef kijken.

Inconsistent. Meestal zijn de boswandelingen die Laura en Hans maken erg leuk, maar nu konden zij elkaar wel schieten. Laura ergerde Hans omdat hij bij elk plantje of beestje op haar moest wachten.

Consistent. Eigenlijk was een scheiding de beste oplossing voor Freek en Mieke. Hun huwelijk was door het vele vreemdgaan namelijk een hel geworden. Freek kwelde Mieke omdat hij na de vele avontuurtjes nog steeds met andere vrouwen sliep.

Inconsistent. Eigenlijk was een scheiding de beste oplossing voor Freek en Mieke. Hun huwelijk was door het vele vreemdgaan namelijk een hel geworden. Mieke kwelde Freek omdat hij na de vele avontuurtjes nog steeds door haar bedrogen werd.

Consistent. Janine en Maarten waren toevallig allebei in de bieb. Zij zaten naast elkaar aan een tafel te studeren. Maarten stoorde Janine omdat hij tijdens het leren de hele tijd zeer irritant kuchte.

Inconsistent. Janine en Maarten waren toevallig allebei in de bieb. Zij zaten naast elkaar aan een tafel te studeren. Janine stoorde Maarten omdat hij tijdens het leren de hele tijd haar gekuch hoorde.

Consistent. Max en Paula discussieerden vaak over de wereldpolitiek. Bij de laatste discussie over de islam schrokken ze een beetje van elkaar. Max verontrustte Paula omdat hij niet eerder zulke discriminerende opmerkingen had gemaakt.

Inconsistent. Max en Paula discussieerden vaak over de wereldpolitiek. Bij de laatste discussie over de islam schrokken ze een beetje van elkaar. Paula verontrustte Max omdat hij niet eerder zulke discriminerende opmerkingen had gehoord.

Consistent. Anouk en Johan hadden te vaak tegen elkaar geschaakt op de schaakclub. Daarom was er tijdens het laatste spelletje ook geen sprake van spanning. Johan won van Anouk omdat hij na enkele tientallen potjes schaak haar spel goed kende.

Inconsistent. Anouk en Johan hadden te vaak tegen elkaar geschaakt op de schaakclub. Daarom was er tijdens het laatste spelletje ook geen sprake van spanning. Anouk won van Johan omdat hij na enkele tientallen potjes schaak haar niet kon verassen.

Consistent. Nico en Lotte vochten samen voor een eerlijke en gelijke wereld. Hun samenwerking bleek uitermate succesvol. Nico inspireerde Lotte omdat hij bij het zien van onrecht altijd direct actie wilde ondernemen.

Inconsistent. Nico en Lotte vochten samen voor een eerlijke en gelijke wereld. Hun samenwerking bleek uitermate succesvol. Lotte inspireerde Nico omdat hij bij het zien van onrecht altijd op haar kon rekenen.

Consistent. De collega's Maria en Karel hebben allebei zeer graag de touwtjes in handen. Laatst escaleerde een situatie op het werk tot een ware machtsstrijd. Karel intimideerde Maria omdat hij in de meeste gevallen dan zijn zin krijgt.

Inconsistent. De collega's Maria en Karel hebben allebei zeer graag de touwtjes in handen. Laatst escaleerde een situatie op het werk tot een ware machtsstrijd. Maria intimideerde Karel omdat hij in de meeste gevallen dan toch snel toegeeft.

Consistent. Daniel en Irma moesten samen in de werkgroep een referaat houden. Zij kregen helaas maar een week voor de voorbereiding. Daniel irriteerde Irma omdat hij op elke bespreking weer andere ideeën opperde.

Inconsistent. Daniel en Irma moesten samen in de werkgroep een referaat houden. Zij kregen helaas maar een week voor de voorbereiding. Irma irriteerde Daniel omdat hij op elke bespreking weer andere ideeën moest aanhoren.

Consistent. De relatie van Sabine en Boris had net een crisis doorstaan. Zij waren allebei vanaf het begin niet zo trouw geweest. Boris bekende alles aan Sabine omdat hij per se de waarheid wilde vertellen.

Inconsistent. De relatie van Sabine en Boris had net een crisis doorstaan. Zij waren allebei vanaf het begin niet zo trouw geweest. Sabine bekende alles aan Boris omdat hij per se de waarheid wilde horen.

