Relative clause avoidance: Evidence for a structural parsing principle ☆

Adrian Staub [a,*], Francesca Foppolo [b], Caterina Donati [c], Carlo Cecchetto [b,d]

[a] Department of Psychological and Brain Sciences, University of Massachusetts Amherst, United States
[b] Department of Psychology, University of Milan-Bicocca, Italy
[c] Laboratoire de Linguistique Formelle, Université Paris 7/CNRS, France
[d] Structures Formelles du Langage, Université Paris 8/CNRS, France

ARTICLE INFO

ABSTRACT

Three eye movement experiments investigated the processing of the syntactic ambiguity in strings such as *the information that the health department provided*, where the *that*-clause can be either a relative clause (RC) or the start of a nominal complement clause (CC; *the information that the health department provided a cure*). The experiments tested the prediction that comprehenders should avoid the RC analysis because it involves an unforced filler-gap dependency. Readers showed difficulty upon disambiguation toward the RC analysis, and showed facilitated processing of the ambiguous material itself when the CC analysis was available; both patterns suggest rapid initial adoption of the CC analysis in preference to the RC analysis. The strength of the bias of a specific head noun (e.g., *information*) to appear with a CC did not modulate these effects, nor were these effects reliably modulated by the tendency of an ambiguous string to be completed off-line as a CC or an RC. These results add to the evidence that structural principles guide the processing of filler-gap dependencies.

Introduction

The comprehension of *filler-gap* dependencies has long been a central topic in sentence processing research (e.g., Clifton & Frazier, 1989; Fodor, 1978). These are structures in which an element (the *filler*, in psycholinguistic terminology) is displaced with respect to the location at which it receives its thematic role (the *gap*). In syntactic theories that posit movement, the filler is analyzed as having moved from the gap site, which contains a movement trace (e.g. Chomsky, 1981). In (1) and (2), filler-gap dependencies appear in the context of a *wh*-question and a relative clause, respectively.

(1) *Which dog* did the family choose _____ ?
(2) The family chose the dog *that* they visited _____ on Wednesday.

The great majority of research on this topic has focused on processing of structures in which the filler appears to the left of the corresponding gap site (as in (1) and (2)), and has addressed questions about how, in the course of incremental processing, the parser identifies the site of the gap once a filler has been identified. One important conclusion is that incremental parsing appears to respect grammatical constraints as to the location of the gap; the parser does not posit a gap within a syntactic 'island' (Ross, 1967), namely a syntactic domain (e.g., adjunct clauses) where a gap cannot occur (e.g., Phillips, 2006; Traxler & Pickering, 1996). A second important conclusion, however, is that the parser appears to posit a gap in any grammatically licit site that it encounters, as soon as such a site becomes apparent in the course of incremental processing. This process has come to be known as 'active gap filling', and the corresponding parsing strategy as the Active Filler Strategy (Clifton & Frazier, 1989; Frazier & Clifton, 1989).

There are multiple lines of evidence for active gap filling. First, processing difficulty ensues when a potential gap site turns out to be occupied by another element. This phenomenon, known as a *filled gap effect*, was first demonstrated by Stowe (1986), and is illustrated in (3):

(3) *Which dog* did the family choose a leash for _____ at the pet store?

In this sentence, difficulty would occur when the reader encounters *a leash*. This effect appears to reflect the syntactic reanalysis that is required when the parser posits a gap after the first licensing verb that it encounters (*choose*), and then finds that this gap site is filled by other material. Second, processing difficulty results when positing a gap in a grammatically licit location yields an implausible interpretation (Traxler & Pickering, 1996), as in (4):

(4)		*Which dog* did the family eat happily with _____?

This difficulty, which would arise upon encountering the gap-licensing verb (*eat*), suggests that the processor has posited a direct object gap after *eat*, and has initially computed an interpretation on which the family has eaten the dog.

Notably, such plausibility effects at the verb do not depend on that verb's subcategorization bias, i.e., the frequency with which the verb occurs with a direct object. In eyetracking and self-paced reading experiments, Pickering and Traxler (2003; Staub, 2007) found that the critical plausibility effect does not depend on whether the verb usually occurs with a noun phrase direct object, or tends to be followed by a prepositional phrase. The increase in reading time on the verb, compared to plausible controls, did not differ between (5a), where the verb *killed* tends to occur with noun phrase direct object, and (5b), where the verb *worried* tends to be followed by a prepositional phrase.

