Reading Sentences with a Late Closure Ambiguity: Does Semantic Information Help?

Sigrid Lipka

University College London, University of London & Laboratory of Experimental

Psychology, University of Sussex, Brighton

Short Title: READING LATE CLOSURE SENTENCES

Correspondence concerning this article should be addressed to Sigrid Lipka, who is now at the Universität Leipzig, Institut für Linguistik, Abt. Psycholinguistik, Brühl 34-50, 04109 Leipzig, Germany. email: lipka@rz.uni-leipzig.de

Abstract

Stowe (1989) reported that semantic information eliminates garden paths in sentences with the direct-object vs. subject ambiguity, such as <u>Even before the police stopped the driver was very frightened</u>. Three experiments are presented which addressed some methodological problems in Stowe's study. Experiment 1, using a word-by-word, self-paced reading task with grammaticality judgements, manipulated animacy of the first subject noun while controlling for the plausibility of the transitive action. The results suggest that initial sentence analysis is not guided by animacy. Experiment 2 and 3, using the self-paced task with grammaticality judgements and eye-tracking, varied the plausibility of the direct-object nouns to test revision effects. Plausibility was found to facilitate revision without fully eliminating garden paths, in line with various revision models. The findings support the view of a sentence processing system relying heavily on syntactic information, with semantic information playing a weaker role both in initial analysis and during revision, thus supporting serial, syntax-first models and ranked-parallel models relying on structural criteria.

Reading Sentences with a Late Closure Ambiguity: Does Semantic Information Help?

The issue of whether semantic information influences early syntactic analysis has been studied with a variety of syntactic ambiguities and is at the centre of much debate (for reviews see e.g., Mitchell, 1994; Tanenhaus & Trueswell, 1995). This paper focuses on sentences in which a potential direct object turns out to be the subject of a following main clause, e.g., Before the police stopped the driver was getting nervous. Readers typically initially analyse the driver as a direct object. They experience processing difficulties, the garden-path effect (Bever, 1970), on the phrase was getting because it forces the driver to be re-analysed as the subject of the main clause (Ferreira & Henderson, 1991; Frazier & Rayner, 1982; Kennedy & Murray, 1984). Stowe (1989) reported that semantic information eliminates these garden paths and argued for an incremental interactive processing model allowing for early effects of non-syntactic information on syntactic analysis. However, the apparent effects of semantic information might have been due to methodological problems. Furthermore, her data do not mesh well with findings reported by Clifton (1993) and Pickering & Traxler (1998). Thus, three new studies investigated the effects of semantic information on initial syntactic analysis and on revision processes by replicating and elaborating Stowe (1989). Two studies using a word-by-word, self-paced reading task with grammaticality judgements and an eye-tracking study will show that semantic information has a limited effect: it does not guide initial sentence analysis and helps revision to a limited extent.

A number of models predict the strong preference in favour of the direct-object analysis but they differ markedly in their assumptions about the early influence of semantic information. The most influential serial, syntax-first model is the Garden Path model developed by Frazier and colleagues (Frazier, 1987; Frazier, 1989; Frazier & Rayner, 1982). It explains the preference as an instance of the Late Closure strategy, which says to "attach new items into the clause or phrase currently being processed (i.e., the phrase or clause postulated most recently)" (Frazier, 1987, p. 562), thus ensuring that at points of syntactic ambiguity the syntactically

simplest structure is represented. To build up the initial sentence analysis, only syntactic

information is processed. Non-syntactic types of information are assumed to play a role only at

a second stage of processing, when the initial syntactic analysis is filtered, evaluated or revised

(e.g., Rayner, Carlson, & Frazier, 1983). Hence, semantic information should not guide initial

sentence analysis but may help revision. There are other serial, syntax-first models which

explain initial structural preferences in different terms but which make the same predictions

about semantic information as the Garden Path model (Pickering, 1994; Pritchett, 1992). In

contrast to serial, syntax-first models, Ford et al. proposed a simple verb-preferences account,

whereby the processor adopts the analysis consistent with the most frequent subcategorization

of a verb (e.g., Ford, Bresnan, & Kaplan, 1982). There are variants of the lexical model which

assume lexical information to have an effect only at a later checking stage (e.g., Mitchell,

1989). Semantic effects are not easily predicted by this model since it does not specify how

sentence-specific semantic information can alter the lexically encoded frequencies.

In contrast to syntax-driven models, parallel models, such as constraint-satisfaction

models, reject the notion of an initial exclusive reliance on structural information (e.g.,

MacDonald, Pearlmutter, & Seidenberg, 1994). Competing analyses are ranked according to

various criteria, and garden paths result if the highest ranked alternative is incompatible with

later parts of the sentence. No distinction is made between an initial stage and a later reanalysis

stage; revision is the re-ranking of alternatives in response to the relative strengths and

availability of multiple constraints (MacDonald, 1994). The strong direct-object preference is

accounted for by a strong local bias based on the high frequency of occurrence of transitive

structures (Tanenhaus & Trueswell, 1995). Semantic effects are predicted if semantic

constraints are stronger than this local bias. It is difficult to distinguish empirically between

serial versus ranked-parallel models (see Frazier, 1998), especially if parallel models assume

structural criteria for ranking alternative analyses (see e.g., Gorrell, 1989) or invoke a

structural local bias. However, it is possible to test whether semantic information determines

initial processing, which is explicitly ruled out by syntax-based serial models, but which is, at least in principle, allowed by parallel models.

The empirical evidence on whether semantic information helps processing sentences with the direct-object vs. subject ambiguity is equivocal (as is also the case for other syntactic ambiguities, see e.g., Rayner, et al., 1983; Trueswell, Tanenhaus, & Garnsey, 1994). Stowe's (1989) first experiment tested whether semantic information available before the ambiguity occurs determines the initial syntactic analysis, directly challenging syntax-first models (for evidence on other types of non-syntactic information see e.g., Altmann, van Nice, Garnham, & Henstra, 1998; Ferreira & McClure, 1997; Lipka, Kopp, & Pechmann, 2000). Stowe manipulated the subject nouns of the subordinate clause which were either animate (police) or inanimate (truck). She reasoned that agents are obligatorily intentional (following Fillmore, 1968) and hence that the thematic role of agent is assigned by default to animate but not to inanimate subject nouns. Such assumptions about default thematic assignments are widely held (see e.g., Clifton, 1993; Ferreira & Clifton, 1986; Just & Carpenter, 1992; Tanenhaus & Trueswell, 1995) and have also been suggested to be operative in language acquisition (see discussions in e.g., Bowerman, 1990; Cromer, 1991). This kind of assumption will be dubbed the Strong Thematic Default Hypothesis. However, Schlesinger (1992) proposed that the agent role is not defined by necessary or sufficient features (such as animacy) but instead by features which are typically found in agents (see also Dowty, 1991; McClelland & Kawamoto, 1986; McRae, Ferretti, & Amyote, 1997). Stowe (1989) suggested that the assumed default thematic assignments have repercussions for initial syntactic analysis and she cleverly used causative/ergative verbs as a test. In their causative reading (e.g., The man stopped the car), such verbs assign two thematic roles, the Agent role to the subject-NP and the Theme role to the direct object NP. In their ergative reading (e.g., The man stopped suddenly), only one thematic role is assigned, namely the Theme role to the subject-NP. Stowe argued that an animate subject-NP in the ambiguous sentences is assigned the Agent role by default (Before

the police stopped ....), and hence the following NP (the driver) is attached as direct object because it can be assigned the remaining Theme role of the causative verb. An inanimate subject-NP, however, is assigned the Theme role by default (Before the truck stopped ....), and the following NP (the driver) cannot be attached as direct object since there are no more thematic roles to be assigned by the ergative verb. Hence, an alternative has to be found and the only one compatible with the grammar is to close the VP and to postulate a new sentence node to which the NP can be attached as a subject-NP. In this case, a following finite verb (was getting nervous) should not cause a garden path. This is exactly what Stowe (1989) reported. This would be strong evidence against syntax-first models because semantic information would have been shown to determine the initial syntactic analysis of the ambiguous constituent. However, a close analysis of Stowe's test sentences suggests that animacy was confounded with sentence plausibility. In the animate condition, subordinate clauses frequently described highly plausible, often stereotypical transitive actions (e.g., the police stopped the driver; the actor finished the last scene; the lecturer began the talk; the children rolled the ball; the owner opened the shop) whereas intransitive actions were often impossible in the intended ergative reading (e.g., his friend was cooking - cannibalism?). In the inanimate condition, however, transitive actions were often impossible (e.g., the pennies rolled the ball; the doors opened the shop; the dinner was cooking the meat) and intransitive actions very plausible. Note that this bias in plausibility is not a necessary consequence of the in/animacy of the subject noun. If Schlesinger is right, there should be sentences in which an inanimate noun is in fact the agent of a plausible transitive action. Experiment 1 used such sentences as a more valid test of whether animacy determines default thematic assignments and initial parsing decisions. If there is a default assignment of the Agent role to animate nouns, as the Strong Thematic Default Hypothesis predicts, garden paths should occur only in the animate condition. The alternative is that inanimate nouns can receive the agent role, too, as Schlesinger suggests, in

which case garden paths should occur in both the animate and the inanimate condition, provided that the plausibility of the transitive actions is controlled for.

