Feature manipulation in sentence comprehension

Charles Clifton Jr., Lyn Frazier & Patricia Deevy

Agreement features participate in several grammatical dependencies such as subject-verb agreement and antecedent-reflexive binding. We explore whether there is any stage of sentence processing in which a linear structure-independent feature matching process occurs, e.g., as a means of identifying candidate antecedents which are then subject to further structural and semantic evaluation. Three self-paced reading studies fail to confirm the predictions of the structure-independent matching hypothesis. Turning to the question of how features are represented in the processing of filler-gap dependencies and subject verb dependencies, two incremental grammaticality judgment experiments suggest that number features are passed through the syntactic phrase-marker with some consequent confusion when a single link in the tree must transmit both a singular feature, transmitted between the subject and a local verb, and a plural feature, transmitted between the filler and a gap occurring after the local verb. However, if the filler- gap path is terminated before the subject and local verb are encountered, then perceivers more accurately detect the violation in ungrammatical sentences where a singular subject is paired with a local plural verb.*

0. Introduction

Number and gender features participate in several linguistic phenomena. Restricting attention to English, an inflected verb must agree in number with its subject (1), a determiner must agree in number with the head N (2), and a relative pronoun or reflexive must agree with its antecedent in number and gender, (3) and (4).

(1) a. The girl is/*are here.
b. The girls *is/are here.

(2) a. this/*these book
b. *this/these books

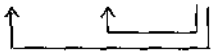
(3) a. [the girl who (sg)/*(pl) is here]
b. [the girls who *(sg)/(pl) are here]

(4) a. John liked himself/*herself/*themselves.
b. John thought Mary liked him/*her.
c. Mary thought John liked *him/her.
d. Mary thought John liked *him/her.

Rivista di Linguistica, 11.1 (1999), p. 11-39 *(ricevuto nel luglio 1998)*

The processor might in principle manipulate or use features in any of several different manners. One possibility is that features are directly matched against each other in a linear or structure-independent manner whenever a syntactic dependency exists between two phrases X and Y which must agree in 'Phi Features' (number, gender). Upon encountering Y, the processor might activate any NP bearing the same Phi features, as schematized in (5), where a means gender and number are compatible:

(5) X_α X_α Y_α



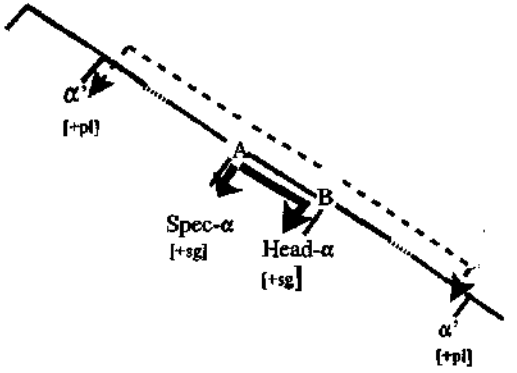
In other words features might guide the processor's hypothesis formation, identifying candidates for further structural and semantic evaluation. Section 1 presents three new experiments and summarizes one in press that allow us to evaluate this possibility in cases where Y is an anaphor.

There are two reasons why Section 1 concentrates on anaphor processing to assess the possibility of structure-independent use of features. First, if features are ever checked independent of the syntactic position of the phrase bearing the features, it is most likely to be in the processing of anaphors. Second, considering research on comprehension (as opposed to production), we are aware of proposals of structure-independent use of features only in the case of antecedent-anaphor relations. Therefore, it is of interest to establish whether the processing of anaphors makes structure-independent use of features.

For the type of feature dependency examined in Section 1, which involves coreference between two independent theta positions, coindexation may be an appropriate way of treating feature checking by the processor. In Section 1 we ask whether there is ever direct, structure-independent feature matching. However, in Section 2, we turn to feature dependencies between pairs of elements where the existence of the second member of the dependency is predictable in advance, such as agreement between a verb and its subject and the sharing of Phi features between a filler and the trace it binds. For these more predictable dependencies, we will assume that feature manipulation is structure-dependent at *all* stages of processing. We address the question of whether agreement-checking is a matter of directly accessing coindexed phrases to ensure feature compatibility (as in the case of anaphors) or whether it sometimes involves passing features

through the phrase structure tree. In Section 2, we will test whether difficulty results when the processor must pass conflicting features through the same local region of the phrase marker as in (6). Here, the branch between A and B is involved in checking the agreement between the number features of a Head and its Spec, as well as passing a plural feature that originates higher in the tree. On a feature-passing view, difficulty is expected. On a coindexiation view, no difficulty is expected. Because of the nature of the question addressed in Section 2, the experimental methods employed as well as the materials used will differ from those in Section 1.

(6)



1. Structure-independent feature-matching?

Psycholinguistic studies of within-sentence anaphora suggest that all and only the grammatically permissible antecedents for an anaphor are activated immediately following the anaphor e.g., a pronoun or reflexive. Specifically, Nicol (1988) demonstrated the existence of cross-modal semantic priming for words semantically-related to the permissible antecedents for a pronoun (underlined in (7)) or reflexive (in (8)), measuring immediately after the anaphor.

(7) The *janitor* told the *landlord* that the fireman with the gas mask would protect *him*.

(8) The janitor told the landlord that the *fireman* with the gas mask would protect *himself*.

In Nicol's study, an NP with appropriate number and gender features for the anaphor was not activated if it occurred in a position

where it could not bind the anaphor according to the dictates of Binding Theory (see Chomsky 1981). [1] With respect to the detailed process of antecedent-anaphor processing, these results are somewhat ambiguous. They might mean that the processor never even considers grammatically illegitimate antecedents at any stage in its search for an antecedent. Alternatively, the results might be interpreted as indicating that cross-modal priming occurs during the semantic evaluation of antecedents. In this case, the results indicate that only grammatically legitimate antecedents are evaluated semantically. This leaves open the possibility that feature-appropriate NPs are considered during the candidate-identification phase of processing even if the NP occurs in a syntactic position which does not satisfy the Binding Theory. These possibilities are summarized in (9).

(9) a. #1. Find NPs with appropriate features (e.g. *janitor, land lord, fireman* in (7) and (8))
#2. Check Binding Theory (e.g. reject *fireman* in (7))
#3. Semantically evaluate permissible candidates

b. #1. Check features and Binding Theory, e.g., in the case of the reflexive, move through the tree to c-commanding positions in the local domain
#2. Semantically evaluate permissible candidates.

Badecker & Straub (1994; Straub & Badecker, 1994) report results most consistent with (9a). They report that the addition of a decoy antecedent (an NP with number and gender features appropriate for the anaphor) increases processing times even if the decoy is not in a syntactically permissible position to serve as antecedent according to Binding Theory. In a self-paced reading experiment, sentences with two gender-appropriate candidate antecedents for a pronoun took longer to process than those with only one. This was true even in sentences like (10), where only one NP is syntactically accessible to the pronoun. The gender and number of the inaccessible NP mattered: reading times were longer when it matched the pronoun (10b) than when it did not (10a). Straub & Badecker reported comparable effects when they used reflexives instead of the pronouns of (10). These results are expected according to the sequence in (9a), where even NPs in inaccessible syntactic positions are checked as possible antecedents if they bear appropriate Phi features for the anaphor. However, it is unexpected if syntactic position is consulted first, or together with Phi features, as indicated in (9b).