Consistent. Marcel en Lisa hadden net weer een winkel overvallen. Zij waren nu onder een viaduct de buit aan het verdelen. Marcel lichtte Lisa op omdat hij de vorige keer een flinke som geld door haar was misgelopen.

Inconsistent. Marcel en Lisa hadden net weer een winkel overvallen. Zij waren nu onder een viaduct de buit aan het verdelen. Lisa lichtte Marcel op omdat hij de vorige keer een flinke som geld van haar had gestolen.

Consistent. De relatie tussen Hester en Gijs was na lang geruzie dan toch stukgelopen. Ook in de laatste ruzie konden zij niet eerlijk tegen elkaar zijn. Gijs loog tegen Hester omdat hij de werkelijke reden achter hun problemen te pijnlijk voor haar vond.

Inconsistent. De relatie tussen Hester en Gijs was na lang geruzie dan toch stukgelopen. Ook in de laatste ruzie konden zij niet eerlijk tegen elkaar zijn. Hester loog tegen Gijs omdat hij de werkelijke reden achter hun problemen te pijnlijk zou vinden.

Consistent. Bijna elk weekend gaan Willem en Merel wel naar een expositie. Maar bij het zien van de Nachtwacht bleek weer eens, dat zij sterk van mening verschilden. Willem verbaasde Merel omdat hij zo een negatieve reactie had gegeven.

Inconsistent. Bijna elk weekend gaan Willem en Merel wel naar een expositie. Maar bij het zien van de Nachtwacht bleek weer eens, dat zij sterk van mening verschilden. Merel verbaasde Willem omdat hij zo een negatieve reactie had gehoord.

Consistent. Tijdens een intense ruzie tussen Marcia en John liepen de emoties hoog op. Hun relatie dreigde nu echt op de klippen te lopen. John smeekte Marcia omdat hij ondanks de vele problemen toch veel van haar hield.

Inconsistent. Tijdens een intense ruzie tussen Marcia en John liepen de emoties hoog op. Hun relatie dreigde nu echt op de klippen te lopen. Marcia smeekte John omdat hij ondanks de vele problemen toch haar grote liefde was.

Consistent. Enthousiast waren Simon en Carien samen aan hun nieuwe opdracht begonnen. Inmiddels zijn zij een stuk minder blij met elkaar. Simon stelde Carien teleur omdat hij toch wel wat meer inzet had kunnen tonen.

Inconsistent. Enthousiast waren Simon en Carien samen aan hun nieuwe opdracht begonnen. Inmiddels zijn zij een stuk minder blij met elkaar. Carien stelde Simon teleur omdat hij toch wel wat meer inzet had willen zien.

Consistent. Roos en Harm hadden elk een groot en bloeiend bedrijf. Zij probeerden als elkaars grootste concurrenten elkaar steeds de loef af te steken. Harm misleidde Roos omdat hij anders de potentiële nieuwe opdrachtgever zou kwijtraken.

Inconsistent. Roos en Harm hadden elk een groot en bloeiend bedrijf. Zij probeerden als elkaars grootste concurrenten elkaar steeds de loef af te steken. Roos misleidde Harm omdat hij anders de potentiële nieuwe opdrachtgever zou wegkopen.

Consistent. Vanochtend waren Jasper en Rianne allebei weer te laat opgestaan voor het college. Zij stonden te dringen in de badkamer. Jasper hinderde Rianne omdat hij voor zijn ochtendritueel alle ruimte gebruikte.

Inconsistent. Vanochtend waren Jasper en Rianne allebei weer te laat opgestaan voor het college. Zij stonden te dringen in de badkamer. Rianne hinderde Jasper omdat hij voor zijn ochtendritueel alle ruimte nodig had.

Consistent. Op het feest werden Loes en Bart door wederzijdse vrienden aan elkaar voorgesteld. Zij vonden elkaar direct leuk en interessant. Bart fascineerde Loes omdat hij een persoon met een sterk karakter bleek te zijn.

Inconsistent. Op het feest werden Loes en Bart door wederzijdse vrienden aan elkaar voorgesteld. Zij vonden elkaar direct leuk en interessant. Loes fascineerde Bart omdat hij een persoon met een sterk karakter altijd interessant vond.

A.2. NP2-biased verbs

Consistent. Ruim vier jaar deden Joost en Karin onderzoek naar het taalvermogen van dolfinen. In die tijd kregen zij veel waardering voor elkaar. Karin respecteerde Joost omdat hij zowel theoretische als praktische vaardigheden in ruime mate bezat.

