

The Effect of Clausal and Thematic Domains on Left Branching Attachment Ambiguities

Patrick Sturt (patrick@psy.gla.ac.uk)

Human Communication Research Centre; 53 Hillhead Street
Glasgow G12 8QF, UK

Holly P. Branigan (holly@psy.gla.ac.uk)

Human Communication Research Centre; 53 Hillhead Street
Glasgow G12 8QF, UK

Yoko Matsumoto-Sturt (Y.M.Sturt@ed.ac.uk)

Centre for Japanese Studies; George Square, Edinburgh
Edinburgh EH8 9JX, UK

Abstract

Recent work has emphasised the importance of thematic domains in sentence processing. Two questionnaire studies examined whether thematic domains influence attachment of relative clauses to complex NPs in Japanese. The results suggest that definitions of thematic domains should be revised to cover left-branching structures, but do not support a distinction between domains associated with clauses and adpositional phrases.

Introduction

In this paper, we present two experiments which examine how postpositional and clausal domains affect the resolution of relative clause ambiguities in Japanese. Our two concerns are firstly to establish whether the notion of thematic domain is relevant to left-branching structures; and secondly to compare the effects of clausal and postpositional phrase thematic domains on relative clause attachments. Our results will clarify whether existing definitions of thematic domain must be modified to cover left-branching structures, and will also provide evidence to distinguish between current theories of thematic domain effects in parsing.

Relative Clause Attachment and Thematic Domains

This paper considers constructions that involve a relative clause which can attach to either one of two possible sites in a complex noun phrase, as in the English example shown in (1), from Cuetos and Mitchell (1988):

- (1) Somebody shot the servant of the actress who was on the balcony.

Several studies have shown that the interpretation of such ambiguities depends partly on the thematic status of the preposition in the complex noun phrase. In (1), the preposition *of* is commonly assumed not to assign a theta role, but appears for purely syntactic reasons (perhaps to assign case to the lower NP, or as a “reflex” of the lower NP’s inherent case (Chomsky, 1986)), mediating the assignment of a theta role from *servant* to *actress*. Cross-linguistically,

the eventual interpretation of ambiguities such as (1) tends to favour high attachment of the relative clause, where it is the *servant* rather than the *actress* who is interpreted as being on the balcony. This high attachment preference has been established in Spanish (Cuetos & Mitchell, 1988), French (Zagar, Pynte, & Rativeau, 1997), Dutch (Brybaert & Mitchell, 1996) and Italian (De Vincenzi & Job, 1995), though in English the preference seems to be much weaker, and if anything favours low attachment (i.e., with the *actress* on the balcony) (Cuetos & Mitchell, 1988). These findings contrast with similar examples where the preposition assigns a thematic role in its own right. For example, consider the following sentence (from De Vincenzi & Job, 1995).

- (2) Everybody admires the man with the daughter who began to sing an opera.

In (2), *with* is commonly assumed to assign a thematic role to *daughter*. Several studies have found that the high attachment preference is reduced for complex NPs containing a thematic preposition such as *with*, compared to similar examples containing a non-thematic preposition such as *of*, often resulting in a low attachment preference when a thematic preposition is used. This has been found in English (Traxler, Pickering, & Clifton, 1998), Italian (De Vincenzi & Job, 1995) and Spanish (Gilboy, Sopena, Clifton, Jr, & Frazier, 1995).

These facts can be explained by adopting the notion of a *thematic processing domain*. Construal Theory (Frazier & Clifton, 1996) defines the *Current Thematic Processing Domain* as “the extended maximal projection of the last theta assigner”. Construal Theory claims that modifiers, such as the relative clauses in (1) and (2), are associated with the current thematic processing domain. Thus any potential attachment site within the current thematic processing domain is accessible, but sites lying beyond the current thematic processing domain are inaccessible. Let us apply this definition to (1) and (2) at the point where the relative clause is first encountered. In (1), assuming that *servant* assigns a thematic role to *actress*, then the last theta assigner is *servant*. Hence the current thematic processing domain

corresponds to the complex NP *the servant of the actress*. Therefore, both *servant* and *actress* are available as heads of attachment sites for the relative clause. In (2), by contrast, the current thematic processing domain corresponds to the PP *with the daughter*, since *with* assigns a theta role, and is therefore the last theta assigner at the point that the relative clause is encountered. Therefore, the higher noun phrase headed by *man* is inaccessible as an attachment site of the relative clause, resulting in a low attachment preference.