(10) a. John thinks that *Mary* will give him a better cut of venison next year
b. John thinks that *Bob* will give him a better cut of venison next year.

Straub & Badecker (1994) pursued the idea that feature-appropriate NPs are considered as antecedents even when they appear in syntactic positions rendering them as inappropriate antecedents. The idea that candidate antecedents are identified initially through structure-independent feature matching fits well with a multiple constraints view of processing, where weighted constraints of all types apply simultaneously. Because an antecedent and anaphor almost invariably match in number and gender, the feature-matching 'constraint' is very strong and would be expected to yield candidate antecedents quickly. By contrast, constraints on the syntactic relation between an anaphor and its antecedent are more variable because both the anaphor and antecedent may be in any of several syntactic positions. Thus syntactic constraints such as those captured by Binding Theory might be expected to be weaker and thus to be effective later than the feature-matching constraint. In short, early identification of potential antecedents based on structure-independent feature matching is not just compatible with a multiple constraints view post hoc but is really expected in advance given that the feature matching constraint is highly predictive. Therefore the feature-matching constraint should be weighted stronger than a weaker or less predictive constraint such as the constraints of Binding Theory.

1.1. Experiment 1

Imagine that direct (linear) feature matching takes place, as illustrated in (5) above. In a sentence like (11a), when the processor encounters the reflexive *himself*, a number appropriate NP *nurse* should be activated first (assuming more recent material is more easily accessed than more distant material). The number appropriate NP headed by *son* should also be accessed.

(11) a. The son of the pretty nurse hurt himself in a bad accident
b. The son hurt himself in a bad accident
c. The son of the pretty nurse hurt Fred in a bad accident
d. The son hurt Fred in a bad accident.

On this account, the syntactically inaccessible singular NP *nurse* should slow the processing of the sentence with a reflexive (11a) more than it should slow the processing of a sentence without a reflexive (11c). To control for the relative difficulty of a reflexive vs. a name per se, the controls in (11b) and (11d) were included. Direct linear features matching predicts an interaction: (11a) - (11b) > (11c) - (11d).

1.1.1. Method

16 sentences like (11) and (12) were constructed, with four versions of each. All sentences appear in Appendix 1. Eight of the sentences began with an 'NP of NP' phrase, and eight began with an 'NP with NP' phrase, but since the same pattern of results was obtained for both types of sentences, the manipulation will not be discussed further. Two versions of each sentence had a reflexive in the VP (a,b) and two had a proper name in its place (c,d). Two (a,c) had complex noun phrases as subject, whereas two had a simple noun phrase (b,d). The noun phrase added in the complex noun phrase condition (a,c) was always biased to a different gender than the head of the subject noun phrase. For example, *handsome* in (12b,d) biases toward a masculine referent for NP2 (*the handsome friend*) whereas the head noun is feminine. The sentences were divided into two regions, as shown, for presentation in a self-paced reading study.

(12) a. The waitress with the handsome friend/ burned herself after spilling the soup
b. The waitress/ burned herself after spilling the soup
c. The waitress with the handsome friend/ burned Sarah after spilling the soup
d. The waitress/ burned Sarah after spilling the soup.

These 16 sentences were embedded in a list of 94 sentences total, exhibiting a wide variety of syntactic constructions (including several "garden-path" constructions). 48 University of Massachusetts undergraduates read the sentences in individually-randomized orders, using a phrase-by-phrase self-paced reading task with a moving window in which successive button-presses caused underscore 'preview' characters to turn into the letters of the sentence and then back into underscores. These subjects were divided into four counterbalancing groups, so that each subject saw four experimental sentences in each experimental condition, and so that, over all subjects, each experimental sentence was tested equally often in each condition. The sentences in Experiment 1 were divided into two phrasal regions for presentation, as indicated by the '/' in (12). The time to read each region was recorded by a microcomputer, as were subjects' answers to two-choice wh-questions about the subject or object of the main verb which followed half the sentences were recorded.

1.1.2. Results

The mean reading times for the second regions of the sentences appear in Table 1. These times are presented in ms per character to obtain rough comparability in reading times among sentences of different lengths, but not to correct for any length differences among the experimental conditions. Reading times in ms, unadjusted for length, yielded fully comparable results. All statistical results presented are significant beyond the p = .05 level, unless otherwise indicated.

Table 1. Mean Reading Times (ms/character), Region 2, Experiment 1

Object Form	Subject Complexity		
	Complex	Simple	Mean
Reflexive	69	60	65
Name	86	72	79
Mean	77	66	72

The second phrases of sentences with simple subjects were read faster than sentences with complex subjects (66 vs 77 ms/character; F1(1,47) = 23.14; F2(1,15) = 15.52). Sentences with reflexives were processed faster than sentences without reflexives (65 vs 79 ms/char; F1(1,47) = 33.23; F2(1,15) = 28.44). There was no hint of an interaction (p > .30).

Questions were answered correctly 78% of all trials.

1.1.3. Discussion

Experiment 1 indicates that sentences with a complex NP took longer to process than sentences with a simple NP. This effect could be attributed to any of a variety of factors (length, complexity, etc.) and is expected on any theory. Sentences without a reflexive were harder to process than sentences with a reflexive. We suspect that this is due to the greater number of discourse participants introduced in the sentences without a reflexive. In terms of a theory like Heim's (1982), only one index card, or discourse referent, needs to be introduced in the sentences with a reflexive whereas two index cards (discourse referents) must be set up in the sentences without a

reflexive. The absence of an interaction suggests that the singular NP which intervened between the head of the subject (*waitress*) and the reflexive (*herself*) did not interfere with binding of the reflexive or the search for its antecedent. Perhaps the intervening NP did not interfere with reflexive binding because the intervening NP was not of the same gender as the reflexive. To test this possibility, we conducted a second experiment using intervening NPs of the same gender as the reflexive.

1.2. Experiment 2

1.2.1. Method

Thirty-two sentences were constructed, with four versions of each, as illustrated in (13).

(13) a. The son of the fireman/ hurt himself/ in a bad accident
b. The son/hurt himself/ in a bad accident
c. The son of the fireman/ hurt Fred/ in a bad accident
d. The son/ hurt Fred/ in a bad accident.

The complex initial NP versions of 16 sentences began with an 'NP of NP' phrase, while the remaining 16 sentences began with an 'NP with NP' phrase. As was the case in Experiment 1, this manipulation resulted in no interesting effects, and will not be discussed further. The four versions of each sentence corresponded precisely to the four versions described in Experiment 1, except that the object of the preposition of or with in the complex initial NP versions of Experiment 2 was of the same gender as the head noun of the NP. Two-choice wh-questions were made for half of the experimental sentences, and these 32 items were embedded in a list with a total of 175 sentences in it, of a wide variety of constructions.