In her second experiment, Stowe (1989) additionally manipulated semantic information at the site of the ambiguity by including nouns which were implausible as direct objects. This manipulation presupposes that the transitive analysis is adopted; hence, the issue is how much semantic information affects revision rather than whether it determines initial analysis. Does semantic information trigger full revision or does it facilitate syntactically triggered revision? Stowe varied animacy and plausibility ("When the police (or: truck) stopped the driver (or: silence) became very frightened (or: frightening)." Implausibility was indeed noticed immediately on the noun ("silence") but surprisingly not only in the animate but also in the inanimate condition. This is odd since Stowe assumed that no transitive analysis is attempted in the inanimate condition; hence the plausibility or otherwise of the noun as direct object should have been irrelevant. Quite possibly, Stowe is wrong on this point, or there was some confounding difficulty with the implausible nouns. In fact, about half were "more abstract than the corresponding plausible noun" (Stowe, 1989, p. 341). What about garden paths? As Stowe predicted, the disambiguating verb was easier in the implausible- than the plausible-object sentences when the subject nouns were animate. However, because there were no unambiguous sentences, there is no way of knowing whether garden paths were eliminated or merely reduced. Furthermore, given that plausibility might have been confounded with other factors, it is not clear which factor eliminated or reduced garden paths.

Stowe's results do not fit easily with two eye-tracking studies by Clifton (1993) where garden paths occurred even when the direct object was implausible. However, he manipulated plausibility across, not within, experiments, and the eye-tracking measures did not always produce converging results. Stowe's data also contrast with Pickering & Traxler (1998, Experiment 1) who found that implausibility helped revision but that even in the implausible condition, garden paths were evident in Total Reading Times but not First Pass Reading Times

and Regressions. This contradictory state of affairs clearly merits further investigation. Experiments 2 and 3 manipulated plausibility within experiments and, in contrast to Stowe's study, included unambiguous sentences both to measure garden paths and to test for any confounding difficulty of the implausible nouns. If plausibility information at the site of the ambiguity affects revision, garden paths should be eliminated or reduced. A further aim of Experiment 3 was to address the possible spill-over of processing difficulty from the implausible noun to the disambiguating verb region.

Experiment 1

Experiment 1 tested whether initial syntactic analysis of ambiguous sentences is determined by semantic information available before the ambiguity. This would be strong evidence against syntax-first models. The critical sentences contained the direct-object vs. subject ambiguity studied by Stowe (1989) but with plausibility of the transitive action controlled for. Table 1 gives an example stimulus set.

_____

Insert Table 1 about here

_____

In the ambiguous sentences, attaching the noun phrase (the wheel) as a direct object turns out to be incorrect when the second verb (came) is read. In the unambiguous sentences the direct-object analysis is blocked by an additional adverbial (very rapidly). This type of control condition is the same as Stowe's to ensure easier comparison. The crucial semantic manipulation occurred on the first subject noun which was either animate or inanimate (child; motor). The Strong Thematic Default Hypothesis predicts garden paths in the animate but not the inanimate condition because only animate nouns should be assigned the agent role and the direct-object analysis should be adopted. The alternative hypothesis, based on Schlesinger, as well as the Garden Path model, predicts garden paths in both conditions because even inanimate nouns should be assigned the agent role. In contrast to Stowe (1989, Experiment 1)

and Clifton (1993, Experiment 2), a pretest was run to ensure equal plausibility of the transitive action in both the animate and inanimate condition. Garden paths are assumed to be reflected in longer decision times on the second verb (came) and possibly on the next words in ambiguous as compared to unambiguous sentences. Furthermore, on the adverbial, the Strong Thematic Default Hypothesis predicts processing to be more difficult in the animate than the inanimate condition because in the animate condition a direct-object NP should be expected instead. No such difficulty is predicted by the alternative hypothesis according to which the adverbial is equally unexpected in both conditions. Finally, if the transitive analysis is adopted regardless of animacy, as predicted by the alternative hypothesis, the ambiguous NP should be equally easy in the animate and inanimate condition.

<div align="center">Pretest</div>

The aim was to establish that the transitive analysis described in the subordinate clauses of the ambiguous sentences was equally plausible in the animate and inanimate conditions.

Method of Pretest

Participants. Ten students of the University of London, all native speakers of English, volunteered to take part. None of them participated in any other experiment reported here.

Materials. There were thirty-six sentence pairs sixteen of which contained verbs from Stowe's set, and twenty a new set of causative/ergative verbs. The sentences were truncated versions of the sentences to be used in Experiment 1, e.g., The police stopped the driver and The truck stopped the driver.

Procedure. Each sentence pair was presented on a separate sheet of paper, making up a booklet of thirty-six pages. The pages were arranged in a different random order for each participant. Sentences with animate subject nouns were the first sentence of a pair as often as sentences with inanimate subject nouns. In a forced-choice paradigm, participants were asked to read both sentences and to tick the response box next to the sentence that "sounded more

natural" to them. They were asked to choose one sentence of the pair even when the two sentences sounded almost "equally OK" to them.

Results of Pretest

A sentence pair was defined as biased if eight or more participants had chosen the same sentence of a pair as the more plausible one. There were ten pairs with a bias for the sentence with the animate subject noun. These were not included in Experiment 1. Five pairs had a bias towards the sentence with the inanimate subject noun. They were included in Experiment 1 since they allow for a strict test of the Strong Thematic Default Hypothesis. Of the stimuli included in the main experiment, animate-subject sentences were chosen on average 42% of the time and inanimate-subject sentences 58%, which did not differ ($\underline{t}$ (23) = -1.83, n.s.).

<div align="center">Main Experiment</div>

Method

Participants. Forty-four students and members of staff of University College London received £1 to participate. They were normally-sighted native speakers of English.

Materials. Twenty-four critical sentence sets like the one shown in Table 1 were used (see Appendix A). They were derived from the sentences used in the pretest excluding those with a clear preference for the animate-subject sentence (two further items were excluded at random because the 2 x 2 design requires a set divisible by 4). In all sentences, at least three words followed the disambiguating verb. Seventy-four unambiguous fillers of various sentence structures were randomly intermixed with the critical sentences. Of the total ninety-eight items, fifty were ungrammatical, ending with a word which belonged to an incorrect syntactic category, or carried a morphological or an agreement error. Sentences did not continue after the ungrammatical word. Four of the critical items were ungrammatical late in the sentence, at least four words after the disambiguating verb. In thirty fillers, the ungrammatical word occurred late (more than nine words into the sentence), and in twenty fillers, early (before the

tenth word of the sentence). There were fourteen practice sentences (eight ungrammatical ones) representing the different types of items and ungrammaticalities used in the experiment.

Task. A cumulative, word-by-word self-paced moving-window paradigm with a grammaticality judgement on each word was used. Participants had to press one of two response keys to indicate whether the sentence up to and including the word they were currently reading was grammatical or not.

Design. The design was a fully factorial 2 x 2 repeated-measures design incorporating a Latin Square. The factors were Ambiguity (ambiguous, unambiguous) and Animacy (animate, inanimate). To ensure that participants saw only one sentence from each stimulus set, four stimulus lists were created and a Latin Square was used to assign sentences to the four conditions. Conditions were intermixed rather than blocked. Participants were assigned at random to a list. Each participant saw six items per condition, and each item was responded to by eleven participants per condition.

