Thomas G. Bever
Brian McElree

Empty Categories Access Their Antecedents during Comprehension

This material may be protected by copyright law (Title 17 U.S. Code).

The distribution of implicit anaphoric entities ("gaps") has played a critical role in the recent development of generative grammar known as Government-Binding Theory (Chomsky (1981; 1986)). One of the most important discoveries of this theory is the generalization that the same set of constraints apply to the relation between explicit anaphoric entities and their antecedents as to the relation between implicit anaphoric entities and *their* antecedents. This emphasizes the formal claim that gaps are literally implicit anaphors, which access their antecedents in the same way as explicit anaphors. Such a claim invites behavioral verification, to the effect that gaps access their antecedents during comprehension in the same way as pronouns. In this research note we report initial success in demonstrating such access by gaps. The demonstration includes direct evidence that pure movement traces access their antecedents more strongly than obligatory PRO, even in constructions such as NP-raising and *tough*-movement; most surprising is evidence that the object NP-trace in the passive construction also accesses its antecedent.

We first validated on single sentences, an experimental technique developed originally for studying the accessing of a discourse antecedent by an explicit anaphoric word (McKoon and Ratcliff (1980), Dell, McKoon, and Ratcliff (1983), Cloitre and Bever (1983)). This technique utilizes a "priming" paradigm. In our studies a subject reads a sentence on a computer screen, presented one phrase at a time each time the subject pushes a button. At the end of the sentence a probe word appears on the screen. The subject must decide quickly whether the probe word was contained in the sentence just presented. (See the appendix for the background of this technique and experimental details.)

We found that the probe technique is sensitive to anaphoric accessing within a single sentence. For example, the time needed to decide that the word *astute* was in sentence (2) is 0.9 seconds, whereas the time for the same decision after sentence (1) is 1 second.[1]

We are indebted to Charles Clifton for comments on the manuscript of this article, most of which we honored.

Texts of all experimental material are available upon request to the first author.

[1] All differences were tested by correlated t-tests on the times for positive responses. This difference was significant, $p < .05$ by subjects (2-tailed), $p < .025$ by items (2-tailed).

Linguistic Inquiry, Volume 19, Number 1, Winter 1988
35-43
© 1988 by The Massachusetts Institute of Technology

(See the appendix for a description of experiment 1. See table 1 and the statistical note.)

(1) The astute lawyer who faced the female judge hated the long speech during the trial. (nonanaphor construction)
(2) The astute lawyer who faced the female judge hoped he would speak during the trial. (pronoun construction)

Probe: *astute*

The error rate was also larger following construction (1) than construction (2); that is, more adjective probes were incorrectly rejected following (1). These differences reflect the fact that the pronoun *he* occurs late in (2) and is interpreted to have the subject phrase containing *astute* as its antecedent. The fact that recognition responses to the probe word in (2) are relatively fast and accurate confirms the view that the pronoun accesses a representation of its antecedent phrase. This access refreshes the salience of the noun phrase containing the adjective and facilitates the subject's immediate decision that the probe adjective was in the text. (We do not discuss pronoun constructions further, since the pronoun is an explicit anaphor and not directly comparable to gaps.)

We used the priming technique to see if gaps also access their antecedents with the same effect as pronouns. Constructions (3), (4), and (5) exemplify our materials and the probe word recognition time they elicited.

(3) The astute lawyer who faced the female judge strongly hoped [PRO] to argue during the trial. (PRO construction)
(4) The astute lawyer who faced the female judge was certain [e] to argue during the trial. (NP-raising construction)
(5) The astute lawyer was hard for the judge to control [e] during the very long trial. (*tough*-movement construction)

The results show that the PRO construction elicited faster probe recognition times compared with the nonanaphor sentence (1).[1] *Tough*-movement constructions elicited faster recognition responses, and fewer errors, than the nonanaphor sentence[2] and PRO.[2] NP-raising constructions elicited faster times and fewer errors than nonanaphor[2] and PRO. The reaction time difference between PRO and raising was only marginally significant. However, the miss rate was significantly higher for PRO than for raising.[3] The facilitation following all the gap constructions relative to the nonanaphor construction is a clear indication that gaps other than *wh*-traces indeed access their antecedents during comprehension, just as suggested by their role in linguistic descriptions. Furthermore, movement traces showed a larger facilitation than PRO. A straightforward interpretation of this difference rests on the distinct formal properties of their relations to their antecedents.