48 University of Massachusetts undergraduates read the sentences in individually-randomized orders. The counterbalancing and presentation procedures were as described in Experiment 1, except that each experimental sentence was presented in three regions, as indicated by the '/' marks in (13) in an attempt to obtain data more sharply focused on the point of possible processing disruption.

1.2.2. Results

The reading times for regions 2 and 3 are presented in Table 2, in ms/character.

As in Experiment 1, the results for unadjusted ms were very similar to the ms/char results.

Sentences with a reflexive were processed faster than sentences without a reflexive in Region 2 (52 vs 61 ms; F1(1,47) = 79.2; F2(1,31) = 51.09). The difference disappeared in Region 3 (F < 1). Sentences with simple NPs were read faster than sentences with a complex NP both in Region 2 (52 vs 61 ms; F1(1,47) = 41.09; F2(1,31) = 73.59) and in Region 3 (61 vs 61 ms; F1(1,47) = 6.65; F2(1,31) = 5.46). The interaction was completely nonsignificant for each region (F < 1 in each case).

Table 2. Mean Reading Times (ms/character), Regions 2 and 3, Experiment 2

Object Form	Region 2			Region 3		
	Subject Complexity					
	Complex	Simple	Mean	Complex	Simple	Mean
Reflexive	56	48	52	63	61	62
Name	65	57	61	61	61	63
Mean	61	52	57	61	61	62

Accuracy to the questions after experimental sentences averaged 96%.

1.2.3. Discussion

The absence of an interaction in Experiment 2 strengthens the conclusion that an NP in the string between the head of the subject and a reflexive does not interfere with binding a reflexive or finding an antecedent for the reflexive. We think this finding is difficult to reconcile with a theory of binding where antecedents are identified by a backward search through the terminal string. It specifically conflicts with theories like Badecker & Straub (1994) and Straub & Badecker (1994), who claim that structural factors (specifically, those described by Binding Theory) do not initially constrain which NPs are considered as antecedents for pronouns and anaphors. Rather, they claim, antecedent NPs for either pronouns or reflexives are identified based on number and gender features and only then excluded if binding theory prohibits a feature-appropriate NP from serving as an antecedent. However, Badecker & Straub could (and in recent unpublished work, do) claim that only antecedent NPs in prominent positions (e.g., topic, or argument of main verb) are considered as antecedents of a pronoun. NPs in relative clauses, like those used in Experiments 1 and 2, may be inaccessible, and hence produce no

interference. Experiment 3 was designed to overcome this interpretive problem by using sentences with potentially-interfering NPs in prominent positions, just as Badecker & Straub did.

1.3. Experiment 3

Experiments 1 and 2 disconfirmed the predictions of a simple version of the direct (linear) feature matching approach. They suggested that an NP that is syntactically inaccessible because of Binding Theory is not checked during early phases of candidate antecedent identification. This does not necessarily disconfirm Badecker & Straub's proposal because in our materials the decoy NP was contained inside a larger NP, while the two NPs were contained in matrix vs. embedded sentences in the Badecker & Straub materials. Hence, Experiment 3 tried to replicate the basic finding of Badecker & Straub, that an inaccessible feature-appropriate NP slows processing of an anaphor.

Consider sentences like those in (14).

(14)
a. John thinks that Bill owes his sister another chance to solve the problem
b. John thinks that Betsy owes his sister another chance to solve the problem
c. Joan thinks that Bill owes his sister another chance to solve the problem
d. John thinks that Bill owes him another chance to solve the problem
e. John thinks that Betsy owes him another chance to solve the problem
f. John thinks that Bill owes himself another chance to solve the problem
g. Joan thinks that Bill owes himself another chance to solve the problem.

Using a self-paced word by word reading technique, Badecker & Straub investigated the reading of sentences like (14 d-g). They found disruptions precisely two words following the pronoun or anaphor in sentences like (14d) and (14f), compared to (14e) and (14g). [2] They argued that the structurally-inaccessible NP in the former sentences was considered as an antecedent for the pronoun or anaphor because of its gender-appropriateness, in spite of its inaccessibilty.

Experiment 3 attempted to replicate this finding, using materials constructed following the description provided by Badecker &

Straub. It also added items (14a-c), because in these, the potentially-interfering NP is structurally accessible as an antecedent for the possessive pronoun, and in fact (14a) represents a full ambiguity which should cause disruption given the assumptions that Badecker & Straub use to interpret their data.

1.3.1. Method

42 sets of seven sentences like those in (14) were constructed and embedded in a list of 102 sentences total. 49 University of Massachusetts undergraduates were tested in a self-paced reading experiment using the procedures described for Experiments 1 and 2, except that sentences were presented word-by-word (as in Badecker & Straub) rather than phrase-by-phrase. Seven counterbalancing groups were used to ensure that each sentence was tested equally often in each version. Two-choice wh-questions were presented after half of the experimental items; 15 of these questioned the antecedent of the pronoun or reflexive or the noun that followed a possessive pronoun.

1.3.2. Results

The mean reading times for each region of the experimental sentences, from the verb of the embedded clause through three words after the pronoun or anaphor, appear in Figure 1. The data are expressed in unadjusted ms rather than ms/character, because the latter measure over-adjusts for length differences in short regions, such as the pronoun region. The upper panel contains the data from the personal pronouns and reflexives (versions (14d-g)) while the lower panel contains the data from the possessive pronoun sentences (versions (14a-c)).

Analyses of variance indicated that the only significant differences in reading times occurred in the region of the pronoun/reflexive and two words after the possessive pronoun. The former effect ($F1(6,48) = 2.73$, $p < .02$; $SE_{difference} = 15.7$ ms) can be attributed to the longer times for reflexives than for personal or possessive pronouns (means of 573 ms vs. 539 ms, with a range only 17 ms among all personal and possessive pronouns). The latter effect ($F1(6,48) = 2.12$, $p = .05$; $SE_{difference} = 12.0$ ms) can be attributed to the fact that reading was slower following personal pronouns whose antecedent was non-local (14b; 419 ms) than following personal pronouns with local antecedents (14c; 396 ms). Reading time for the second word following a pronoun with a non-local antecedent (14b) was also slower than reading time following any sentence with a personal or reflexive pronoun, which averaged 387 ms with a range of 384 to 389 ms. No