Procedure and Apparatus. Participants were tested individually in small, sound-attenuated cubicles. Presentation of instructions and stimuli, and recording of responses was controlled by Elonex PCs (with an adapted MEL program; see Schneider, 1988). The plus-key to the right of the numeric keypad was labelled the yes-key, and the minus-key just above it the no-key. Participants were seated in front of the computer, read the Instructions and did the practice trials. The instructions stated that "half of the sentences are ungrammatical, that is they are incorrect in any context one could think of". After the practice trials participants could ask questions about the procedure. The experimenter encouraged them to respond as fast and as accurately as possible and left the cubicle. Each trial started with the message "Press the YES-key to see the first word and then the YES- or NO-key", which was presented at the top of the screen and remained there till the end of a trial. Sentences were presented cumulatively one word at a time, on one or two horizontal lines a third of the way down the screen, in white letters on black background. After the response to the last word of a sentence (either a

grammatical word plus full stop, or an ungrammatical word), the sentence disappeared and a feedback message ("Well done!" or "Wrong!") flashed up for 1000 ms in the centre of the screen. Then the message "Press the space bar to begin the next trial" was shown at the top of the screen and remained there until the participant started the next trial. Following Stowe, explicit feedback was given only at the end of sentences. Participants had been informed that a sentence would only continue if the last word they had seen was grammatical.

Scoring regions. For purposes of data analysis, sentences were divided into four regions. The noun region consisted of the head of the ambiguous noun phrase (wheel), the verb region was the syntactically disambiguating verb (came), the postverb region was the word following it (loose) and the adverbial region was the first word of the adverbial (very).

Results

Only a few false negative decisions (3.2%) were made. In this and the following experiments, effects were analysed by subjects ($F1$) and by items ($F2$) for response times for correct decisions (see Table 2).

—————————————

Insert Table 2 about here

—————————————

Separate 2-way analyses of variance (2 levels of Animacy x 2 levels of Ambiguity) were performed on the data from each region. Data for the noun region and the postverb region were logtransformed first. Exactly the same pattern of results was found when the analyses of variance were run on the non-transformed data. To help compare data across regions, geometric means are reported for all regions in Table 2.

The adverbial, forcing the intransitive analysis in the unambiguous sentences, should be more difficult in the animate than the inanimate condition, according to the Strong Thematic Default Hypothesis. However, paired t tests did not reveal any significant difference between the two unambiguous conditions,[1] nor did Wilcoxon Matched-pairs Signed-ranks Tests reveal

a difference in the number of extremely long decisions.[2] The <u>noun</u> region of the two types of ambiguous sentences did not differ either, against the Strong Thematic Default Hypothesis, indicating that the transitive analysis is adopted equally easily, regardless of animacy. Unexpectedly, for unambiguous sentences, response times were higher in the animate than the inanimate condition ($ps < .01$, two-tailed).

The data from the <u>verb</u> region are of special interest because it is here that garden paths might show up. Reading times were longer for ambiguous than unambiguous sentences [$F1(1, 40) = 67.86$, $p < .001$; $F2(1, 20) = 60.50$, $p < .001$] and longer for sentences with animate rather than inanimate subject-nouns [$F1(1, 40) = 26.22$, $p < .001$; $F2(1, 20) = 22.29$, $p < .001$]. The interaction Ambiguity x Animacy was also significant [$F1(1, 40) = 18.07$, $p < .001$; $F2(1, 20) = 18.23$, $p < .001$]. Planned paired t tests indicated that the animate-ambiguous condition was more difficult than the inanimate-ambiguous condition ($ps < .001$, one-tailed). However, ambiguous sentences were read more slowly than unambiguous sentences in the by-subjects and the by-items analyses both in the animate condition ($t(43) = 6.96$, $p < .001$, one-tailed; $t(23) = 6.55$, $p < .001$, one-tailed) and in the inanimate condition ($t(43) = 4.37$, $p < .001$, one-tailed; $t(23) = 3.99$, $p < .001$, one-tailed). On the <u>postverb</u> region,[3] the analysis of variance revealed no significant response-time differences across conditions.

<u>Discussion</u>

Experiment 1 showed that garden paths occurred not only in sentences with animate subject nouns but also in sentences with inanimate subject nouns. Ambiguity did interact with Animacy, as predicted by the Strong Thematic Default Hypothesis, but this reflected the <u>extent</u> of the disruption - it was greater in the animate than the inanimate sentences - and not <u>whether</u> there was a disruption. These results contrast with Stowe (1989, Experiment 1) who found garden paths only in sentences with animate subject nouns, but not in those with inanimate subject nouns. The conflicting results cannot be due to differences in methodology or type of unambiguous condition since these were identical. Stowe's results probably indicate that her

manipulation of animacy was confounded with the plausibility of the transitive action. The current findings indicate that if this confounding factor is successfully ruled out, animacy information on its own does not determine initial syntactic analysis.

In the unambiguous sentences, the intransitive analysis enforced by the adverbial appeared to be more difficult in the animate than the inanimate condition, but this effect was delayed, showing up on the noun following the adverbial. This could be seen as some support for the Strong Thematic Default Hypothesis. Note, however, that in the ambiguous sentences, the transitive analysis was adopted even in the inanimate condition, contrary to the Strong Thematic Default Hypothesis.

These results are, partly, in line with Clifton (1993, Experiment 2), who also found garden paths in the inanimate condition which in fact were stronger than in Experiment 1 since most of his measures yielded Ambiguity main effects. In contrast to Clifton (1993, Experiment 2), Experiment 1 produced no evidence for slower reading times on the ambiguous noun in the inanimate compared to the animate ambiguous sentences. He interpreted this slowing down as reflecting the initiation of reanalysis triggered by the "unusual" assignment of the agent role to an inanimate subject and the "modest implausibility of the postverbal NP as direct object" (p. 239, 240) in his stimuli. However, in Experiment 1, thematic assignment was less unusual and NPs quite plausible, given the results of the pretest.

<div align="center">Experiment 2</div>

Experiment 2 tested to what extent semantic information available at the site of the ambiguity helps revision in sentences with the direct-object vs. subject ambiguity. A strong semantic cue against the preferred transitive analysis was provided by manipulating the plausibility of the potential direct-object noun. Previous studies with unambiguous sentences showed that readers immediately register semantic mismatches between a verb and its subcategorised nouns (e.g., Lipka, 1993; Rösler, Pütz, Friederici, & Hahne, 1993) but the evidence on how useful semantic mismatches are for syntactic revision is mixed. Clifton

(1993), Pickering and Traxler (1998, Experiment 1) and Stowe (1989, Experiment 2) all addressed this question but, as discussed in the Introduction, the findings were inconclusive.

_____

Insert Table 3 about here

_____

Table 3 gives an example item set. Ambiguous NPs were either semantically acceptable or anomalous as direct objects (dog vs. hat). All models predict an initial preference for the transitive analysis. If semantic information is processed rapidly, there should be a semantic anomaly effect on the noun, with longer response times for semantically anomalous ambiguous sentences compared to all other sentences. If semantic anomaly triggers full revision, the next word (attracts), forcing the alternative intransitive analysis, should be processed without difficulty. Garden paths should occur only for ambiguous sentences in which the transitive analysis is semantically acceptable.

Method

Participants. Forty normally-sighted students of University College London, all native speakers of English, received course credit for their participation.