[2] $p < .025$ by subjects (2-tailed), $p < .01$ by items (2-tailed).
[3] $p < .025$ by a Wilcoxon matched pairs, signed ranks test. Other significant error rate differences were: PRO vs. *tough*-movement ($p < .05$) and vs. pronoun ($p < .05$); nonanaphor vs. pronoun ($p < .025$), vs. raising ($p < .01$) and vs. *tough*-movement ($p < .025$).

Table 1
Response times (seconds) to recognize that the probe word was in the preceding sentence (error response times are not included in the mean reaction times). % error rates are in (parentheses); % subjects with at least 1 error on a given construction are in [brackets].

Experiment		*1*	
Nonanaphor (type (1))	1.05	(12)	[43]
Pronoun (type (2))	0.93	(6)	[33]
PRO (type (3))	0.96	(15)	[50]
NP-raising (type (4))	0.92	(7)	[27]
Tough-movement (type (5))	0.87	(7)	[27]
Experiment	*2*	*3*	*4*
Adjective/Active (types (7), (9))	1.12 (8)	1.01 (11)	0.93 (7)
Short/Full passive (types (6), (8))	1.0 (7)	0.93 (9)	0.84 (7)

Movement traces are roots of chains with their antecedents as heads, in which the trace position defines the thematic role, and the phrase structure configuration of the sentence specifies the relation between the trace and its antecedent. That is, the relation between movement trace and its antecedent is uniquely determined and crucial for assignment of a thematic role to the antecedent. PRO is not in a chain with its antecedent, as reflected in the fact that the antecedent is assigned a thematic role by the matrix verb, whereas PRO itself is assigned a separate thematic role by the complement verb. That is, the coreference relation between PRO and its antecedent is not determined by the phrase structure configuration alone, nor does PRO provide the thematic role for its antecedent. Accordingly, traces have a more structurally bound and more thematically consonant relation to their antecedents than does PRO. If we took these formal distinctions to correspond to mental ones, it would be consistent with the fact that during comprehension, movement traces excite more priming of their antecedents than does PRO.

The findings in experiment 1 validate our method of assessing the extent to which gaps access their antecedents in single sentences, with an effect similar to that of explicit pronouns. The results also show that a distinction made by the grammar between different kinds of gaps is reflected behaviorally in a plausible way. The gaps in the cases we have examined so far all have an intuitive basis as well as a formal one: in (3)–(5) naive informants will often agree that the gap position corresponds in some way to the designated antecedent. Thus, although it is gratifying that these positions act like anaphoric elements in comprehension, it is not completely surprising. We now turn to the

case of NP movement-trace in the passive construction, which is totally motivated by considerations internal to linguistic theory, with no directly supporting intuitions that the movement gap exists.

In the passive construction the subtheories of GB Theory combine to require that the surface subject is the head of a chain with a trace in object position; that is, the passive construction has a structure like (6), which contrasts with the superficially similar adjectival construction in (7).

(6) The astute lawyer who faced the female judge was suspected [e] constantly.
(7) The astute lawyer who faced the female judge was suspicious constantly.

We ran a probe recognition test similar to the preceding one (we used only passive preterites that passed all the distributional tests for being true verbal passives; see experiment 2 in the appendix). The time to recognize that *astute* was in the sentence was faster following the passive than following the adjectival construction (table 1).[4] This result might seem to have been due to the fact that subjects can control the time they spend reading each phrase in the sentence. On this view, subjects might have increased the time reading the final phrase of the passive sentence, so they were more ready than following the active sentence to decide about the following probe word (our computer program did not collect phrase reading times, but other studies suggest that there may be differences in sentence-final reading time for passive and active constructions; see Carrithers and Bever (1984), Carrithers (1985), Bever, Carrithers, and Townsend (1987)). To control for this, we ran the same materials in a paradigm in which the last phrase was presented for a constant time, not under the subject's control (see the description of experiment 3).[4] The results were similar to those of the first study. It was also the case that the surface structure subject thematic role in the adjectival in experiment 2 is different from the thematic role of the corresponding surface subject in the passive. To control for the possibility that this difference affected the results, we ran a third study with materials like (8) and (9), in which the two thematic roles are the same but in reverse order (we also used optionally intransitive verbs in all our active constructions (9), which could stop with a complete sentence right after the verb, just as in the passives (8)). (See experiment 4.)

(8) The astute lawyer who faced the female judge was suspected by the boys.
(9) The astute lawyer who faced the female judge had spoken to the boys.