Relative Clause Attachment and Left-Branching Structure

In head-final languages, where recursive structure is predominantly left-branching, the relative clause *precedes* both of the nouns which it may eventually modify. Consider the Japanese translation of (1), given in (3):

- (3) barukonii ni iru joyuu no mesitukai wo
 balcony LOC is actress GEN servant ACC
 dareka ga utta.
 somebody NOM shot.

This contrasts with predominantly right branching languages such as English or Italian, where both possible attachment sites have been read in the input at the point where the relative clause is attached.

The postposition *no* in (3) is similar to the English preposition *of*. We assume that, like *of*, it “transmits” a thematic role from the higher noun to the lower NP. Henceforth, we will call such postpositions *non-thematic* postpositions. Now consider a pair of complex NPs which differ only in their postposition. In (4a) below, the postposition is again the non-thematic postposition *no*; but in (4b) it is another postposition, *kara-no*, which we will assume *assigns* rather than *transmits* a thematic role to its complement. Henceforth, we will call such postpositions *thematic* postpositions.¹ Figure 1 shows the thematic structure of the two complex NPs in (4):

- (4) a. genjuumin ga odosita tankentai no
 natives NOM threatened expedition-force of
 taichou
 commander
 “The commander of the expedition force that the natives threatened”
 b. genjuumin ga odosita tankentai
 natives NOM threatened expedition-force
 kara-no taichou
 from commander
 “The commander from the expedition force that the natives threatened”

¹*Kara-no* is actually composed of two postpositions, *kara* and *no*.

Consider (4a) and (4b) in the light of the Construal definition of thematic domains. An on-line study of constructions similar to (4a) by Kamide and Mitchell (1997) suggests that the relative clause is initially attached to the first noun that becomes available (i.e. the *low* site). In (4), this noun is *tankentai* (“expedition force”). However, under the definition above, the Construal notion of thematic domain corresponds to the maximal projection of the last theta assigner. At the point where the relative clause is initially attached, the last theta assigner is actually the verb inside the relative clause itself. Hence Construal would predict the domain for attachment of the relative clause to be the relative clause itself, which clearly makes no sense. Therefore, the Construal definition cannot be applied to left-branching constructions such as (4). This suggests that the Construal notion of thematic domains would not apply in such cases.

However, in Sturt (1997) and Sturt and Crocker (1997), we develop a definition of thematic domains which is applicable in such cases. These domains are defined in terms of thematic nodes, which for the purposes of this paper can be assumed to correspond to the (extended) maximal projections of theta-assigners. The thematic domain for a node N corresponds to the subtree rooted at the thematic node which most immediately dominates N if such a dominating thematic node exists, and corresponds to the root of the entire tree otherwise (see Sturt (1997) and Sturt and Crocker (1997) for more precise definitions). In left-branching structures, the thematic domain for a node N may be established after N has been processed, as subsequent theta assigners are read in the input. Once established, such a thematic domain defines the set of alternative attachment sites to which N may be easily reanalyzed (see Figure 1).

Once the relative clause has been attached in the low site (Kamide & Mitchell, 1997), the maximal projection of a *thematic* postposition will demarcate the thematic domain for the relative clause. Hence reanalyzing the initially low-attached relative clause to the high site is predicted to be dispreferred in such cases, since it entails moving the relative clause outside its domain (4b). However, if the postposition is *non-thematic*, then the thematic domain for the relative clause will be established when the higher of the two nouns is read. In the case of (4a), this noun is *taichou* (“leader”). This noun will assign a theta role to the lower NP headed by *taikentai* (“expedition force”). Hence, the thematic domain will correspond to the entire complex NP headed by *taichou*, and it is predicted that, even if the relative clause is initially attached to the NP headed by the lower noun, it can subsequently be freely reanalyzed to the other possible attachment site, since it is within the same thematic domain. In fact, there is evidence that such reanalysis takes place when the postposition is non-thematic. In an off-line questionnaire, Kamide and Mitchell (1997)