Materials. There were twenty-four critical sentence sets (see Appendix B) like the one presented in Table 3 plus twenty-four fillers of various sentence structures. For the critical sentences, sixteen of the causative/ergative verbs used in Experiment 1 were used again plus eight new verbs (to ride, to walk, to drive, to break up, to fly, to develop, to return, to finish). Ambiguous nouns were matched for length (semantically acceptable: mean number of letters 5.2, SD = 1.6; anomalous: 5.5 and 1.9, respectively) and word frequency (the mean frequency per one million words for written texts, taken from the Celex data base (Version 2.5), was 104 (SE = 23) for the acceptable, and 143 (SE = 36) for the anomalous condition, which did not differ in an independent-samples t test). Several norming studies were carried out. These were paper-and-pencil rating tasks with different samples of native speakers, none of whom took

part in the main experiment or other norming studies. In each task, two lists were used to ensure that subjects only saw one version of each item set. Semantically anomalous nouns lacked a semantic feature which the verb required its direct object to have - they did not allow semantic "fusion" sensu Jackendoff (1987, p. 383). This was checked in a norming study in which ten participants rated truncated versions of the stimuli, i.e., simple declarative sentences describing transitive actions (e.g., 'Alice walks her funny old dog / hat'), on a 3-point scale from 1 ('OK in terms of meaning') to 3 ('anomalous in terms of meaning'). A paired t test confirmed that the two conditions differed significantly (semantically acceptable: M = 1.2; SD = .12; semantically anomalous: M = 2.8; SD = .27); t(23) = -24.92, p < .001). To ensure that semantically acceptable and anomalous sentences were equally natural in their unambiguous versions, ten participants rated, again on a scale of 1 to 3, whether a sentence was 'OK in terms of meaning'. The mean ratings for the two unambiguous versions were virtually identical (acceptable: 1.5; anomalous: 1.7) and did not differ in a t test. In all sentences, at least three words followed the disambiguating word. Half of the sentences were ungrammatical (see Experiment 1). Four of the critical sentences (one per condition) ended with an ungrammaticality which occurred six or more words after the disambiguating verb. In ten ungrammatical sentences, the error occurred within the first nine words of the sentence, and in fourteen sentences, the error occurred more than nine words into the sentence. Critical and filler items were mixed at random. There were ten practice items (five were ungrammatical) which were representative of the sentence structures used.

Task, Procedure and Apparatus: These were the same as in Experiment 1.

Design. A 2 x 2 fully factorial repeated-measures design incorporating a Latin Square was used. The factors were Ambiguity (ambiguous, unambiguous) and Semantic Status (semantically acceptable versus anomalous). Conditions were intermixed rather than blocked. Participants were assigned randomly to a list. Each of them read six items per condition, and each item was responded to by ten participants.

Scoring regions. The noun region was the head of the ambiguous noun phrase (dog or hat), the verb region was the syntactically disambiguating verb (attracts) and the postverb region the word following it.

Results

Response times for correct decisions and percentages of correct decisions on each region are presented in Table 4. Separate 2-way analyses of variance (2 levels of Semantic Status x 2 levels of Ambiguity) were performed on the logtransformed response times.

_____

Insert Table 4 about here

_____

Response times on the noun produced a semantic anomaly effect: there were significant main effects for Ambiguity [$F1(1, 36) = 10.40$, $p < .01$; $F2(1, 20) = 11.40$, $p < .01$] and Semantic Status [$F1(1, 36) = 29.90$, $p < .001$; $F2(1, 20) = 17.67$, $p < .001$], and a significant interaction Ambiguity x Semantic Status [$F1(1, 36) = 5.99$, $p < .025$; $F2(1, 20) = 10.78$, $p < .01$]. Planned paired t tests indicated that mean response times were significantly higher in the semantically anomalous ambiguous condition than either the semantically acceptable ambiguous or the semantically anomalous unambiguous conditions (ps $< .01$, one-tailed).

On the verb, response times were longer for ambiguous than unambiguous sentences [$F1(1, 36) = 92.51$, $p < .001$; $F2(1, 20) = 35.77$, $p < .001$] and longer for semantically acceptable than anomalous sentences [$F1(1, 36) = 22.16$, $p < .001$; $F2(1, 20) = 8.79$, $p < .01$]. The interaction of Ambiguity x Semantic Status was significant [$F1(1, 36) = 11.27$, $p < .01$; $F2(1, 20) = 13.86$, $p < .01$]. Planned paired t-tests demonstrated that the two ambiguous conditions differed significantly, with semantically acceptable sentences being more difficult than semantically anomalous ones (ps $< .001$, one-tailed). However, ambiguous sentences were more difficult than their unambiguous control sentences in the by-subjects and the by-items analyses for both the semantically acceptable condition ($t(39) = 7.73$, $p < .001$, one-tailed;

$t(23) = 6.79$, p < .001, one-tailed) and the semantically anomalous condition ($t(39) = 4.77$, p < .001, one-tailed; $t(23) = 3.38$, p < .01, one-tailed). On the postverb region, response times for correct decisions[4] did not differ across conditions.

Discussion

The semantic anomaly effect found in Experiment 2 is strong evidence for incremental processing and shows that semantic information is processed immediately, in line with Pickering & Traxler (1998). The effect cannot be due to any confounding difficulty with the nouns because they were read as fast as the semantically fitting nouns when the sentences were unambiguous. Experiment 2 also showed that semantic information against the transitive analysis clearly facilitated revision but did not eliminate garden paths.

One might object (G. T. M. Altmann, personal communication, October 1993) that the increased response times on the disambiguating word in the semantically anomalous sentences are not evidence for garden paths but instead reflect a spill-over of processing difficulty from the immediately preceding semantically anomalous noun. Spill-over might be particularly likely in cumulative self-paced paradigms which might invite subjects to 'pace through' the sentence (Ferreira & Henderson, 1990; Just, Carpenter, & Wooley, 1982). While some studies produced circumscribed effects (see Holmes, 1987, Experiment 2; Stowe, 1989), Boland, Tanenhaus, Garnsey, & Carlson (1995, Experiments 4 and 5) found plausibility effects in filler-gap sentences which started on the first possible word but continued on the following words. Note however that their items were globally rather than locally implausible, possibly inviting effects on more than one word. Experiment 3 helps to rule out the possibility of spill-over by adding an extra phrase after the noun. Note that Pickering & Traxler (1998) also included an extra postnoun region, which produced an implausibility effect for Regressions and Total Time data, leaving open the possibility that their weak garden path for implausible sentences in the Total Time data was due to spill-over.

Experiment 3

Experiment 3 was a replication of Experiment 2 with two modifications: an eye-tracking methodology was used to establish the exact time course of semantic effects and to offer a more realistic reading situation, and the potential spill-over confound in Experiment 2 was addressed. To soak up any spill-over, ambiguous NPs (my dogs / hats) were lengthened by a postmodifier (that look so funny). Table 5 presents an example stimulus set.

_____

Insert Table 5 about here

_____

The predictions about the semantic anomaly effect on the ambiguous noun and garden paths on the disambiguating verb are identical to Experiment 2. The added phrase allows us to test if the potential spill-over of the semantic anomaly effect has dissipated.

Method

Participants. Thirty-six normally-sighted students of the University of Sussex, all native speakers of English, were paid £6 an hour to participate.

Materials. Twenty-eight critical sentence sets as illustrated in Table 5 were constructed (see Appendix C). In twenty, the causative/ergative verbs were the same as in Experiment 2 but the sentences were reworded to make them more natural given the lengthening of the NP. NPs were lengthened by adding either a relative clause, a prepositional phrase or a reduced relative clause. The critical nouns were matched for length (mean number of letters was identical in the acceptable and anomalous versions, i.e., M = 4.5; SD = .51) and word frequency per one million words for written texts, according to Celex (acceptable: M = 105; SE = 23; anomalous: M = 168; SE = 65) which did not differ for the two versions as shown by an independent samples t test. Norming studies were conducted in the same way as for Experiment 2. The semantic anomaly of the implausible nouns was again confirmed in a rating study with ten participants rating truncated versions of the stimuli, e.g., 'I walk my dogs / hats that look so funny' (acceptable: M = 1.4; SD = .30; anomalous: M = 2.6; SD = .32; $t(27)$ =

-13.17, p < .001). The unambiguous versions were equally natural in the judgements of ten participants (acceptable: $\underline{M}$ = 1.6; anomalous: $\underline{M}$ = 1.7; paired $\underline{t}$ test: n.s.). There were seventy filler items of a variety of sentence structures. Half of the critical and filler trials were followed by a simple comprehension question, half of which required a Yes- and half a No/Can't tell response. Questions focused on different parts of the sentences. Critical and filler items were intermixed at random. There were eight unambiguous practice items of a variety of sentence structures, half of which were followed by a question.

Design. This was the same as in Experiment 2. In each condition, there were seven observations in the by-subjects design and nine observations in the by-items design.

Apparatus. The infrared limbus eye-tracking system at the Sussex eye lab (Opto-electronic Developments type 54) was used. It samples the horizontal signals every 5 ms. Viewing was binocular but the data from only one eye (generally the left) were stored. With readers sitting about 60 cm from the screen, the eye-tracker had a resolution of one character.