Inconsistent. Ruim vier jaar deden Joost en Karin onderzoek naar het taalvermogen van dolfinen. In die tijd kregen zij veel waardering voor elkaar. Joost respecteerde Karin omdat hij zowel theoretische als praktische vaardigheden in een persoon waardeerde.

Consistent. De boezemvrienden Anna en Lars zijn sinds kort ook burens van elkaar. Na een drukke periode hadden ze eindelijk tijd gevonden voor een dineetje. Anna bedankte Lars omdat hij de afgelopen weken bij de verhuizing zeer vaak onmisbaar was gebleken.

Inconsistent. De boezemvrienden Anna en Lars zijn sinds kort ook burens van elkaar. Na een drukke periode hadden ze eindelijk tijd gevonden voor een dineetje. Lars bedankte Anna omdat hij de afgelopen weken bij de verhuizing haar hulp zeer gewaardeerd had.

Consistent. Lex en Suzan waren van die personen die elkaar absoluut niet kunnen uitstaan. Elke keer dat zij elkaar tegenkwamen, liep het op ruzie uit. Suzan verafschuwde Lex omdat hij werkelijk altijd domme en botte opmerkingen tegen haar maakte.

Inconsistent. Lex en Suzan waren van die personen die elkaar absoluut niet kunnen uitstaan. Elke keer dat zij elkaar tegenkwamen, liep het op uit. Lex verafschuwde Suzan omdat hij werkelijk altijd domme en botte opmerkingen ruzie van haar hoorde.

Consistent. Els en Bas deden samen mee aan een vraaggesprek over de liefde. Toen hen gevraagd werd waarom ze van elkaar hielden, was het antwoord duidelijk. Els hield van Bas omdat hij altijd zichzelf en ontzettend vrolijk was in haar aanwezigheid.

Inconsistent. Els en Bas deden samen mee aan een vraaggesprek over de liefde. Toen hen gevraagd werd waarom ze van elkaar hielden, was het antwoord duidelijk. Bas hield van Els omdat hij altijd zichzelf en ontzettend vrolijk bij haar kon zijn.

Consistent. Guido en Heleen hadden hard gestudeerd voor een tentamen. Jammer genoeg haalden ze het allebei niet. Heleen troostte Guido omdat hij na het horen van de uitslag toch wel erg teleurgesteld had gereageerd.

Inconsistent. Guido en Heleen hadden hard gestudeerd voor een tentamen. Jammer genoeg haalden ze het allebei niet. Guido troostte Heleen omdat hij na het horen van de uitslag zag hoe moeilijk ze het had.

Consistent. Sofie en Martijn hielden van jongs af aan al van rollenspelletjes. Zij speelden elke week weer een andere situatie na. Sofie strafte Martijn omdat hij deze week een zeer vervelende schooljongen was.

Inconsistent. Sofie en Martijn hielden van jongs af aan al van rollenspelletjes. Zij speelden elke week weer een andere situatie na. Martijn strafte Sofie omdat hij deze week een zeer vervelende schoolmeester was.

Consistent. Sinds lange tijd waren Bob en Wendy werkzaam bij een advocatenbureau. Als zij samenwerkten vlogen de verwijten over en weer. Wendy bekritiseerde Bob omdat hij de hele tijd alle moeilijke beslissingen aan haar over liet.

Inconsistent. Sinds lange tijd waren Bob en Wendy werkzaam bij een advocatenbureau. Als zij samenwerkten vlogen de verwijten over en weer. Bob bekritiseerde Wendy omdat hij de hele tijd alle moeilijke beslissingen helemaal alleen moest nemen.

Consistent. Olga en Sander zaten in hetzelfde korfbalteam. Ze hadden zojuist met dit team een belangrijk toernooi gewonnen. Olga feliciteerde Sander omdat hij na zo'n goed gespeelde finale een schouderklopje zeker wel verdiende.

Inconsistent. Olga en Sander zaten in hetzelfde korfbalteam. Ze hadden zojuist met dit team een belangrijk toernooi gewonnen. Sander feliciteerde Olga omdat hij na zo'n goed gespeelde finale een schouderklopje wel terecht vond.