(5)	a.	That's the country that the soldier killed for during the war in Korea.
	b.	That's the car that the dog worried about after going to the vet because of an injury.

Thus, active gap filling appears to be a general parsing heuristic rather than a frequency-sensitive mechanism, as a direct object gap is actively posited in the first grammatically licit site that is encountered during incremental processing of the sentence, even when a direct object would not be expected in this location based on the licensing verb's subcategorization bias.

This work has established that the parser actively posits a gap once a filler has been identified to its left. A much smaller literature has investigated processing of sentences in which an element is extraposed to the right of its canonical position in English, so that a gap appears before the corresponding filler (Levy, Fedorenko, Breen, & Gibson, 2012; Staub, Clifton, & Frazier, 2006). Staub et al. (2006) conducted two eyetracking experiments investigating the processing of Heavy NP Shift, in which a verb's direct object is shifted to the right, over another element. The occurrence of Heavy NP Shift is conditioned by the phonological weight of the object (i.e., length) and by its discourse status (Arnold, Losongco, Wasow, & Ginstrom, 2000; Wasow, 1997). An example from Staub et al. (2006) is in (6):

(6)		Jack watched _____ from the stands his daughter's attempt to shoot a basket.

Staub et al. observed a pattern of processing difficulty suggesting that whenever the verb does not categorically require a direct object (e.g., *watched*), the intransitive analysis is initially adopted in preference to the Heavy NP Shift analysis. In this case, encountering the shifted direct object resulted in disruption. Notably, this was true even for verbs that were preferentially, but not obligatorily, transitive. Only when the verb was obligatorily transitive (e.g.,

praised) was a gap readily posited in the post-verbal position. In this case, there was some disruption on the intervening prepositional phrase (e.g., *from the stands*), but the shifted direct object was processed easily. Similarly, based on a series of self-paced reading studies examining processing of extraposed relative clauses, Levy et al. (2012) also reach the conclusion that the comprehender expects an extraposed RC only when an RC is made essentially obligatory.

Thus, the literature presents a highly coherent picture of how the processor deals with ambiguity as to whether a gap is present at a given location in a string. When a filler has already been unambiguously identified, a gap is posited in the first grammatically licit location. But when no filler has been identified, the parser avoids positing a gap. In both cases, the parser's behavior appears to be more-or-less categorical, i.e., insensitive to subcategorization biases of lexical heads.

The existing evidence, then, is consistent with a general proposal by De Vincenzi (1991) governing the processing of filler-gap dependencies, dubbed the *Minimal Chain Principle* (MCP):

> MCP: Avoid postulating unnecessary chain members at S-structure, but do not delay required chain members.

The second clause of the MCP describes active gap filling, and the first clause states that when there is ambiguity as to whether a filler-gap dependency is present, an analysis that does not require such a dependency should be adopted in preference to an analysis that would require one. These two clauses of the MCP can be regarded as consequences of a single principle minimizing the maintenance of dependencies: Do not form dependencies that are not required, and when they are required, make them as short as possible. The motivation for the MCP is similar to the motivation for other structural constraints that have been argued to govern the behavior of the parser, such as Minimal Attachment and Late Closure (Frazier, 1978; Frazier, 1987). This motivation is to avoid the memory burden associated with maintaining unstructured material. The two clauses of the MCP have the effect of assigning each (noun) phrase a thematic role in the sentence as rapidly as possible, and assigning each thematic role to a (noun) phrase as rapidly as possible.

In the present work, we test a further prediction arising from the first clause of the MCP. In addition to predicting that the parser will avoid positing a gap when a filler has yet to be identified, this clause predicts that the parser should avoid analyzing material as the filler in a long-distance dependency if an alternative parse in which that material is not a filler is available.

In sentences containing *wh*-questions and relative clauses, there is typically no syntactic ambiguity upon reaching the filler; in (2), for example, encountering the relative pronoun *that* unambiguously introduces a relative clause, and therefore triggers a search for the corresponding gap position. But, this is not always the case. Consider (7):

(7)	The information that the health department provided

This string is ambiguous between an analysis on which *that the health department provided* is part of a nominal complement clause (CC), and an analysis on which it is a relative clause (RC). These two analyses would give rise, respectively, to continuations like in (8a–b):

(8)	a.	The information that the health department provided a cure reassured the tour operators.
	b.	The information that the health department provided reassured the tour operators.