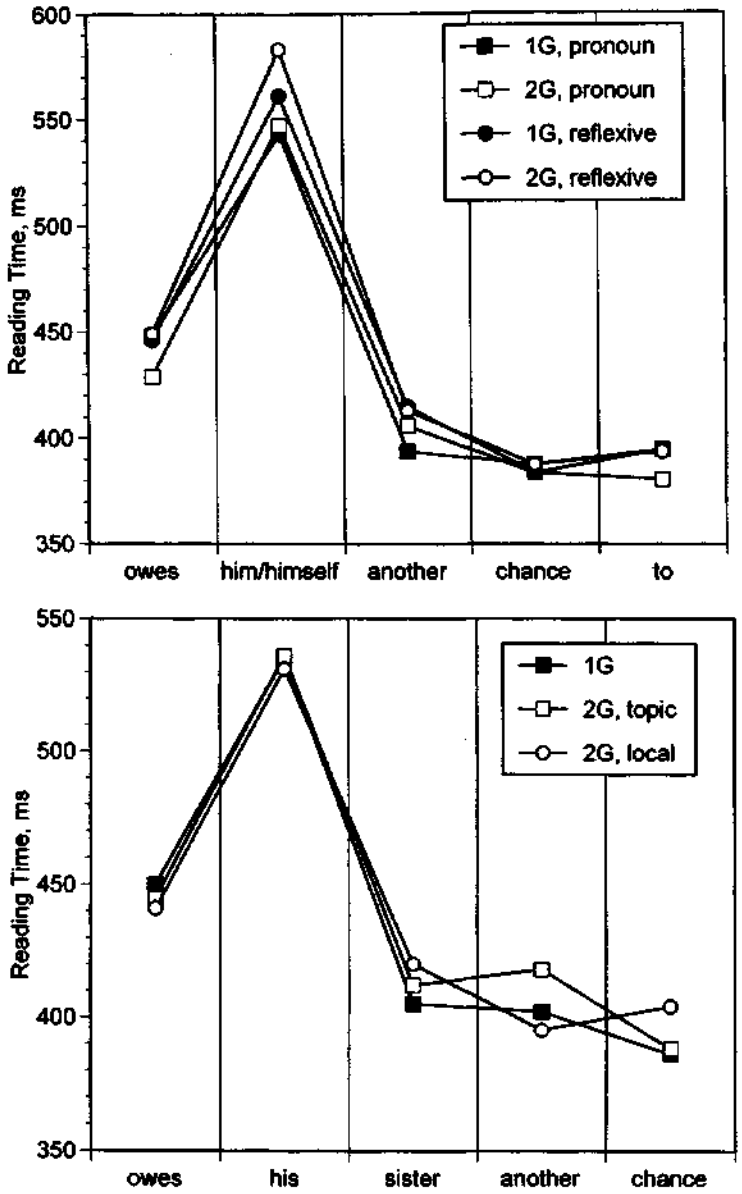


Figure 1. Mean reading times (ms), Experiment 3, for each word from the subordinate clause verb through three words after the pronoun or anaphor. Top panel: Personal pronouns and reflexives (14d-g). 1G = possible antecedents (accessible and inaccessible) have the same gender; 2G = possible antecedents have different genders. Bottom panel: Possessive pronouns. Topic/local designates the position of the gender-appropriate antecedent.

other difference among mean reading times in this or any other region approached significance.

Questions that involved potential pronoun antecedents followed 15 of the experimental sentences. Considering only questions which followed sentences with distinct-gender antecedents (i.e., unambiguous sentences), 76% of those with possessive pronouns were answered accurately, and 82% of those with simple personal pronouns.

1.3.3. Discussion

The results provide no support for interference from the presence of a feature-appropriate NP in either syntactically accessible (14a) or inaccessible (14d,f) positions. In the second position following the pronoun, which is where Badecker & Straub reported finding a difference as a function of gender of inaccessible antecedents, the range of mean reading times was a minuscule 5 ms (384 to 389 ms). We are not certain why we failed to obtain the results reported by Straub & Badecker. Subtle procedural differences may be responsible. For instance, Straub & Badecker followed every sentence with a speeded probe recognition task (apparently never probing a potential pronoun antecedent), but followed only 1/4 of the sentences with a verification question (none of which required resolution of the pronoun-antecedent relation to answer). It is logically possible that the probe task influenced the reading process. Since the probe task required verbatim memory only of words other than pronouns and their antecedents, it may have encouraged superficial processing of the sentences. Another possibility focuses on the fact that 15 of our 21 post-sentence questions required resolution of pronoun reference, while none of Badecker & Straub's did. This is perhaps related to the fact that our subjects read the pronoun or anaphor much slower than the surrounding words, while any slowing in Badecker & Straub's data was minimal. It may be that our subjects were more careful to resolve pronoun reference immediately.

Assuming that our data accurately reflect the processing demands of our sentences, we conclude that there is little support for the idea that any stage of antecedent identification activates NPs in syntactically inaccessible positions due to a feature match between the NP and an anaphor. The only trustworthy effect in the data, aside from longer reading times for reflexives than for personal or possessive pronouns (which is probably simply a function of the length of the word), was that a local antecedent was preferred over a non-local topic antecedent for a possessive pronoun (*his/her*). This effect is of interest because it establishes that the antecedent of a

pronoun is assigned within a word or two of reading the pronoun (cf. Di Domenico & De Vincenzi 1996, who suggest that their evidence indicates that pronoun antecedents are not always assigned on-line).

Direct (linear or structure-independent) feature matching has not been supported by any of the studies reported here. It is also at odds with the results of the studies reported in Albrecht & Clifton (1998). In several reading studies, sentence reading times were longer for a sentence containing a pronoun whose unambiguous antecedent was introduced inside a conjoined NP, *Stan and Pam* in (15), compared to examples where the antecedent was introduced as a simple NP (*Stan*).

(15) The cinema was quite full for the movie premier. Stan and Pam asked the usherette for assistance. He quickly followed the usherette to the seats. Then the usherette returned to the ticket office.

Albrecht & Clifton argue that NP-splitting is time-consuming and they account for the experimental results by invoking this cost. Notice that direct feature matching makes the wrong predictions here. If there is only one [+masculine +singular] NP, it must be the true antecedent for *he*. If direct feature matching guides antecedent selection in such cases, reading time should be fast whether the antecedent was introduced in a conjoined NP or not. It was not.

Considering the data discussed to this point, direct feature matching incorrectly predicted difficulty in (16a-c) due to the presence of a syntactically inaccessible but feature appropriate NP (marked X in (16)).

(16) a. $[\ [X_\alpha]]_\alpha$ ... $himself_\alpha$ (Experiments 1, 2)
b. $[X]_\alpha$ $[[\ ]_\alpha$ $himself_\alpha]$ (Experiment 3, (14f,g))
c. $[\ ]_\alpha$ $[[X_\alpha]$ $him_\alpha]$ (Experiment 3, (14d,e))
d. $[[\ _\alpha]$ and $[\ _\alpha]]_{-\alpha}$ ... He_α (Albrecht & Clifton 1998).

It also incorrectly predicts no difficulty due to the (feature inappropriate) conjoined NP in (16d). We conclude that features are not used in a structure-independent manner. Rather NPs which are tempting antecedents in terms of features, structure and discourse properties are evaluated as potential antecedents for an anaphor. We cannot pretend to answer all questions about how features are used during anaphor processing but we think it is clear that they are not used in a direct structure-independent feature match. [3] Our guess is that a bound anaphor initiates a search moving up through the tree, favoring the highest phrase in the local domain of the anaphor. By contrast, a pronoun may initiate a search for a topical or centered discourse entity in the discourse representation (see Gordon & Hendrick 1997).