Procedure. Participants were tested individually in a session of about 60 to 90 minutes. Instruction booklets explained the calibration procedure and the general purpose of the study. Participants sat down in front of a PC-screen, head and chin rests were adjusted, a bite bar prepared and the 'glasses' measuring eye movements put on. During practice and experiment, calibrations taking about thirty seconds were made every two trials. To answer comprehension questions, two response boxes labelled Y and N were placed on the table in front of the readers so that their dominant hand could comfortably reach the Yes button and their non-dominant hand the No/Can't tell button. Each trial consisted of this sequence of events: two vertical bars, one on top of the other, were displayed on the left, half way down the screen to alert readers that a sentence was about to be presented. Readers pressed the Yes button when they were ready and a sentence was displayed horizontally on the screen, with the first letter of the first word in the space between the two bars of the alerting signal. After reading a sentence, readers pressed the Yes button when they thought that they had understood it. Then,

either a question or the alerting signal for the next trial was displayed, or a calibration was made. Readers could come off the bite bar any time but were encouraged not to do so during a trial.

Scoring regions. The NP region consisted of the determiner and head of the ambiguous noun phrase (my dogs), the spill-over region was the modifier following the noun phrase (that look so funny), the verb region was the syntactically disambiguating verb (gain), and the postverb region was the word following it (much).

Results

Separate 2-way analyses of variance (2 levels of Semantic Status x 2 levels of Ambiguity) were performed on the eye-tracking data for each region. The measures were First Pass Reading Times per character (the duration of all fixations in a region before the region is left in either direction), the probability of making a First Pass Regression (a regression to an earlier region made during the first pass through a region) and Total Reading Times per character (the duration of all fixations in a region, including fixations during the first sweep and refixations after the region had been left). The average proportion of fixating any of the regions during the first pass was 0.78; there were no systematic differences across experimental conditions. Occasions on which a region was not fixated at all during the first pass were treated as missing values in the analyses of the first-pass data. Exactly the same pattern of results was obtained when these occasions were included in the analyses as real zeros. On one occasion, all datapoints for a subject in a condition were missing and a replacement value was calculated with the Winer (1962) formula as the mean of the subject-means across all conditions, plus the condition mean minus the grand mean. Table 6 summarizes the eye-tracking data by region and condition.

_____

Insert Table 6 about here

_____

Only First Pass Regressions produced evidence for the semantic anomaly effect in the NP region. There was a significant Ambiguity x Semantic Status interaction ($\underline{F}1(1, 35) = 8.08$, $\underline{p} < .01$; $\underline{F}2(1, 27) = 5.14$, $\underline{p} < .05$) and planned comparisons revealed that for ambiguous sentences, probabilities of First Pass Regressions were (marginally) higher for semantically anomalous than semantically acceptable sentences ($\underline{F}1(1, 35) = 4.49$, $\underline{p} < .05$; $\underline{F}2(1, 27) = 3.35$, $\underline{p} < .10$).

In the spill-over region, there was no semantic anomaly effect for any measures. First Pass Reading Times unexpectedly produced a significant main effect for Semantic Status ($\underline{F}1(1, 35) = 13.32$, $p < .001$; $\underline{F}2(1, 27) = 16.51$, $\underline{p} < .001$) due to longer reading times for semantically acceptable compared to anomalous sentences. First Pass Regressions also produced a main effect for Semantic Status ($\underline{F}1(1, 35) = 16.36$, $p < .001$; $\underline{F}2(1, 27) = 10.61$, $\underline{p} < .01$) but it was due to anomalous sentences leading to more regressions than semantically acceptable sentences. For neither of these two measures was the interaction significant. Planned comparisons confirmed that these effects for Semantic Status also held for the unambiguous control sentences, albeit marginally for First Pass Regressions.

Did semantic implausibility eliminate garden paths? On the disambiguating verb, there were main effects, sometimes marginal, for Ambiguity (for Total Reading Times: $\underline{F}1(1, 35) = 41.54$, $\underline{p} < .0001$; $\underline{F}2(1, 27) = 16.24$, $\underline{p} < .001$; for First Pass Reading Times: $\underline{F}1(1, 35) = 7.58$, $\underline{p} < .01$; $\underline{F}2(1, 27) = 3.25$, $\underline{p} < .10$; for First Pass Regressions: ($\underline{F}1(1, 35) = 5.67$, $\underline{p} < .025$; $\underline{F}2(1, 27) = 3.57$, $\underline{p} < .10$) and main effects for Semantic Status (for Total Reading Times: $\underline{F}1(1, 35) = 7.23$, $\underline{p} < .025$; $\underline{F}2(1, 27) = 4.70$, $\underline{p} < .05$; for First Pass Regressions $\underline{F}1(1, 35) = 2.90$, $\underline{p} < .10$; $\underline{F}2(1, 27) = 5.45$, $\underline{p} < .05$), but the interaction was significant only for Total Reading Times ($\underline{F}1(1, 35) = 5.57$, $\underline{p} < .05$; $\underline{F}2(1, 27) = 7.22$, $\underline{p} < .025$). Planned comparisons revealed that the ambiguous conditions differed significantly, with greater difficulty in semantically acceptable than semantically anomalous sentences for Total Reading Times ($\underline{F}1(1, 35) = 7.38$, $\underline{p} < .025$; $\underline{F}2(1, 27) = 9.70$, $\underline{p} < .05$) and First Pass Regressions ($\underline{F}1(1, 35) = 5.15$,

$p < .05$; $F2(1, 27) = 6.92$, $p < .025$) but not for First Pass Reading Times, where the two ambiguous conditions produced virtually identical results. In the semantically acceptable condition, ambiguous sentences were more difficult than unambiguous sentences for all measures (for First Pass Regressions: $F1(1, 35) = 5.47$, $p < .05$; $F2(1, 27) = 5.57$, $p < .05$; marginal for First Pass Reading times: $F1(1, 35) = 4.52$, $p < .05$; $F2(1, 27) = 3.38$, $p < .10$; for Total Reading Times: $F1(1, 35) = 30.27$, $p < .0001$; $F2(1, 27) = 20.96$, $p < .001$). In the semantically anomalous condition, ambiguous sentences were (marginally) more difficult than unambiguous sentences only for Total Reading Times ($F1(1, 35) = 4.38$, $p < .05$; $F2(1, 27) = 3.16$, $p < .10$). On the postverb region, there was an Ambiguity main effect for Total Reading Times ($F1(1, 35) = 23.97$, $p < .001$; $F2(1, 27) = 14.46$, $p < .001$) and marginally for First Pass Regressions ($F1(1, 35) = 6.20$, $p < .05$; $F2(1, 27) = 3.86$, $p < .10$). Planned comparisons confirmed that Total Reading Times were significantly higher for ambiguous sentences than their unambiguous counterparts both in the semantically acceptable condition ($F1(1, 35) = 6.36$, $p < .05$; $F2(1, 27) = 4.27$, $p < .05$) and in the semantically anomalous condition ($F1(1, 35) = 29.98$, $p < .0001$; $F2(1, 27) = 11.51$, $p < .01$). Finally, Total Reading Times produced Ambiguity main effects for the NP- and spill-over region, probably indicating that these regions were re-read after garden paths occurred on the disambiguating regions.

Discussion

When the transitive analysis was semantically anomalous rather than acceptable, disruption on the syntactically disambiguating verb was less severe as shown by Total Reading Times and First Pass Regressions but not First Pass Reading Times,[5] suggesting that semantic information did not always facilitate revision. Total Reading Times[6] on the verb and the postverb region provided evidence for garden paths in anomalous sentences. This cannot be due to a spill-over of the semantic anomaly effect because on the added postnoun phrase there was no specific difficulty for the ambiguous-anomalous sentences. Thus, the overall pattern of

results replicates Experiment 2, which also found that semantic information helped revision but did not completely eliminate garden paths.