Again, probe adjective recognition times were faster following the passive.[4] At this point we concluded that the finding was not an obvious artifact of methods or materials. The simplest interpretation consistent both with the structural claim that passive has a trace and with the results of our first study is that the NP-trace in passive is assigned during comprehension and that it accesses its antecedent, with the usual priming effect.

Much remains to be done. For one thing, we have not spelled out the comprehension mechanism within which the antecedent accessing and probe priming takes place (see

[4] $p < .01$ by subjects (2-tailed), $p < .025$ by items (2-tailed).

Cloitre and Bever (1983; in preparation), McElree and Bever (in preparation)). Furthermore, our probes occur at the end of the sentence, so we have no information about the time course of accessing the antecedent. Nor have we used definitively unassailable materials—one could try to construct post hoc analyses of each of our individual sentence types to predict the different results without appealing to empty categories at all. But we have shown that the most consistent interpretation of the data assumes the "psychological reality" of the implicit categories and of the distinctions between them. It is our claim that any theory of comprehension will have to incorporate a mechanism that is responsive to empty categories as assigned and differentiated in linguistic theory.

We have used GB Theory to assign the formal distinctions between the kinds of empty categories. However, our behavioral results do not definitively reject any particular linguistic theory as a whole, so long as the theory respects these distinctions at some level of representation. For example, although Generalized Phrase Structure Grammar (Gazdar et al. (1985)) and Lexical Functional Grammar (Bresnan (1978)) do not postulate and make the distinctions among all the empty categories within their syntactic structure, they certainly do—or could—at the semantic level. The force of our results is to invalidate those grammars only insofar as their syntactic components are taken literally as complete models of sentence parsing. Since such proposals have appeared (Crain and Fodor (1985a,b) for GPSG; Bresnan and Kaplan (1982) for LFG), some reformulations may be in order. Correspondingly, the results do not demonstrate the psychological validity of GB Theory as a parser, any more than they invalidate other grammatical theories. One *could* take such results as demonstrating the independent operation during comprehension of distinct subtheories of GB Theory, as proposed for other phenomena by Freedman and Forster (1985), Frazier (1985), Clifton and Frazier (in press), Gerken and Bever (1986); however, we have not described a linking hypothesis that connects the grammatical model to comprehension behavior in a way that would allow empirical results to bear uniquely on each grammatical subtheory (nor may have the authors just cited; see Bever (in press)). Hence, our results only confirm the behavioral relevance of the distinctions made by the grammar in the abstract representations assigned to sentences. As always, it is a separate question what the mechanism is that respects and imposes these distinctions as we comprehend language.

Appendix: Background, General Method, and Procedure

Previous research on gap filling during comprehension has used sentences with two types of gaps with identifiable fillers: *wh*-gaps, whose antecedent "fillers" are identifiable before the gap, either by explicit *wh*-marking or by their presence in a unique *wh*-filler position; and obligatory PRO, whose filler is uniquely specified by a controller matrix verb (Wanner and Maratsos (1978), Frazier, Clifton, and Randall (1983), Clifton and Frazier (1986), Crain and Fodor (1985a,b)). One controversial question in this research involves the relative importance and independence of structural and lexical strategies that listeners use to assign a filler to a gap. In general, the studies utilize measures of complexity to show that sentences that conform to a particular kind of strategy are

relatively easy to comprehend. Tanenhaus, Carlson, and Seidenberg (1985) and Clifton and Frazier (in press) report pilot experiments suggesting that the antecedents of *wh*-gaps are "primed." Tanenhaus, Carlson, and Seidenberg used a "lexical decision" task, in which subjects decide whether a letter sequence is a word: the lexical decision about probe words following a *wh*-gap is faster than the decision not following a gap, when the antecedent of the gap rhymes with the probe word. Clifton and Frazier (using materials supplied by David Swinney) used a "word-reading time" paradigm, in which subjects read a probe word out loud when it appeared during a sentence: probe words presented 3 words after the *wh*-gap that were high associates of the gap's antecedent noun were read 4% faster than probe words that were low associates.