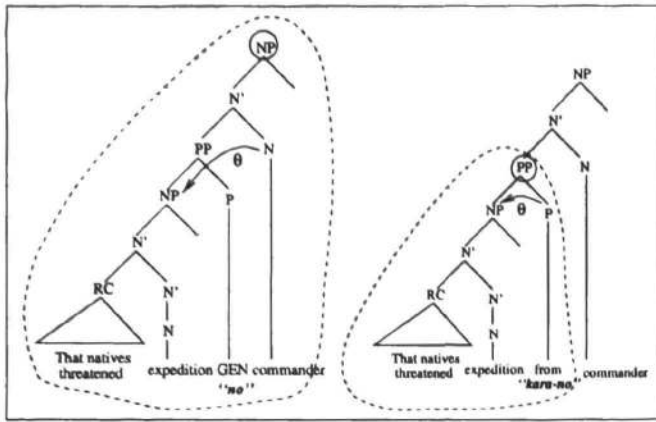


Figure 1: Syntactic structure of (4a) and (4b). The circles indicate thematic assigner nodes, and the dotted loop indicates the domain within which the relative clause can be easily reanalyzed

found that the final attachment preference for constructions similar to (4a) is for the high site, despite the fact that an on-line self paced reading study had found evidence for an initial attachment to the low site.

If thematic domains do have an effect on these ambiguities, then clearly the definition given in Construal should be revised along the lines of Sturt (1997) and Sturt and Crocker (1997), where thematic domains can be applied to left-branching structures. The purpose of Experiment 1 was to establish whether such an effect could be found.

Experiment 1

Method

Participants The participants were 24 native speakers of Japanese attending a course at the Park Language School in Sheffield.

Materials and Design The materials were 24 pairs of complex NPs involving a relative clause and two NPs, similar to the example given in (4). In each pair, the two possible attachment sites for the relative clause were separated by either a non-thematic postposition (4a) or a thematic postposition (4b); the two members of the pair differed only in the postposition separating the two noun attachment sites. To control for possible effects of animacy, half of the 24 items included an animate noun in the high site and an inanimate noun in the low site, and half involved the reverse configuration. We will call the former configuration "animate-high" and the latter "animate-low". We used 26 fillers, which consisted of NPs with various types of internal structure, and various types of ambiguity.

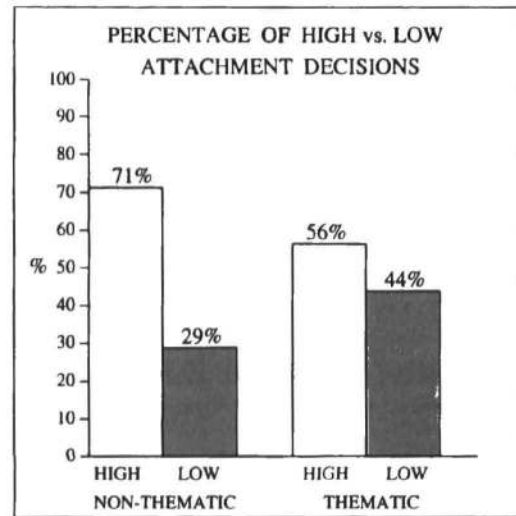


Figure 2: Results of Experiment 1

Procedure Participants saw individually randomized lists of materials, which were presented in printed booklets. The materials were rotated in a Latin Square design, so that each participant saw only one version of each item. The fillers were interspersed among the materials in a random order, and in such a way that no two experimental items appeared adjacent to each other. Underneath each experimental or filler item, two sentences were included, which indicated the two possible interpretations of the item. For example, in the item given in (4), the two possible interpretations were indicated by (Japanese translations of) the following:

- The natives threatened the commander.
- The natives threatened the expedition force.

Participants had to indicate whether a. or b. corresponded to their first interpretation. The sentences were balanced so that the a. sentence indicated the high attachment interpretation half the time and vice versa.

Results

The results are summarized in Figure 2. Overall, participants chose the high attachment interpretation more often than the low attachment interpretation. Collapsing over the two postposition types, we compared the proportion of high attachment decisions with 50% (i.e. the population mean expected on the null hypothesis). This revealed a significant preference ($t_1(23) = 3.67, p < .01$; $t_2(23) = 3.93, p < .001$). However, there were significantly more high-attachment decisions in the non-thematic condition than in the thematic condition ($t_1(23) = 3.03, p < .01$; $t_2(23) = 3.76, p < .01$).² The proportion of high

²The Wilcoxon Matched-Pairs Signed-Ranks Test yielded equivalent results for all pairwise comparisons in this paper.

attachment decisions for the non-thematic condition was significantly greater than 50% ($t_1(23) = 4.96, p < .001$; $t_2(23) = 6.66, p < .001$). In contrast, the proportion of high attachment decisions for the thematic condition was not significantly greater than 50% ($t_1(23) = 1.24, p > .1$; $t_2(23) = 1.24, p > .1$). This indicates that the overall high attachment preference is driven by the non-thematic condition.