Consistent. Hoewel David en Mirjam al sinds lange tijd samen op pianoles zaten, hadden ze totaal geen waardering voor elkaar. Mirjam minachtte David omdat hij na al die jaren les nog steeds geen vooruitgang boekte.

Inconsistent. Hoewel David en Mirjam al sinds lange tijd samen op pianoles zaten, hadden ze totaal geen waardering voor elkaar. David minachtte Mirjam omdat hij na al die jaren les nog steeds geen vooruitgang zag.

Consistent. Afgelopen dinsdag vierden Nina en Vincent ieder hun elfde verjaardag. Ze hadden allebei om een nieuwe mountainbike gevraagd. Nina benijdde Vincent omdat hij de zo fel begeerde mountainbike wel had gekregen.

Inconsistent. Afgelopen dinsdag vierden Nina en Vincent ieder hun elfde verjaardag. Ze hadden allebei om een nieuwe mountainbike gevraagd. Vincent benijdde Nina omdat hij de zo fel begeerde mountainbike niet had gekregen.

Consistent. Luuk en Marije kwamen elkaar tegen in de studio's van Hilversum. Zij hadden beide auditie gedaan voor een soapserie. Marije prees Luuk omdat hij weer een zeer goede prestatie voor de camera had neergezet.

Inconsistent. Luuk en Marije kwamen elkaar tegen in de studio's van Hilversum. Zij hadden beide auditie gedaan voor een soapserie. Luuk prees Marije omdat hij weer een zeer goede prestatie van haar had gezien.

Consistent. Carlijn en Steven hadden vanuit hun studievereniging voor het eerst samen een congres georganiseerd. Na afloop moesten zij alleen nog de laatste rommel opruimen. Carlijn complimenteerde Steven omdat hij alles zo efficiënt en snel geregeld had.

Inconsistent. Carlijn en Steven hadden vanuit hun studievereniging voor het eerst samen een congres georganiseerd. Na afloop moesten zij alleen nog de laatste rommel opruimen. Steven complimenteerde Carlijn omdat hij alles zo efficiënt en snel geregeld vond.

Consistent. De rechtbank had het bedrijf van Michiel en Cindy failliet verklaard. Zij moesten nu hun schulden aflossen. Cindy stelde Michiel verantwoordelijk omdat hij het financieel beleid van begin af aan had tegengewerkt.

Inconsistent. De rechtbank had het bedrijf van Michiel en Cindy failliet verklaard. Zij moesten nu hun schulden aflossen. Michiel stelde Cindy verantwoordelijk omdat hij het financieel beleid van begin af aan had afgekeurd.

Consistent. Sandra en Hugo hadden een nogal verhitte geschiedenis vol ruzies en rijd. Zij konden elkaar inmiddels wel schieten. Sandra haatte Hugo omdat hij zo veel achterbakse en vuile streken had uitgehaald.

Inconsistent. Sandra en Hugo hadden een nogal verhitte geschiedenis vol ruzies en rijd. Zij konden elkaar inmiddels wel schieten. Hugo haatte Sandra omdat hij zo veel achterbakse en vuile streken had moeten incasseren.

Consistent. Guus en Zita werkten beiden op dezelfde afdeling van een bedrijf voor mobiele telefonie. Helaas, door slechte winstresultaten moest het bedrijf snijden in het personeel. Zita ontsloeg Guus omdat hij nu eenmaal niet genoeg vertrouwen van de leiding had.

Inconsistent. Guus en Zita werkten beiden op dezelfde afdeling van een bedrijf voor mobiele telefonie. Helaas, door slechte winstresultaten moest het bedrijf snijden in het personeel. Guus ontsloeg Zita omdat hij nu eenmaal niet genoeg vertrouwen in haar capaciteiten had.

Consistent. Op de Rietveld-academie waren Ellen en Frits elkaar voor het eerst tegen gekomen. Sindsdien deelden zij samen een atelier. Ellen bewonderde Frits omdat hij een groot talent voor beeldhouwen had laten zien.

Inconsistent. Op de Rietveld-academie waren Ellen en Frits elkaar voor het eerst tegen gekomen. Sindsdien deelden zij samen een atelier. Frits bewonderde Ellen omdat hij een groot talent voor beeldhouwen in haar had herkend.

Consistent. Tot een jaar geleden hadden Mark en Rachel samen voor een internetbedrijf gewerkt. Zij waren helaas op slechte voet uit elkaar gegaan. Rachel klaagde Mark aan omdat hij een grote som geld van het bedrijf verduisterd bleek te hebben.