Why did the semantic anomaly effect only show up in First Pass Regressions? Pickering & Traxler (1998, Experiments 1 and 2) found similar patterns, but Clifton (1993, Experiment 1) reported effects for First Pass Times and Gaze durations. The number of implausible items was quite similar across these studies and thus cannot account for the discrepant results. Pickering & Traxler (1998) suggest that the stronger the implausibility, the more likely a reader is to stop looking at the implausible word (trying to find an interpretation) and to initiate a regressive eye movement instead. It is possible that in Clifton's materials the implausibility was less striking, given that in half of his implausible stimuli the subordinate clause had an inanimate subject, which might have made the transitive analysis less compelling. Effects on implausible direct-object nouns might also depend on the verb being biased towards taking a direct-object noun, as Garnsey et al. reported for sentences with direct-object versus sentence complement ambiguities (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; but see Pickering, Traxler, & Crocker, 2000). Clearly, the experimental task also affects the strength of the semantic anomaly effect: the self-paced reading task with grammaticality judgements used in Experiment 2 produced a strong effect, probably because it encourages fuller semantic processing on each word.

<div align="center">General Discussion</div>

Experiment 1 showed that semantic information, available before the ambiguous constituent, reduced but did not prevent garden paths in sentences with the direct-object vs. subject ambiguity. Experiments 2 and 3 demonstrated that semantic information available once the ambiguous constituent has been attached facilitated revision but did not completely eliminate garden paths. The results of Experiment 1 rule out initial parsing models guided by animacy, such as Stowe's (1989), and support the serial, syntax-first Garden Path model, and ranked-parallel models relying on structural criteria for ranking the alternatives. What about the

finding that garden paths were reduced in inanimate-subject sentences? It is possible but unlikely that the reduction was due to semantic information determining the initial analysis on a few trials. All stimuli were carefully selected to allow the transitive analysis, hence there is no reason why inanimacy should have had an effect on some trials but not others. Furthermore, if inanimacy had forced early closure on some trials, the ambiguous NP, which does not fit this analysis, should have been difficult to process, but there was no evidence for that. Instead, the reduction of garden paths suggests that the parser uses animacy information during revision, which is compatible with both serial Garden-Path type models and parallel models. In all three experiments, the effect of semantic information during revision was limited as garden paths were reduced but not eliminated. This suggests that Pickering & Traxler (1998) and Clifton (1993) are probably basically right whereas Stowe's claim that semantic information can eliminate garden paths is questionable.

The experiments shed new light on various revision models but they were not designed to decide between them. In some revision models, only syntactic information triggers revision (e.g., Ferreira & Henderson, 1991; Fodor & Inoue, 1994), but in others, semantic information can be the trigger (Clifton, 1993; Rayner et al., 1983). If syntactic information is the trigger, semantic information can help either by making the initial analysis easier to give up or the alternative analysis easier to adopt. For both options, various mechanisms have been proposed. The initial analysis might be easier to give up if the sentence has not yet been semantically interpreted (Frazier & Rayner, 1982) or if the degree of semantic commitment is weak. Experiments 2 and 3 confirm Pickering & Traxler's (1998) proposal that semantic commitment is weak for implausible transitive sentences, and Experiment 1 produced new evidence suggesting that semantic commitment is weak for transitive sentences with inanimate-subject nouns, probably because the agent role is less usual for them. Ferreira & Henderson (1991) discuss how semantic information might make it easier to adopt the alternative analysis. They assume that a thematic processor receives from

the lexicon all thematic and subcat frames of a verb. Once the parser, for whatever reason, has attached the NP as direct object, the thematic processor assigns the Theme role to the head of the NP, thereby selecting the transitive frame and prompting the decay of the intransitive frame. They found reanalysis to be more difficult for long NPs in which the head noun is followed by several words as compared to short NPs (and in equally long NPs, for those NPs where the head occurs earlier in the phrase) because, assumedly, the thematic processor is committed to the initial analysis earlier and hence the alternative decays earlier and is less available for reanalysis (but see Sturt, Pickering, & Crocker, 1999). The current findings that semantic information reduced garden paths suggest that the thematic processor is slow to select the transitive frame if subject nouns are inanimate or if direct-object nouns are implausible, thus causing the intransitive frame to start decaying late and thereby making it more easily available for revision. It is also possible that the thematic processor commits less strongly to the transitive frame under these conditions, or that it abandons it on some trials, which would lead to a lesser decay of the alternative intransitive frame.

How do the current findings relate to the view that semantic information itself triggers revision? On this account one has to explain why reanalysis was not complete before the syntactically disambiguating verb occurred, or, in terms of constraint-based models, why semantic information was such a weak constraint. Maybe the parser reacted to the semantic error signal only on some trials. But given that the animacy manipulation in Experiment 1 was clear-cut, and that in Experiments 2 and 3 all NPs were clearly anomalous - as shown by the norming studies and the semantic anomaly effect - there is little reason why the parser did not always react. Possibly, the parser did not have enough time to complete revision. This is unlikely, too, since in Experiment 3, there was a whole phrase between the semantic error signal and the disambiguating verb. A more likely explanation for why semantic information did not trigger full revision is that it was not informative (sensu Fodor & Inoue, 1994) with respect to the correct alternative. In the current sentences, only the syntactic error signal (the

verb) carries a clue to the correct alternative: the verb forces the ambiguous NP to be reanalysed as a subject-NP. In contrast, semantic information, e.g., an implausible direct-object noun, may well serve as an error signal but, even if the parser were to accept it as a trigger to attempt reanalysis, it could not attach the NP anywhere else in the previously computed syntactic structure. Of course, even the syntactic error signal was not sufficiently informative to ensure cost-free revision since after all there were garden paths (for accounts of what makes revision difficult, see Fodor & Ferreira, 1998; Sturt et al., 1999).

To conclude, three experiments showed that semantic information had a limited effect in processing sentences with the direct-object vs. subject ambiguity: it did not determine initial analysis, and it facilitated revision rather than triggered full revision. This provides strong evidence for a sentence processing system in which syntactic information plays a stronger role than semantic information, as predicted by serial, syntax-first models and ranked-parallel models relying on structural criteria for ranking the alternatives.

References

Altmann, G. T. M., van Nice, K. Y., Garnham, A., & Henstra, J.-A. (1998). Late closure in context. Journal of Memory and Language, 38, 459-484.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), Cognition and the development of language (pp. 279-362). New York: Wiley.

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. Journal of Memory and Language, 34, 774-806.

Bowerman, M. (1990). Mapping thematic roles onto syntactic functions: Are children helped by innate linking rules? Linguistics, 28, 1253-1289.

Clifton, C., Jr. (1993). Thematic roles in sentence parsing. Canadian Journal of Experimental Psychology, 47 (2), 222-246.

Cromer, R. F. (1991). Language and thought in normal and handicapped children. Oxford: Basil Blackwell.

Dowty, D. (1991). Thematic proto-roles and argument selection. Language, 67 (3), 547- 619.

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

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

Ferreira, F., & Henderson, J. M. (1991). How is verb information used during syntactic parsing? In G. B. Simpson (Ed.), Understanding word and sentence (pp. 305-330). Amsterdam: North-Holland.

Ferreira, F., & McClure, K. K. (1997). Parsing of garden-path sentences with reciprocal verbs. Language and Cognitive Processes, 12 (2/3), 273-306.

Fillmore, C. (1968). The case for case. In E. Bach & R. T. Harms (Eds.), Universals in linguistic theory (pp. 1-90). New York: Holt, Rinehart & Winston.

Fodor, J. D., & Ferreira, F. (Eds.) (1998). Reanalysis in sentence processing. Dordrecht, The Netherlands: Kluwer.

Fodor, J. D., & Inoue, A. (1994). The diagnosis and cure of garden paths. Journal of Psycholinguistic Research, 23 (5), 407-434.

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

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), Attention and Performance XII: The psychology of reading (pp. 559-585). Hove: LEA.

Frazier, L. (1989). Against lexical generation of syntax. In W. Marslen-Wilson (Ed.), Lexical representation and process (pp. 505-528). Cambridge, MA: MIT Press.

Frazier, L. (1998). Getting there (slowly). Journal of Psycholinguistic Research, 27(2), 123-146.

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

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

Gorrell, P. (1989). Establishing the loci of serial and parallel effects in syntactic processing. Journal of Psycholinguistic Research, 18 (1), 61-73.

Holmes, V. M. (1987). Syntactic parsing: In search of the garden path. In M. Coltheart (Ed.), Attention and Performance XII: The psychology of reading (pp. 587-599). Hove: LEA.