Discussion

The study replicated the off-line high-attachment preference found by Kamide and Mitchell (1997) for non-thematic postpositions, but also demonstrated that postposition type has a reliable effect on attachment decisions in Japanese. This demonstrates that thematic domains are relevant to processing in left-branching constructions, and therefore the definition of thematic domain in Construal theory should be revised.

Our findings are explicable if we assume that reanalyzing a relative clause to a position outside its thematic domain is dispreferred in relation to a reanalysis which does not cross a domain boundary in this way. However, the fact that no low-attachment preference was observed for the thematic condition indicates that the processor does sometimes re-analyze across a thematic domain boundary, even in the absence of a syntactic or semantic cue.

Experiment 2

Experiment 1 demonstrated what we can call a “containing effect” for postpositional thematic domains: Assuming an initial low attachment, the presence of a thematic domain boundary reduces the chances of the relative clause being re-attached high. In Experiment 2, we compared such domains with clausal thematic domains. We might expect this containing effect to be “stronger” in cases where the thematic domain boundary is clausal than in cases where it is merely postpositional. This is the prediction of Gibson et al.’s (1996) *Predicate Proximity* principle, which specifies that an attachment is costly if it is not as close as possible to a *predicate phrase* in the tree. Gibson et al intend *predicate phrases* to correspond to verbal projections, but not prepositional or postpositional projections.³

Gibson et al also postulate a pure recency preference, under which an attachment is costly to the extent that it is made to non-recent material. An attachment which is favoured by both recency and predicate proximity will be preferred over an attachment that is favoured by only one of these principles. Now consider (4b) which includes a postpositional thematic domain boundary (between *kara-*

³This is clearly presupposed in their explanation for attachment preferences in Spanish and English 3-site relative clause ambiguities.

no and *taichou*), and compare it with (5), which includes a clausal thematic domain boundary (marked *S*). Note that (5) involves two relative clauses; the initial relative clause, whose attachment is ambiguous, and a second relative clause, which modifies the higher of the two attachment sites (see also Figure 3).

(5) [*RC* *genjuumin ga odosita*] *tankentai wo*
 natives NOM threatened expedition ACC
hikiita S] *taichou*
 led commander

“The commander who led the expedition force that the natives threatened”

In (4b), there is no predicate phrase outside the relative clause itself. In (5), by contrast, there is a predicate phrase headed by the verb *hikiita* (“led”). Therefore, the lower noun *tankentai* would be favoured as an attachment site for the relative clause by both recency and predicate proximity in (5), while this noun would be favoured only by recency in (4b). Hence, predicate proximity would predict more low attachments in (5) than in (4b).

Thus in the type of ambiguity under discussion here, predicate proximity effectively predicts that clausal domains, rather than thematic domains in general, will affect the attachment of the relative clause. Of course, this also means that predicate proximity cannot account for the results of Experiment 1, since the relevant domain there was postpositional rather than clausal. This could perhaps be remedied by invoking an extra principle, for example, *PP-proximity*. This extra principle would have to be weaker than predicate proximity, however, to account for the findings presented in Gibson et al. (1996). Hence, whether or not an extra principle is invoked, clausal domains would be predicted to have a stronger containing effect than postpositional thematic domains.

We can contrast this view with the Construal notion of thematic domains, which does not differentiate between clausal and pre/postpositional domains. Given that the Construal definitions could be amended to allow for left branching constructions, as discussed above, then clausal domains would be predicted to have a containing effect of equal strength to postpositional domains.