Inconsistent. Tot een jaar geleden hadden Mark en Rachel samen voor een internetbedrijf gewerkt. Zij waren helaas op slechte voet uit elkaar gegaan. Mark klaagde Rachel aan omdat hij een grote som geld van het bedrijf tegoed bleek te hebben.

Consistent. Op de schermsschool hadden Tanja en Karl zich ingeschreven voor een wedstrijd. Omdat zij voor het eerst elkaars tegenstander waren, overwogen zij elkaars kansen. Tanja vreesde Karl omdat hij door vele jaren van ervaring meer kans had.

Inconsistent. Op de schermsschool hadden Tanja en Karl zich ingeschreven voor een wedstrijd. Omdat zij voor het eerst elkaars tegenstander waren, overwogen zij elkaars kansen. Karl vreesde Tanja omdat hij door vele jaren van ervaring haar klasse zag.

Consistent. Hercules en Athena woonden temidden van andere goden beiden op de Olympus. Zij kenden elkaar letterlijk al eeuwen. Athena aanbad Hercules omdat hij als een echte Griekse god zo vaak al haar wensen had vervuld.

Inconsistent. Hercules en Athena woonden temidden van andere goden beiden op de Olympus. Zij kenden elkaar letterlijk al eeuwen. Hercules aanbad Athena omdat hij als een echte Griekse god zo vaak viel voor de dochters van Zeus.

Consistent. Babet en Thijs waren al jaren verbonden aan hetzelfde natuurkundig instituut. Als team hadden zij al voor meerdere theoretische doorbraken gezorgd. Babet waardeerde Thijs omdat hij als wetenschapper altijd goede feedback aan haar gaf.

Inconsistent. Babet en Thijs waren al jaren verbonden aan hetzelfde natuurkundig instituut. Als team hadden zij al voor meerdere theoretische doorbraken gezorgd. Thijs waardeerde Babet omdat hij als wetenschapper altijd goede feedback van haar kreeg.

References

- Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247–264.
- Arnold, J. (1998). Reference form and discourse patterns. Unpublished doctoral dissertation, Stanford University, Stanford, CA.
- Arnold, F. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes*, 31, 137–162.
- Arnold, J., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eye tracking. *Cognition*, 76, B13–B26.
- Au, T. K. (1986). A verb is worth a thousand words: The causes and consequences of interpersonal events implicit in language. *Journal of Memory and Language*, 25, 104–122.
- Boland, J. E., Acker M. T., & Wagner L. (1998). *The use of gender features in the resolution of pronominal anaphora* (Cognitive Science Technical Reports No. #17). The Ohio State University Center for Cognitive Science.
- Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition: International Journal of Cognitive Science*, 14, 237–273.
- Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior*, 16, 601–609.
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12, 335–359.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.
- Ehrlich, K. (1980). Comprehension of pronouns. *The Quarterly Journal of Experimental Psychology*, 32, 247–255.
- Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Hove: Psychology Press.
- Garnham, A., Traxler, M., Oakhill, J., & Gernsbacher, M. A. (1996). The locus of implicit causality effects in comprehension. *Journal of Memory and Language*, 35, 517–543.
- Garvey, C., & Caramazza, A. (1974). Implicit causality in verbs. *Linguistic Inquiry*, 5, 459–464.
- Garvey, C., Caramazza, A., & Yates, J. (1975). Factors influencing assignment of pronoun antecedents. *Cognition*, 3, 227–243.
- Gerrig, R., & McKoon, G. (1998). The readiness is all: The functionality of memory-based text processing. *Discourse Processes*, 26, 67–86.
- Gordon, P. E., Hendrick, R., & Foster, K. L. (2000). Language comprehension and probe list memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 766–775.
- Greene, S. B., & McKoon, G. (1995). Telling something we can't know: Experimental approaches to verbs exhibiting implicit causality. *Psychological Science*, 6, 262–270.
- Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49, 133–156.
- Kutas, M., & Van Petten, C. K. (1994). Psycholinguistics electrified. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83–143). San Diego, CA: Academic Press.
- Long, D. L., & De Ley, L. (2000). Implicit causality and discourse focus: The interaction of text and reader characteristics in pronoun resolution. *Journal of Memory and Language*, 42, 545–570.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676–703.
- Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, 57, 203–220.
- McDonald, J. L., & MacWhinney, B. (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language*, 34, 543–566.
- McKoon, G., Greene, S. B., & Ratcliff, R. (1993). Discourse models, pronoun resolution, and the implicit causality of verbs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1040–1052.