Jackendoff, R. (1987). The status of thematic relations in linguistic theory. Linguistic Inquiry, 18 (3), 369-411.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. Psychology Review, 99(1), 122-149.

Just, M. A., Carpenter, P. A., & Woolley, J. (1982). Paradigms and processes in reading comprehension. Journal of Experimental Psychology: General, 111, 228-38.

Kennedy, A., & Murray, W. S. (1984). Inspection times for words in syntactically ambiguous sentences under three presentation conditions. Journal of Experimental Psychology: Human Perception and Performance, 10 (6), 833-849.

Lipka, S. (1993). Semantic processing in natural language understanding. Unpublished doctoral dissertation, University of London.

Lipka, S., Kopp, F., & Pechmann, T. (2000). Referential context effects for subject/object relative-clause ambiguities: The role of the Reading Span. In E. Schröger, A. Mecklinger, & A. D. Friederici (Eds.), Working on working memory. Leipzig: Leipzig University Press.

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

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

McClelland, J. L., & Kawamoto, A. H. (1986). Mechanisms of sentence processing: Assigning roles to constituents of sentences. In J. L. McClelland, D. E. Rumelhart, and the PDP Research Group (Eds.), Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2: Psychological and biological models (pp. 272-325). Cambridge, MA: MIT Press.

McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. Language and Cognitive Processes, 12 (2/3), 137-176.

Mitchell, D. C. (1989). Verb guidance and other lexical effects in parsing. Language and Cognitive Processes, 4 (3/4), SI 123-154.

Mitchell, D. C. (1994). Sentence parsing. In M. A. Gernsbacher (Ed.), Handbook of psycholinguistics (pp. 375-409). San Diego: Academic Press.

Pickering, M. J. (1994). Processing local and unbounded dependencies: A unified account. Journal of Psycholinguistic Research, 23(4), 323-352.

Pickering, M. J., & Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. Journal of Experimental Psychology: Learning, Memory, and Cognition, 24 (4), 940-961.

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

Pritchett, B. L. (1992). Grammatical competence and parsing performance. Chicago: The University of Chicago Press.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. Journal of Verbal Learning and Verbal Behavior, 22, 358-374.

Rösler, F., Pütz, P., Friederici, A.D., & Hahne, A. (1993). Event-related brain potentials while encountering semantic and syntactic constraint violations. Journal of Cognitive Neuroscience, 5 (3), 345-362.

Schlesinger, I. M. (1992). The experiencer as an agent. Journal Memory and Language, 31, 315-332.

Schneider, W. (1988). Micro Experimental Laboratory: An integrated system for IBM-PC compatibles. Behavior Research Methods, Instrumentation, and Computers, 20, 206-217.

Stowe, L. A. (1989). Thematic structures and sentence comprehension. In G. N. Carlson & M. K. Tanenhaus (Eds.), <u>Linguistic structures in language processing</u> (pp. 319-357). Dordrecht: Kluwer.

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

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

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. <u>Journal of Memory and Language, 33</u>, 285-318.

Winer, B. J. (1962). <u>Statistical principles in experimental design</u>. New York: McGraw-Hill.

Appendix A

Stimuli For Experiment 1

Note: An X preceding a word indicates that the sentence becomes ungrammatical at that point. Participants did not see the X. The first subject is animate, the second after the slash inanimate. The adverbials of the unambiguous conditions are given in square brackets.

1. After the guide/downpour had ended [in the afternoon] the walk was welcomed by everybody.

2. Before the butcher/theatre opened [in the evening] the door was jammed by the great number of Xwhere

3. When the toddler/van shifted [along the path] the bricks were scattered everywhere and they were in everybody's way.

4. While the skipper/wipers was/were drying [in the sun] the windscreen was being cleaned thoroughly.

5. Although the partners/company split up [after a quarrel] the property remained in joint ownership.

6. While his friend/the oven was warming up [in the kitchen] the meat was defrosting on the shelf.

7. Before the child/motor turned [very rapidly] the wheel came loose and frightened everybody.

8. After Dame Kiri/the duet concluded [triumphantly] the concert was praised by everybody.

9. While the lad/crane was swinging [in the air] the tyre fell to the ground.

10. Although the cheerful clowns/campaign had changed [beyond recognition] the mood remained the same for a while.

11. Before the captain/boat sailed [to America] the Atlantic ocean seemed very calm and the weather was good.

12. When the tiger/noise started [so violently] the panic spread quickly around the town.

13. Although the giant/drill was shaking [dangerously] the wall was not damaged at all.

14. As the man/his hand moved [carelessly] the cup fell to the floor and broke.

15. As the woman/rock sank [beneath the waves] the lifeboat was on its way back to the Xwithin

16. Because the boy/kiln was burning [fiercely] the pots were quickly filled to extinguish the Xthese

17. Because the girl/tree was growing [that summer] lots of fruit was consumed in the house.

18. After the clumsy maid/continuous draught had circulated [through the room] the dust took a long time to settle.

19. Although the delivery boy/fierce tempest swirled [in the streets] the newspapers did not fly away.

20. Although the farmer/water was running [fast] the mill burnt to the ground.

21. When the flock/landslides decreased [in size] the fertile land became more valuable than before Xthan

22. Although the driver/lorry bounced [on the road] the precious goods reached their destination on time.

23. When the technicians/engine had rotated [according to plan] the propellers were cleaned thoroughly in the factory.

24. Although the tourist/water was boiling [in the heat] the eggs were very cold and needed reheating.

Appendix B

Stimuli For Experiment 2

Note: An X preceding a word indicates that the sentence becomes ungrammatical at that point. Participants did not see the X. The first potential direct object noun is semantically acceptable, the second, after the slash, is semantically anomalous. The adverbials of the unambiguous conditions are given in square brackets.

1. After the teacher had finished [at lunchtime] the adventurous walk/boy was welcomed by everybody in the school which Xwhy

2. Before the butcher opened [in the morning] the new shop/path had been thoroughly cleaned by his assistant.

3. Before we had dried [in front of the fire] our damp clothes/ the strong rain had really begun to get on our nerves.

4. Although the partners split up [after a quarrel] their shared property/ history was enough reason to help them come back together again.

5. When his friend had warmed up [after a while] the meal/night was a real success as had been expected.

6. As the man turned [into the harbour] the enormous boat/noise took everybody completely by surprise.

7. After Dame Kiri had ended [triumphantly] the fantastic concert/audience was highly praised by the organisers.

8. While the girl was swinging [in the playground] her arms/thoughts were all over the place.

9. Although the cheerful clowns had changed [beyond recognition] the gloomy mood/sky remained the same for the rest of the day.

10. Before the captain sailed [to America] the Atlantic ocean/climate seemed very pleasant to him and the Xthem

11. When the workmen started [early in the morning] the very frightening noise/storm forced everybody to seek refuge.

12. Although the furious elephant was shaking [with rage] the strong trees/ wind hindered each one of his aggressive attacks.

13. As the children moved [carelessly] their new toys/lawn got very badly damaged by accident.

14. After the pirates sank [without a struggle] the armoured vessel/ the stormy weather moved on fast and soon reached the mainland.

15. Although the tourist was boiling [in the summer] the water/house remained surprisingly cool for a long time Xhours

16. Before the king rides [in the morning] his beautiful big horse/park is always looked after by his staff.

17.Wherever Alice walks [in the city] her funny old dog/hat attracts a lot of attention.

18. When the salesman was driving [down the road] the brand new car/tyre struck an old lady and hurt her slightly.

19. While the fire burnt [ferociously] a huge amount of petrol/water escaped from several big tanks.

20. As the girl had grown [in the past few months] her beautiful black hair/dress suited her better than ever before Xbefore

21. Although the couple had broken up [after several years] their family/children continued to be their joint responsibility.

22. As the pilot flew [during the storm] the shaky unreliable aeroplane/stewardess caused great anxiety amongst everybody.

23. Although the baby was developing [quite normally] an awful chronic disease caused him great discomfort and suffering.

24. After the head-teacher had returned [from holiday] the first important report/meeting was widely publicized in the school magazine.

Appendix C

Stimuli ForExperiment 3

Note: The first potential direct object noun is semantically acceptable, the second, after the slash, is semantically anomalous. The adverbials of the unambiguous conditions are given in square brackets.