In Section 1 we have investigated examples where feature agreement is involved in a 'right-to-left' dependency, i.e. where agreement is at issue only when the later phrase (the anaphor) is processed. Given that anaphors are not obligatorily present in a sentence, the dependency involved between an anaphor and its antecedent comes into play only once the anaphor is processed. In this sense, Section 1 concerned 'backwards' agreement dependencies. This contrasts with the standard examples of chain formation, e.g. relative pronoun-trace chains, where the existence of the trace is predictable as soon as the relative pronoun is encountered. In examples of binder-trace relations, agreement features may be carried forwards (from 'left-to-right') down the tree. It is this aspect of feature processing that we explore in Section 2.

2. *Feature-passing mechanisms*

In a fascinating squib, Kimball & Aissen (1971) make an interesting observation about their dialect (spoken by some speakers in the Boston area). The observation is summarized in (17). Kimball & Aissen find that a local subject-verb number mismatch in (17a) is acceptable for them if it is on the extraction path between the relative pronoun and its trace.

(17) a. (*)Lucine dislikes the people who the manager think know the answers.
b. Lucine dislikes the people who the managers think know the answers.
c. *Lucine dislikes the people who think the manager know the answers.
d. Lucine dislikes the people who think the managers know the answers.

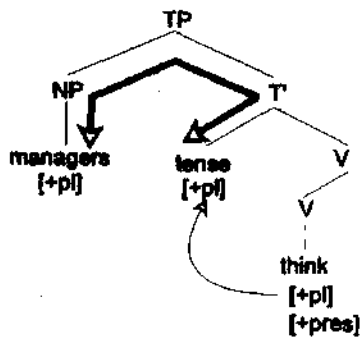
The mere presence of a c-commanding NP does not suffice, as illustrated by the obvious ill-formedness of (17c).

Kimball & Aissen were describing their own grammaticality/acceptability judgments. In their dialect, (17a) is wellformed, comparable to the expected form with *thinks* in place of *think*. We do not

share these grammaticality judgments nor do any of the subjects that we've tested (including 48 who were given a formal post-experiment interview in a pilot experiment conducted before Experiments 4 and 5). However, we suspect that the basis for the dialect Kimball & Aissen describe may derive from a more general processing effect that may hold even for dialects like ours where (17a) is ungrammatical but 'sounds better' than (17c).

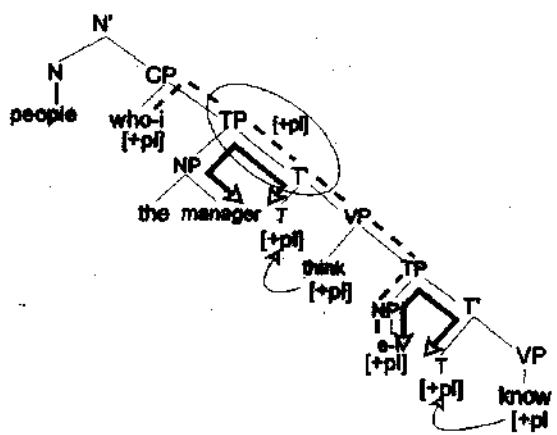
Here's our suspicion: perceivers check (subject verb) agreement in the phrase marker insuring that SPEC-head features are compatible, as illustrated in (18), see Deevy (1999) for details and evidence.

(18)



Now consider a sentence like (17a), illustrated in (19). When the processor checks the agreement features of *think*, there will be a plural feature in the local tree structure, circled in (19). But this feature does not arise from the feature set of *manager* (the subject of *think*).

(19)



Instead it is passed through the phrase marker along what we will call the 'feature transmission path' from *who* to the trace *who* binds (e_i). [4] Nevertheless, at points of high processing complexity, it may appear to the processor that a plural is available locally to satisfy the plural requirement imposed by *think*. Assuming that singular is the unmarked value of number (and that the processor does not fill in unmarked feature values), *the manager* will not introduce a contradictory feature specification to draw the processor's attention to the violation - singular is simply the default. (This assumption could be checked in future work to see if, as expected, a singular $who_i...e_i$ chain does not interfere with checking plural subject-verb agreement).

Let's dub the interference from the plural generated from the head of a chain the "distant plural" effect. The distant plural effect cannot be attributed to the mere presence of a higher plural when the lower subject-verb agreement between *manager* and *think* is checked. If the distant plural is marked on a chain that has already been completed before the lower subject-verb agreement is processed, then no effect is detectable using intuitions. Thus, in (17c) one readily detects the agreement violation. Indeed, it is the sensitivity of the distant plural effect to structure that most interests us. It suggests that the processor carries features through the phrase marker as illustrated in (19). [5]

2.1. Experiment 4

Experiments 4 and 5 are preliminary attempts to investigate the distant plural effect experimentally. Both experiments used a word-by-word grammaticality judgment technique. In Experiment 4, we tested materials like those in (17). Each sentence had four versions. In one (17a), a plural verb (*think*) had a singular subject but the nonagreeing subject and verb shared a link on the path connecting *who* and its trace in the subject of a sentential complement of the plural verb (*think*). In another version (17b), the subject of the plural verb was itself plural, so no agreement violation was present. In the remaining versions, the trace of *who* appeared before the plural verb (*think*). In one form (17c), the subject of the plural verb was singular and in another (17d) the subject of the plural verb was plural.

2.1.1. Method

16 sentences were constructed in four versions each, as shown in (17) (reproduced here, with positions of relevant empty elements and

number features indicated; all sentences appear in Appendix 2)

(17) a. (*)Lucine dislikes the people$_{i,pl}$ who$_{i,pl}$ the manager$_{sg}$ think$_{pl}$ e_i know$_{pl}$ the answers.
b. Lucine dislikes the people$_{i,pl}$ who$_{i,pl}$ the managers$_{pl}$ think$_{pl}$ e_i know$_{pl}$ the answers.
c. *Lucine dislikes the people$_{i,pl}$ who$_{i,pl}$ e_i think$_{pl}$ the manager$_{sg}$ know$_{pl}$ the answers.
d. Lucine dislikes the people$_{i,pl}$ who$_{i,pl}$ e_i think$_{pl}$ the managers$_{pl}$ know$_{pl}$ the answers.

These 16 experimental sentences and nonsentences were embedded in a list of 102 items total, of which 28 were clearly ungrammatical (e.g., subject-verb number disagreement, missing particles or arguments, extra arguments, etc.) or anomalous and another 20 which were marginally acceptable. Experimental subjects were first presented with an acceptability pretest in which they indicated the acceptability of a list of 10 sentences. The pretest contained a variety of ungrammatical and anomalous, as well as grammatical, items (but none similar to those in (17)). The very rare disagreements between a subject's judgment and the experimenter's judgment were discussed with the subject. The subject, after a short practice session, then participated in a word-by-word incremental acceptability judgment experiment. Each time the subject pressed a key with the right hand, the next word in an item would appear in a moving window, replacing underscore 'preview' marks (and the previous word would be replaced by underscore marks). The subject was instructed to continue pressing the right-hand key as long as the sentence remained acceptable, but to press a left-hand key as soon as it became unacceptable. Pressing the left-hand key terminated presentation of the sentence.