Experiment 2 aimed to distinguish between these two alternative accounts of domains. If the clause-based notion of domains implied by predicate proximity is correct, then there should be fewer high attachment conditions where a clause boundary separates the two possible noun attachment sites, as in (5), than where a postpositional thematic domain separates these nouns, as in (4b). However, if the more general notion of thematic domains implied by Construal is correct, there should be no difference between (5) and (4b); but both of these should differ from cases where no thematic domain boundary intervenes between the two

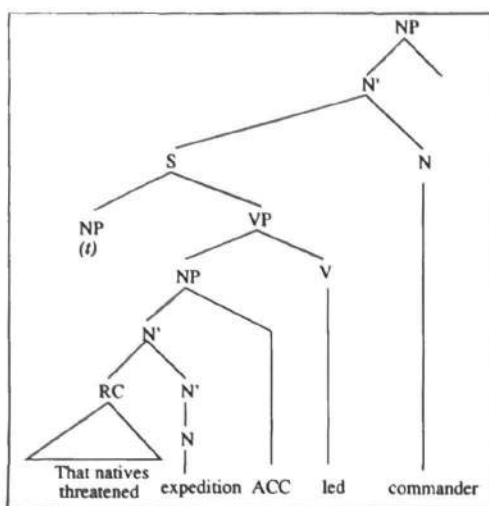


Figure 3: Syntactic structure of (5)

possible noun attachment sites, as in (4a). The use of clausal domains (and thus verbal theta assigners) also allowed us to generalize the findings of Experiment 1 to domains involving a wider range of thematic roles.

Method

Participants Participants were 24 Japanese native speakers who were attending a course at the Institute of Applied Language Studies at the University of Edinburgh.

Materials and Design The materials consisted of 24 triples of complex NPs. Each triple consisted of three conditions. The first two were a *non-thematic* condition and a *thematic condition*, which together correspond to the two conditions of Experiment 1; the third was a *clausal* condition, as in (5) (see also Figure 3).

Twelve items were animate-high and 12 were animate-low. In all the animate-high items in the clausal condition, the second relative clause was a *subject relative* (i.e., included a subject gap, coindexed with the (animate) higher noun). In all the animate-low items, the second relative clause was a *non-subject* relative (i.e., included a non-subject gap, coindexed with the (inanimate) higher noun). This was necessary because of the strong preference for transitive verbs to have animate subjects in Japanese.

The 24 experimental items were selected from a larger group of materials (including the materials from Experiment 1) in a pretest run at the University of Surugadai. This pretest ensured that the high and low interpretations were matched for plausibility (both t 's < 1). The materials for Experiment 1 were found to have had a slight plausibility bias towards the low attachment reading (the opposite direction to the high attachment preference actually found in experiment 1). Experiment 2 used the same fillers as Ex-

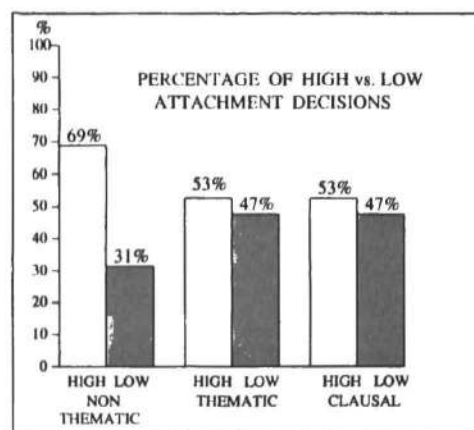


Figure 4: Results of Experiment 2

periment 1, and the procedure was identical.

Results

The results are summarized in Figure 4. Overall there was a preference for high attachment. t -tests comparing the overall proportion of high attachments with 50% revealed that this preference was significant by subjects but not by items ($t_1(23) = 2.5, p < .05; t_2(23) = 1.47, p = 0.16$). Planned contrasts revealed that there were significantly more high-attachment decisions in the non-thematic condition than in the thematic condition ($t_1(23) = 4.53, p < .001; t_2(23) = 3.82, p < .001$), and also that there were significantly more high attachment decisions in the non-thematic condition than in the clausal condition ($t_1(23) = 4.17, p < .001; t_2(23) = 2.81, p < .02$). However, there was no difference between the thematic and the clausal conditions—in fact the means were numerically identical ($t_1 = t_2 = 0, p = 1$).

We also found a significant effect of animacy in this experiment, with an overall high attachment preference for the animate-low items and an overall low attachment preference for the animate-high items. As this factor was not the focus of the experiment, and did not interact with the thematicity effect, we will not discuss it further here (though see Sturt (1997) for details).

Discussion

Experiment 2 replicated the finding of Experiment 1 that a thematic postposition neutralizes the high attachment preference found with the non-thematic postposition. However, Experiment 2 shows no evidence that a clausal thematic domain is treated any differently from a postpositional one. These results support (a left-branching version of) the Construal notion of thematic domains, but do not support the clause-based notion of domain implied by Predicate Proximity.