1. As I was hiding [somewhere] the snake/rain that I had feared did not bother me.

2. When Dame Kiri finished [triumphantly] the opera/stage shown on TV was much praised.

3. As the baker closed down [last year] the shop/cash that his wife owned was some help.

4. As I was driving [slowly] the truck/pond owned by Jim came into view.

5. Before he dried [eventually] his hands/the winds that were very cold made him uneasy.

6. When I was sailing [yesterday] the boat/gull that moved so well came very close to me.

7. After I'd warmed up [for a while] the lunch/party that I'd hated was more enjoyable.

8. As Ann had grown [fast] her hair/coat that was so pretty made her look wonderful.

9. As the acrobat balanced [perilously] a sword/noise that was huge made him nervous.

10. When the fish-monger opened [last month] the shop/town near the sea did not look nice.

11. When she drowned [last night] the cats/shoes that she loved were not found.

12. When she was boiling [in the heat] the water/shade that was so cool made her feel good.

13. As we turned [slowly] the canoe/noise from which we'd escaped was soon gone.

14. As the steamship burnt [ferociously] the fuel/water from the tank ran down the hull.

15. While Sue was trying to swing [on a rope] her arms/eyes that were sore got even worse.

16. If the director returns [today] the memo/fear about safety will not be discussed openly.

17. When he stretched [clumsily] the scarf/shelf that he liked fell down to the floor.

18. If we break up [next month] our group/child that was so strong will not survive.

19. After the pirates sank [in the sea] the boat/storm from Siberia was gone quickly.

20. As the traffic slowed down [somewhat] the coach/noise from the hotel was more noticeable.

21. Before I rode [yesterday] my horse/coat that was quite old had not been prepared well.

22. After he had exercised [briefly] his legs/scar that had been sore hurt even more.

23. When the band started [eventually] the music/crowd that was too noisy was very annoying.

24. If we do split up [so soon] the land/time that we've shared will lose its value.

25. Wherever I walk [in the city] my dogs/hats that look so funny gain much attention.

26. As I hadn't stirred [all day] the soup/skirt that I was making was not ready on time.

27. Unless you move [to the side] the vase/lake that is really pretty will not be visible.

28. Though I was shaking [with fear] the woman/crowd next to me did not come to my help.

Author Note

Footnotes

[1] Further paired _t_ tests were carried out on subject- and item means for the adverbial in unambiguous sentences after calculating the data different ways (taking the means of subject- and item-medians, logtransforming decision times before taking means, adjusting decision times by word length, and calculating mean decision times per character for all words making up the adverbial), all of which failed to reach significance.

[2] Decisions were defined as extremely long when they were greater than 1321 ms which is the value that lies 1.5 times the interquartile range above the 75th percentile of the distribution of response times collapsed over conditions. This definition is sensitive to the spread of the middle 50% of the data without being affected by the highest values of the distribution which are likely to inflate the definition of a cut-off based on standard deviations.

[3] Response times for the postverb region were only analysed if there was no error on this region plus no error on the preceding regions, leading to the exclusion of 20% of the data in the animate-ambiguous condition, 4% in the animate-unambiguous condition, 8% in the inanimate-ambiguous condition and 1% in the inanimate-unambiguous condition.

[4] Response times for the postverb region were only analysed if there was no error on this region plus no error on the preceding regions, leading to the exclusion of 8.7% of the data in the semantically-acceptable ambiguous condition, 0.8% in the semantically-acceptable unambiguous condition, 4.6% in the semantically-anomalous ambiguous condition, and 1.7% in the semantically-anomalous unambiguous condition.

[5] nor First Fixation Durations, not reported here.

[6] and Total Number of Regressions, not reported here.

Table 1

Example Stimulus Set for Experiment 1

_____

Animate-ambiguous

Before the child turned the wheel came loose and frightened everybody.

Inanimate-ambiguous

Before the motor turned the wheel came loose and frightened everybody.

Animate-unambiguous

Before the child turned very rapidly the wheel came loose and frightened everybody.

Inanimate-unambiguous

Before the motor turned very rapidly the wheel came loose and frightened everybody.

_____

Table 2

Experiment 1: Mean Response Times for Correct Decisions (Geometric Means in
Milliseconds) with Percentage Correct "Grammatical" Decisions per Sentence Condition and
Region

	Animate subject		Inanimate subject	
Region	Ambiguous	Unambiguous	Ambiguous	Unambiguous
Adverbial		696 (1.04)		670 (1.03)
		96.2		98.9
Noun	742 (1.03)	792 (1.04)	729 (1.03)	658 (1.03)
	99.2	97	99.6	99.2
Verb	1260 (1.05)	704 (1.03)	803 (1.04)	640 (1.02)
	83	99.2	93	99.2
Postverb	728 (1.04)	765 (1.03)	688 (1.03)	731 (1.03)
	94	99.2	98	99.6

Note. Standard Errors (based on raw data) are shown in parentheses.

Table 3

Example Stimulus Set for Experiment 2

_____

Semantically acceptable and ambiguous

Wherever Alice walks her funny old dog attracts a lot of attention.

Semantically anomalous and ambiguous

Wherever Alice walks her funny old hat attracts a lot of attention.

Semantically acceptable and unambiguous

Wherever Alice walks in the city her funny old dog attracts a lot of attention.

Semantically anomalous and unambiguous

Wherever Alice walks in the city her funny old hat attracts a lot of attention.

_____

Table 4

Experiment 2: Mean Response Times for Correct Decisions (Geometric Means in Milliseconds) with Percentage Correct "Grammatical" Decisions per Sentence Condition and Region

Region	Semantically acceptable		Semantically anomalous	
	Ambiguous	Unambiguous	Ambiguous	Unambiguous
Noun	683 (1.03)	680 (1.03)	884 (1.04)	720 (1.03)
	100	99.6	97.5	99.2
Verb	1080 (1.05)	686 (1.02)	838 (1.04)	671 (1.02)
	93.3	99.6	98.3	99.6
Postverb	698 (1.04)	687 (1.03)	633 (1.03)	645 (1.02)
	97.5	99.6	98.3	99.2

Note. Standard Errors (based on raw data) are shown in parentheses.

Table 5

Example Stimulus Set for Experiment 3

Semantically acceptable and ambiguous

Wherever I walk my dogs that look so funny gain much attention.

Semantically anomalous and ambiguous

Wherever I walk my hats that look so funny gain much attention.

Semantically acceptable and unambiguous

Wherever I walk in the city my dogs that look so funny gain much attention.

Semantically anomalous and unambiguous

Wherever I walk in the city my hats that look so funny gain much attention.

Table 6

Experiment 3: First Pass Reading Times, First Pass Regressions and Total Reading Times by Region and Condition

	Semantically acceptable		Semantically anomalous	
Region	Ambiguous	Unambiguous	Ambiguous	Unambiguous
First Pass Reading Times (ms per character)				
NP	37.2 (1.1)	38.0 (1.2)	39.2 (1.2)	38.3 (1.2)
Spill-over	40.6 (1.3)	40.5 (1.4)	35.6 (1.2)	34.2 (1.1)
Verb	65.6 (2.9)	55.0 (2.0)	65.1 (2.3)	58.2 (2.5)
Postverb	59.6 (2.0)	61.1 (2.3)	60.7 (2.1)	57.4 (1.6)
Probability of First Pass Regression				
NP	0.137 (0.02)	0.203 (0.02)	0.199 (0.02)	0.155 (0.02)
Spill-over	0.136 (0.02)	0.144 (0.02)	0.269 (0.03)	0.220 (0.03)
Verb	0.313 (0.03)	0.209 (0.03)	0.215 (0.02)	0.189 (0.02)
Postverb	0.465 (0.03)	0.334 (0.03)	0.480 (0.03)	0.394 (0.03)
Total Reading Times (ms per character)				
NP	68.8 (3.1)	59.5 (2.2)	72.6 (3.0)	64.1 (3.0)
Spill-over	87.4 (4.0)	57.6 (1.9)	70.7 (2.8)	60.0 (2.7)
Verb	100.3 (5.3)	59.4 (3.2)	78.1 (4.7)	64.0 (3.7)
Postverb	73.8 (4.1)	60.0 (3.5)	76.4 (4.4)	53.8 (2.8)

Note. Standard Errors (based on raw data) are shown in parentheses.