52 University of Massachusetts undergraduates were tested in individual sessions, receiving individually-randomized lists. Four different counterbalancing conditions were used so that each subject saw four experimental sentences in each version, and so that over all subjects, each sentence was tested equally often in each version.

2.1.2. Results

The primary data are the cumulative percentages of rejecting sentences, plotted word-by-word in Figure 2. Note that sentences like (17a) (late gap) and (17c) (early gap) become ungrammatical at the word *think* and *know*, respectively. The basic logic of the experiment is that the ungrammaticality in sentences like (17a) and (17c) may be

overlooked, at least briefly, if some factor interferes with noting the subject-verb number mismatch. Our hypothesis is that the shared links in the projection path of (17a) will be such a factor, predicting lower rejection rates for (17a) relative to (17c).

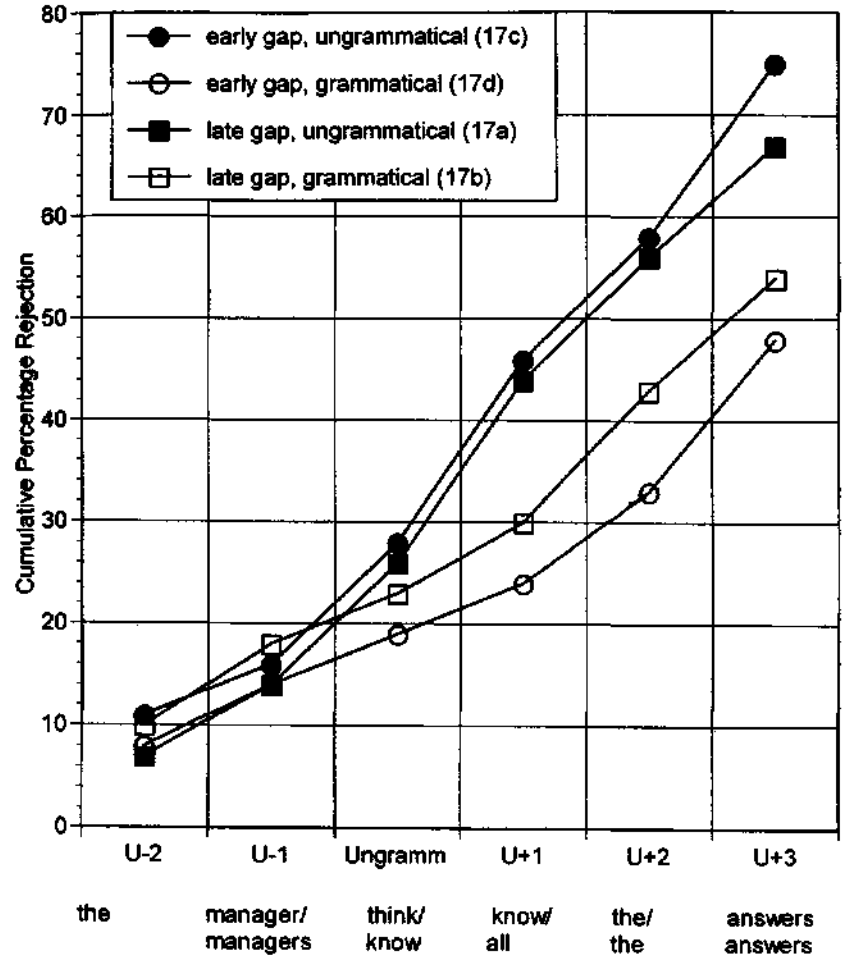


Figure 2. Mean cumulative percentages of rejecting items as unacceptable, Experiment 4, word by word, from two words before the (possible) ungrammaticality through three words after the point of ungrammaticality.

The data (see Figure 2) support this prediction. First, consider the total percentage of rejections at or before the point three words

following the point of ungrammaticality (by the end of most sentences). Ungrammatical sentences were rejected more often, 70% vs 55% ($F1(1,51) = 33.36$; $F2(1,15) = 32.56$). In addition, there was a nearly-significant interaction between grammaticality and late vs early gap ($F1(1,51) = 3.99$, $p < .06$; $F2(1,15) = 3.21$, $p < .10$). The main effect of grammaticality became significant at the point of ungrammaticality, while the interaction approached significance two words after the ungrammaticality.

The interaction takes the form of less frequent rejections of late-gap sentences like (17a) than early-gap sentences like (17c), contrasted with more frequent rejections of the grammatical late-gap sentences (17b) than early-gap sentences (17d). The former difference is significant in a by-subjects t-test three words after the ungrammaticality ($t1(51) = 1.99$) (but not in a by-items t-test; $t1(15) = 1.50$, $p > .10$). The difference between the two grammatical sentences is of lesser interest, and may simply reflect the longer path between filler and gap in (17b) than (17d).

2.1.3. Discussion

As expected, subjects in Experiment 4 tended to find it easier to reject (17c), with no still active plural dependency, than (17a) which exhibited the predicted distant plural effect. In general, rejection rates were high for these sentences. This is not terribly surprising given the complexity of the sentences and the intuitive difficulty of establishing a gap in an embedded complement's subject when the complement is embedded inside a relative clause, especially when the filler-gap path is long (and subjects are encouraged to detect all unacceptable sentences). The fact that the difference in acceptability between the two forms of grammatical items was opposite in direction to the predicted difference between the two ungrammatical forms suggests that this latter difference would be even larger if it weren't confounded with a difference in length of the filler-gap path.

The difference among the rejection rates for the two types of ungrammatical sentences appeared not at the point of ungrammaticality but two words downstream at the end of the sentence. We cannot be sure why the difference appeared at this point rather than earlier. Possibly it is due to the rhythm subjects get into when they proceed through a sentence with word-by-word self-paced presentation of the sentence.

We must emphasize that we take the results of Experiment 4 to reflect a processing confusion that arises because a shared link of the

projection path from the verb to its local subject has a multiple or contradictory specification of the value for the number feature due to the plurality of the unassigned filler. The results cannot be attributed to the grammaticality-status of (17a) given that our subjects, like us, uniformly classify these sentences as being ungrammatical when off-line judgments are elicited, as in a post-experiment questionnaire administered to 48 subjects of a pilot study. In other words, sentences like (17a) are classified as being ungrammatical; nevertheless they contrast with the equally-ungrammatical sentences like (17c) presumably due to the processing difficulty of checking the source of the plural specification on the link of the projection path between TP and T' (the link shared by the local subject and verb and by the path between *who* and its trace).