The experiments in the present study were devoted to exploring the direct evidence that an unmarked antecedent has been accessed. That is, we wanted to show direct evidence of the renewed mental presence of the antecedent shortly after the anaphoric gap. Our experiments all use a priming technique presented originally in studies of antecedent priming in multisentence discourses. (See McKoon and Ratcliff (1980), Dell, McKoon, and Ratcliff (1983), Cloitre and Bever (1983). See also Cloitre (1984), Cloitre and Bever (in preparation), and McElree and Bever (in preparation), for additional studies and analysis.) During each trial subjects repeatedly press a key on a computer keyboard; each key-press displays the next phrase of a sentence, presented in the center of the screen, so that the phrase replaces the previous phrase. Phrases were defined by making a phrase break from left to right at the nearest intuitively deepest break such that phrases were all 2–5 words long (we adopted some conventions to deal with intuitively problematic constituencies; see Janus and Bever (1985) for further details). The sentences were presented with normal use of upper and lower case and with normal punctuation. At some point the next key-press displayed the probe word (except in experiment 3; see below). To distinguish it from the text, the probe word appeared in the center of the screen, bounded by asterisks and printed in capital letters.

Subjects responded "yes" or "no" to indicate whether the probe word had appeared in the immediately preceding text, by pressing labeled keys on the keyboard. Subjects were asked to respond as quickly as they could. On one-fourth of the trials, comprehension questions about the text followed the response to the probe word, to keep the subjects' comprehension processes active. About one-fourth of the trials had experimental sentences like those in (1)–(9), in which the probe word was an adjective that occurred at the end of the sentence. The other three-fourths of the sequences were distractor trials. Some of the distractor trials had texts 2 or 3 sentences long, to discourage subjects from anticipating that the probe word would always occur after a single sentence; some had probe words not drawn from the preceding text; most had probe words other than adjectives, taken from a variety of serial positions in the text. Cloitre (1984) found that these features are an important part of the experimental design: subjects are extremely sensitive to any obvious characteristics of the experiment that allow them to make predictions about the probe word.

All the studies used a TRS-80 model 1 computer to present the stimuli and record

the responses; subjects sat about 2 feet from a standard display monitor. In experiment 1 the subjects were native speakers of English between 17 and 25 years of age, drawn from a paid and unpaid summer subject pool of students at Columbia University; in experiment 2 the subjects were all from a paid pool at Columbia; and in experiments 3 and 4 they were from an unpaid subject pool, participating in partial fulfillment of an undergraduate psychology course.

Experiment 1. There were 50 content sets with different content words; each set had 5 corresponding structural types of sentences, modeled on examples (1)–(5). The following distances were controlled to be the same (in words): in each structural type, the distance between the antecedent noun phrase and the probe word; in types (2)–(5), the distance between the antecedent and its anaphoric position, and the distance between the anaphoric position and the probe. In (2) only the first noun phrase was the same sex as the pronoun. The verbs in the complement clause were all middle verbs, with a mixture of transitive/intransitive preferences (see Connine et al. (1984)). There were 5 counterbalanced experimental groups, each with a different structural type of sentence drawn from a given content set; thus, each structural type of a given content set was responded to by a different group of subjects. The groups were also designed so that each subject read 10 trials of each structural type. This design, similar to all the ones below, allowed us statistically to examine effects separately by subject and content set ("items"). There were 30 subjects in this study.

Experiment 2. There were 10 sets of 2 sentences, modeled on (6) and (7), with a design and presentation method similar to that of experiment 1. There were 30 new subjects in this study.

Experiment 3. The materials and design were similar to experiment 2, except that 16 content sets were used. However, the subject-paced phrase presentation was slightly altered. Subjects had no control over the duration of the last phrase. When they pressed the button for the last phrase on critical trials, it appeared for 0.9 seconds, the average self-paced reading time duration that subjects had required for that phrase in experiment 2. An interval of 0.9 seconds after the final phrase disappeared, the probe word appeared. There were 24 new subjects in this study.

Experiment 4. There were 24 sets of 2 sentences modeled on (8) and (9), with a design and presentation similar to that of experiment 2. There were 20 new subjects in this study.

References

Bever, T. G. (in press) "Three Paradigms for the Study of Language and Cognition," in M. Macken, ed.

Bever, T. G., C. Carrithers, and D. T. Townsend (1987) "A Tale of Two Brains: The Sinistral Quasimodularity of Language," in *Proceedings of the IX Cognitive Science Society*, 764–774.

Bresnan, J. (1978) "A Realistic Transformational Grammar," in M. Halle, J. Bresnan, and G. Miller, eds., *Linguistic Theory and Psychological Reality*, MIT Press, Cambridge, Massachusetts.

Bresnan, J., ed. (1982) *The Mental Representation of Grammatical Relations*, MIT Press, Cambridge, Massachusetts.