General Discussion

Our experiments demonstrate that altering thematic characteristics of a complex NP influences the eventual attachment preferences for relative clauses in Japanese. The effect is identical, whether the domain in question is clausal or postpositional. However, the presence of a thematic domain boundary does not *block* the high attachment of the relative clause.

We can exclude an alternative explanation of our results in terms of the relative phonological weight of the attaching constituent and its potential sister (Fodor, 1998; Hirose, Inoue, Fodor, & Bradley, 1997). The length of the complex NP to which the relative clause may attach differs between conditions in our materials (e.g. *N kara-no N* is longer than *N no N*). According to the phonological weight theory, this could have affected the attachment preferences in our materials. However, post-hoc analyses correlating the number of high attachment decisions with the length (in terms of the number of morae) of a. the relative clause, b. the complex NP, and c. the length of the relative clause minus the length of the complex NP, showed no significant correlations.

It is not clear how current theories can account for the overall high attachment preference that was found in both Experiments 1 and 2. One possible explanation could be some version of *Relativised Relevance* (Frazier, 1990), which predicts a preference for associating a modifier with the "main assertion of the sentence". We note that, as our study examined NPs, which are not sentences and do not denote "assertions", this principle would have to apply either to non-overt clausal structure or to the "main content" of the utterance, whatever its categorial expression or semantic type.

In conclusion, our results support the general notion of thematic domains given in *Construal* (Frazier & Clifton, 1996); but any such definition must be altered to account for left-branching structures such as those considered here. Our data argue against theories in which clausal domains are seen as stronger than pre/postpositional domains.

Acknowledgements

This research was supported by British Academy Postdoctoral Fellowships awarded to HB and PS, and an ESRC studentship awarded to PS. We thank Kuabrat Branigan, Masayo Iida, Ruth Linden, Jim Hutton, Peter Sells and Toyomi Takahashi.

References

- Brybaert, M., & Mitchell, D. C. (1996). Modifier attachment in sentence parsing: Evidence from dutch. *Quarterly Journal of Experimental Psychology, Section A—Human Experimental Psychology*, 49(3), 664–695.
- Chomsky, N. (1986). *Knowledge of language: its nature, origin and use*. New York: Praeger.
- Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in Spanish. *Cognition*, 30, 72–105.
- De Vincenzi, M., & Job, R. (1995). An investigation of late closure: the role of syntax, thematic structure and pragmatics in initial and final interpretation. *Journal of Experimental Psychology. Learning, Memory and Cognition*, 21(5), 1303–1321.
- Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, 27(2), 285–319.
- Frazier, L. (1990). Parsing modifiers: Special purpose routines in the HPSM? In D. A. Balota, G. B. F. d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 303–331). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Frazier, L., & Clifton, C. (1996). *Construal*. Cambridge MA: MIT Press.
- Gibson, E., Pearlmutter, N., Canesco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, 59(1), 23–59.
- Gilboy, E., Sopena, J. M., Clifton, Jr, C., & Frazier, L. (1995). Argument structure and association preferences in Spanish and English compound NPs. *Cognition*, 54, 131–167.
- Hirose, Y., Inoue, A., Fodor, J. D., & Bradley, D. (1997, March). *Adjunct ambiguity in Japanese: The role of constituent weight*. Poster presented at the 11th Annual CUNY Conference on Human Sentence Processing, New Brunswick, NJ.
- Kamide, Y., & Mitchell, D. C. (1997). Relative clause attachment: Nondeterminism in Japanese parsing. *Journal of Psycholinguistic Research*, 26(2), 247–254.
- Sturt, P. (1997). *Syntactic reanalysis in human language processing*. Unpublished doctoral dissertation, Centre for Cognitive Science, University of Edinburgh, Edinburgh, Scotland.
- Sturt, P., & Crocker, M. W. (1997). Thematic monotonicity. *Journal of Psycholinguistic Research*, 26(3), 297–322.
- Traxler, M. J., Pickering, M. J., & Clifton, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, 39, 558–592.
- Zagar, D., Pynte, J., & Rativeau, S. (1997). Evidence for early-closure attachment on first-pass reading times in french. *Quarterly Journal of Experimental Psychology, Section A—Human Experimental Psychology*, 50(2), 421–438.