However, as Marcus Bader (personal communication) pointed out, one could interpret the Experiment 4 results as a straightforward complexity effect, assuming that late gap sentences are more difficult to process accurately. This is more readily seen if one plots the results in terms of percentage correct, not percentage rejections: Early gap sentences are 'more accurate' than late gap sentences, both for ungrammatical and for grammatical items. This suggestion assumes that the complexity of late gap items resulted in failure to detect the agreement violation in the ungrammatical sentences, and (presumably) failure to process properly some of the legitimate grammatical relations of the grammatical sentences. One difficulty with the suggestion, however, is that the ungrammatical sentences have a large number of legitimate grammatical relations (apart from the subject-verb number mismatch) that could have been misprocessed in the late gap condition. Such misprocessing would have resulted in a rejection of the late gap ungrammatical sentences (as it presumably did in the late gap grammatical sentences). This would have increased the observed frequency of rejections of the late gap ungrammatical sentences, making them more nearly comparable to early gap ungrammatical sentences than was observed.

Experiment 5 provides a more direct test of the importance of the disruptive influence of a distant plural feature being passed through the portion of a tree used for checking a subject-verb agreement, as opposed to any possible late gap complexity effects. It tests for such disruption by comparing whether the sentence does or does not introduce a plural feature to pass through the relevant portion of the tree, rather than comparing late vs. early gap position.

2.2. Experiment 5

Experiment 5 tested the prediction of our account of the distant plural effect that the difficulty of detecting an ungrammaticality should disappear when no plural feature is introduced early in the sentence to interfere with checking a later subject-verb agreement. Experiment 5 tested sentences like those in (20), which compared plural and singuar *who*.

(20) a. (*)Lucine dislikes the people$_{i,pl}$ who$_{i,pl}$ the manager$_{sg}$ think$_{pl}$ e_i know$_{pl}$ the answers.
b. Lucine dislikes the people$_{i,pl}$ who$_{i,pl}$ the managers$_{pl}$ think$_{pl}$ e_i know$_{pl}$ the answers.
c. *Lucine dislikes the person$_{i,sg}$ who$_{i,sg}$ the manager$_{sg}$ think$_{pl}$ e_i knows$_{sg}$ the answers.
d. Lucine dislikes the person$_{i,sg}$ who$_{i,sg}$ the managers$_{pl}$ think$_{pl}$ e_i knows$_{sg}$ the answers.

Sentence (20a) is the same as (17a) and, as before, we expect a distant plural effect. Sentence (20c) has a singular antecedent for *who* and thus subjects should find it easier to detect the number agreement violation in (20c) than in (20a). (20b) and (20d) are the grammatical counterparts of (20a) and (20c) where no agreement violation occurs. (20a,b) are the same as (17a,b), but (20c,d) introduce singular NPs as head of the relative clause while maintaining the late gap structure of (20a,b).

2.2.1. Method

The sixteen sentences of Experiment 4 were modified to make sentences like (20). Forty-eight University of Massachusetts undergraduates were tested (none of whom had participated in Experiment 4). Otherwise, the procedures were identical to those described in Experiment 4.

2.2.2. Results

Figure 3 presents the cumulative percentages of rejections. Of most interest, the interaction between grammaticality and singular/plural was significant three words after the point of ungrammaticality ($F_1(1,47) = 5.78$; $F_2(1,15) = 5.60$), and in fact was significant at each region from one word after the ungrammaticality through the end of the sentence. Except for the two words immediately after the ungrammaticality, the difference between (20a) and

(20c) was significant or nearly so at each word by both subjects and items (e.g., three words after the point of ungrammaticality, $t_1(47) = 2.00$, $p < .05$; $t_2(15) = 1.56$, $p < .15$). Oddly, the difference was nearly significant in the opposite direction one word after the ungrammaticality ($t_1(51) = 2.92$, $p < .01$; $t_2(15) = 1.95$, $p = .07$).

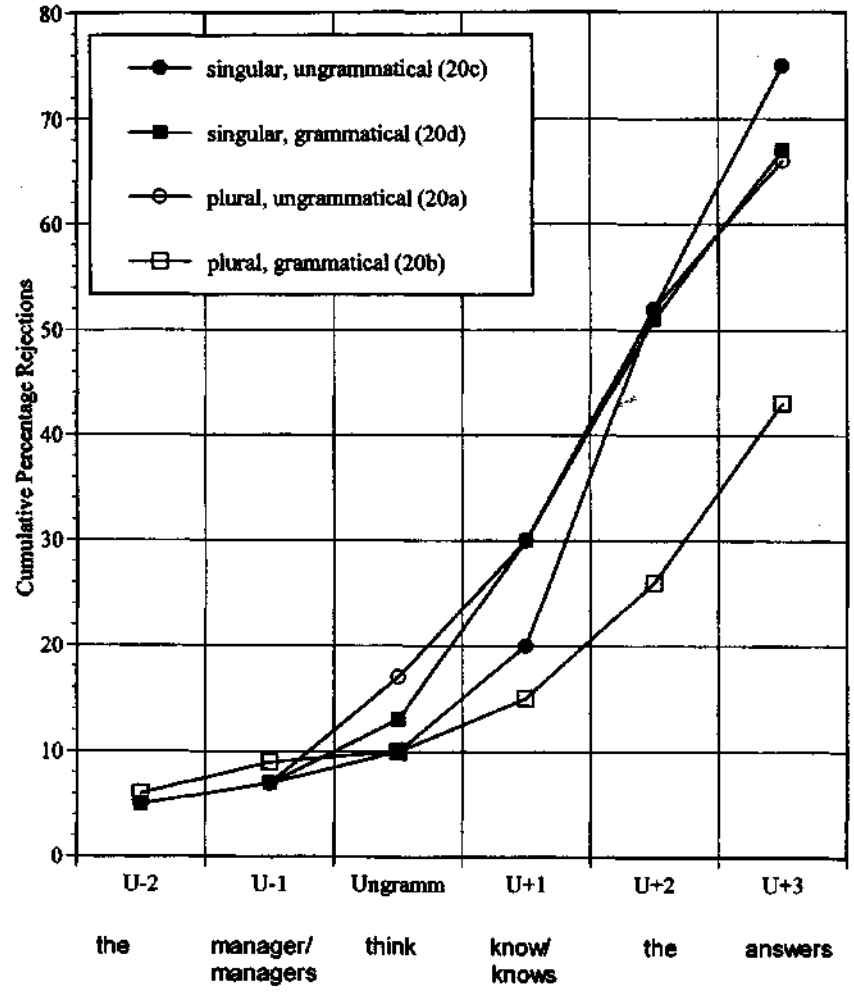


Figure 3. Mean cumulative percentages of rejecting items as unacceptable, Experiment 5, word by word, from two words before the (possible) ungrammaticality through three words after the point of ungrammaticality.

In addition, the main effects of grammaticality and singular/plural were significant at all regions two or more words after the ungrammaticality. For example, three words after the ungrammaticality 70% of the ungrammatical and 55% of the grammatical sentences had been rejected ($F_1(1,47) = 24.16$; $F_2(1,15) = 9.08$). Similarly, at the same point, the plural sentences were being rejected 55% of the time while the singular sentences were rejected 71% of the time ($F_1(1,47) = 17.46$; $F_2(1,15) = 22.79$). This latter effect is due in part to the higher rejection rate for singular ungrammatical sentences like (20c) than for plural ungrammatical sentences like (20a), a difference which we predicted. However, it also reflects the very high rate of rejecting singular grammatical sentences like (20d).