Bresnan, J. and R. Kaplan (1982) "Grammars as Mental Representations of Language," in J. Bresnan, ed. (1982).

Carrithers, C. and T. G. Bever (1984) "Eye-fixation Patterns Confirm Theories of Sentence Comprehension," *Cognitive Science* 8, 157–172.

Carrithers, C. (1985) *The Comprehension of Canonical and Non-canonical Sentence Constructions*, Doctoral dissertation, Columbia University, New York.

Chomsky, N. (1981) *Lectures on Government and Binding*, Foris, Dordrecht.

Chomsky, N. (1986) *Knowledge of Language: Its Nature, Origin, and Use*, Praeger, New York.

Clifton, C. and L. Frazier (1986) "The Use of Syntactic Information in Filling Gaps," *Journal of Psycholinguistic Research*.

Clifton, C. and L. Frazier (in press) "Processing Sentences with Long Distance Dependencies," in M. Tanenhaus and G. Carlson, eds., *Linguistic Structure in Language Processing*, Reidel, Dordrecht.

Cloitre, M. (1984) *The Effects of Kind of Anaphor on the Accessibility of Antecedent Information*, Doctoral dissertation, Columbia University, New York.

Cloitre, M. and T. G. Bever (1983) "The Processing Efficiency of Pronoun Anaphors," *EPA Abstracts*, Spring.

Cloitre, M. and T. G. Bever (in preparation) "Pronouns and Noun Anaphors Access Distinct Levels of Representation during Discourse Processing," ms.

Connine, W. E., F. Ferreira, C. Jones, C. Clifton, Jr., and L. Frazier (1984) "Verb Frame Preferences: Descriptive Norms," *Journal of Psycholinguistic Research* 13, 307–319.

Crain, S. and J. D. Fodor (1985a) "How Can Grammars Help the Parser?" in D. R. Dowty, L. Karttunen, and A. M. Zwicky, eds., *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*, Cambridge University Press, New York.

Crain, S. and J. D. Fodor (1985b) "Rules and Constraints in Sentence Processing," in *Proceedings of the 15th Annual Meeting of NELS*, GLSA, University of Massachusetts, Amherst.

Dell, G. S., G. McKoon, and R. Ratcliff (1983) "The Activation of Antecedent Information during the Processing of Anaphoric Reference in Reading," *Journal of Verbal Learning and Verbal Behavior* 22, 121–132.

Frazier, L., C. Clifton, and J. Randall (1983) "Filling Gaps: Decision Principles and Structure in Sentence Comprehension," *Cognition* 13, 187–222.

Frazier, L. (1985) "Modularity and the Representational Hypothesis," in *Proceedings of the 15th Annual Meeting of NELS*, GLSA, University of Massachusetts, Amherst.

Freedman, S. E. and K. I. Forster (1985) "The Psychological Status of Overgenerated Sentences," *Cognition* 19, 101–132.

Gazdar, G., E. Klein, G. Pullum, and I. Sag (1985) *Generalized Phrase Structure Grammar*, Harvard University Press, Cambridge, Massachusetts.

Gerken, L. A. and T. G. Bever (1986) "Linguistic Intuitions Are the Result of Interactions between Perceptual Processes and Linguistic Universals," *Cognitive Science* 10, 457–476.

Janus, R. and T. G. Bever (1985) "Processing of Metaphoric Language: An Investigation of the 3-Stage Model of Metaphor Comprehension," *Journal of Psycholinguistic Research* 14, 473–489.

McElree, B. and T. G. Bever (in preparation) "The Psychological Reality of Linguistically Defined Gaps," ms.

McKoon, G. and R. Ratcliff (1980) "The Comprehension Process and Memory Structure Involved in Anaphoric Reference," *Journal of Verbal Learning and Verbal Behavior* 19, 668–682.

Tanenhaus, M. K., G. N. Carlson, and M. S. Seidenberg (1985) "Do Listeners Compute Linguistic Representations?" in D. R. Dowty, L. Karttunen, and A. M. Zwicky, eds., *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*, Cambridge University Press, New York.

Wanner, E. and M. Maratsos (1978) "An ATN Approach to Comprehension," in M. Halle, J. Bresnan, and G. Miller, eds., *Linguistic Theory and Psychological Reality*, MIT Press, Cambridge, Massachusetts.

(Bever)
Department of Psychology
University of Rochester
Rochester, New York 14627

(McElree)
Department of Psychology
Columbia University
New York, New York 10027