- Millis, K. K., & Just, M. A. (1994). The influence of connectives on sentence comprehension. *Journal of Memory and Language*, 33, 128–147.
- Mitchell, D. M. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 15–32). New York: Psychology Press.
- Nieuwland, M. S., & Van Berkum, J. J. A. (in press). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*.
- Oakhill, J. V., Garnham, A., Reynolds, D., & Wilshire, C. (1998). Implicit causality effects in the interpretation of pronouns. In *Poster presented at the Eighth Annual Meeting of the Society for Text and Discourse*, Madison, Wisconsin, July 1998.
- Osterhout, L., & Holcomb, P. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785–806.
- Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739–773.
- Otten, M., & Van Berkum, J. J. A. (2004). *Discourse-based lexical anticipation during language processing: Prediction or priming?* Annual meeting of the Cognitive Neuroscience Society (CNS-2004), San Francisco, April 18–20. Supplement to the *Journal of Cognitive Neuroscience*, p. 68.
- Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with “The language-as-fixed-effect fallacy”: Common misconceptions and alternative solutions. *Journal of Memory and Language*, 41, 416–426.
- Rayner, K., & Sereno, S. C. (1994). Eye movements in reading: Psycholinguistic studies. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 57–81). San Diego, CA: Academic Press.
- Rudolph, U., & Försterling, F. (1997). The psychological causality implicit in verbs: A review. *Psychological Bulletin*, 121, 192–218.
- Semin, G. R., & Marsman, J. G. (1994). Multiple inference-inviting properties of interpersonal verbs: Event instigation, dispositional inference, and implicit causality. *Journal of Personality and Social Psychology*, 67, 836–849.
- Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, 9, 519–548.
- Stewart, A. J., Pickering, M. J., & Sanford, A. J. (2000). The time course of the influence of implicit causality information: Focusing versus integration accounts. *Journal of Memory and Language*, 42, 423–443.
- St. George, M., Mannes, S., & Hoffman, J. E. (1994). Global semantic expectancy and language comprehension. *Journal of Cognitive Neuroscience*, 6, 70–83.
- Tanenhaus, M. K., & Trueswell, C. (1995). Sentence comprehension. In J. L. Miller & P. D. Eimas (Eds.), *Speech, language, and communication* (pp. 217–262). San Diego, CA: Academic Press.
- Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives in sentence processing*. Lawrence Erlbaum.
- 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.
- Van Berkum, J. J. A. (2004). Sentence comprehension in a wider discourse: Can we use ERPs to keep track of things. In M. Carreiras & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension: Eye tracking, ERP and beyond*. Psychology Press.
- Van Berkum, J. J. A. (in preparation). The electrophysiology of discourse and conversation (working title). In M. Spivey, K. McRae, & M. Joannisse (Eds.), *The Cambridge Handbook of Psycholinguistics*. Cambridge University Press.
- Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999a). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, 41, 147–182.
- Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999b). When does gender constrain parsing? Evidence from ERPs. *Journal of Psycholinguistic Research*, 28(5), 555–571.
- Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005a). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, Cognition*, 31(3), 443–467.
- Van Berkum, J. J. A., Van den Brink, D., Tesink, C., Kos, M., Müller, O., & Hagoort, P. (2005b). That’s odd for you to say! An EEG study on how voice-inferred speaker characteristics affect language comprehension. In *9th International Conference for Cognitive Neuroscience (ICON-2005)*, Havana, September 5–10.
- Van Berkum, J. J. A., Zwitserlood, P., Bastiaansen, M. C. M., Brown, C. M., & Hagoort, P. (2004). So who’s “he” anyway? Differential ERP and ERSP effects of referential success, ambiguity and failure during spoken language comprehension. In *Annual meeting of the Cognitive Neuroscience Society (CNS-2004)*, San Francisco, April 18–20. Supplement to the *Journal of Cognitive Neuroscience*, p. 70.
- Van Berkum, J. J. A., Zwitserlood, M. C. M., Brown, C. M., & Hagoort, P. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research*, 17, 701–718.
- Van Gompel, R. P. G., & Liversedge, S. P. (2003). The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 128–139.
- Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, 16, 1272–1288.