2.2.3. Discussion

Experiment 5 again revealed a distant plural effect: Subjects had difficulty rejecting ungrammatical sentences like (20a) compared to ones like (20c), perhaps because an interfering feature was carried along a common projection path only in the former. However, the experiment also showed that sentences like (20d) were often rejected. We conjecture that sequences of verbs that must be analyzed as constituents of separate clauses are difficult for perceivers or, equivalently, subjects find it difficult to use the second of two adjacent verbs to (i) postulate an embedded complement (at least inside a relative clause), (ii) project a subject position for the embedded clause headed by the second verb, (iii) postulate a trace in that subject position and (iv) coindex the trace with its binder *who*. An alternative account of the difficulty of (20d) is that the processor does not pass the unmarked singular feature of *person* down the tree in (20d), and that the only recently-experienced feature to check against the singular verb *knows* is the plural feature of *managers think*. This account assumes, contrary to our claims, that the plural feature of *think* is not discharged when it is checked against the plural *managers* but instead remains available to the processor. We cannot securely reject this account, but do note that its structure-independent nature provides no explanation of the difference between (17a) and (17c) in Experiment 4. We also note that the account must claim that sentences like (20d) should begin being rejected one word later than the ungrammatical sentences (20a,c). Figure 3 provides no support for this claim; any delay in rejection seems to be associated not with (20d) but with (20c), which (for reasons we do not understand) is less often rejected one word after the ungrammatical verb than (20a).

3. Conclusions

It has been claimed that Phi features are used in a direct feature-matching process before the constraints from Binding Theory are consulted (Straub & Badecker 1994). The experiments in Section 1 directly disconfirm this hypothesis. Instead, we suggest, features of pragmatically tempting antecedents are checked, but only if they are in syntactically accessible positions.

Section 2 investigated a different aspect of feature manipulation to determine whether local agreement processes (checking subject-verb agreement) are influenced by the Phi features passed from the head of a chain to its tail. Experiments 4 and 5 suggest that there is such an influence and they revealed experimental evidence for the existence of difficulty when a link of the tree must carry number information for two phrases which conflict in number. We have dubbed this the 'distant plural effect'. Further study of the distant plural effect is warranted, we think, because it has the potential to reveal just how features are manipulated during sentence processing.

In particular we wish to confirm that the distant plural effect arises not simply because the features of a 'filler' (moved) constituent are active until its chain is completed but because of the link of the projection path that is shared by two distinct syntactic dependencies. In either case, however, it is clear that feature-use is highly structure dependent. Features do not act like 'pointers' directly granting memory access to phrases with compatible feature specifications.

The grammatically-regulated expression of features can be influenced by features of grammatically-irrelevant phrases. This type of "feature contamination" has been documented in many production studies (e.g., Bock & Eberhard 1993, Eberhard 1997, Nicol *et al.* 1997, Nicol & O' Donnell, this volume). The fact that features can have grammatically unregulated effects is of great potential interest. Whether the pattern of contamination effects will prove to be similar in production and comprehension remains to be determined.

Addresses of the Authors:

Charles Clifton: Department of Psychology, University of Massachusetts, Amherst, MA 01003 (USA), e-mail: *cec@psych.umass.edu*
Lyn Frazier: Department of Linguistic, University of Massachussetts, Amherst, MA 01009 (USA)
Patricia Deevy: Department of Audiology and Speech Sciences, Purdue University, West Lafayette, IN 47907 (USA), e-mail: deevy@purdue.edu

APPENDIX 1

MATERIALS FOR EXPERIMENT 1. Optional material indicated by [ ]

The son [of the pretty nurse] / hurt [[himself | Fred] in a bad accident.
The mother [of the man] / was talking to [herself | Sally] outside.
The daughter [of the fireman] / shot [herself | Brenda] last year.
The brother [of the woman] / committed [himself | Tom] recently.
The girlfriend [of the Congressman] / killed [herself | Marcia] this morning.
The mistress [of the executive] / was amusing [herself | Tricia] on the roof.
The bodyguard [of the actress] / invited [himself | Melissa] to the party.
The nephew [of the ballerina] / admired [himself | Patrick] in the mirror.
The niece [with the weird boyfriend] / hid [herself | Sandra] in the TV room.
The waitress [with the handsome friend] burned [herself | Sarah] after spilling the soup.
The postman [with the lady] / boasted about [himself | Stanley] for hours.
The pretty woman [with the bearded doctor] / watched [herself | Linda] in the mirror.
The kind girl [with the plumber] / sang to [herself | Amy] loudly.
The guy [with the girl] / embarrassed [himself | Samuel] in the park.
The man [with the actress] / was trying to teach [himself | Roger] Spanish.
The cameraman [with the starlet] / was grooming [himself | Fido] in the hall.

APPENDIX 2

MATERIALS FOR EXPERIMENT 4. Optional plural indicated by (s).Conditions 1 and 2 shown before the | ; Conditions 3 and 4 after

Lucine dislikes the people who the manager(s) think know the answers. | Lucine dislikes the people who think the manager(s) know all the answers.
Kathy already met the administrators who the dean(s) argue deserve strong support. | Kathy already met the administrators who argue the dean(s) deserve much stronger support.
The reporter interviewed repeatedly the janitors who the student(s) claim promote the union. | The reporter interviewed repeatedly the janitors who claim the student(s) promote the union actively.
Martin questioned again the managers who the researcher(s) claim ignore the needs of working mothers. | Martin questioned again the managers who claim the researcher(s) ignore the needs of working mothers.
Patrick didn't like the visitors who the maid(s) say want to move here. | Patrick didn't like the visitors who say the maid(s) want to move near here.
Stanley will encounter the guides who the scout(s) think carry supplies. | Stanley will encounter the guides who think the scout(s) carry extra supplies.
Kathy remembered fondly the teachers who the parent(s) claim behave irresponsibly. | Kathy remembered fondly the teachers who claim the parent(s) behave quite irresponsibly.
The director already fired the supervisors who the worker(s) maintain merit a raise. | The director already fired the supervisors who maintain the worker(s) merit a big raise.
Karen particularly likes the explorers who the historian(s) think act impulsively. | Karen particularly likes the explorers who think the historian(s) act too impulsively.
Mr. Gould chastised the children who the neighbor say make noise at night. | Mr. Gould chastised the children who say the neighbor(s) make loud noise at night.
Molly really despises the politicians who the mayor(s) say support welfare cuts. | Molly really despises the politicians who say the mayor(s) support unfair welfare cuts.
Sally befriended the newcomers who the boy(s) pretend come from space. | Sally befriended the newcomers who pretend the boy come from outer space.
Sam never actually met the relatives who Mother's cousin(s) claim live in Warsaw. | Sam never actually met the relatives who claim Mother's cousin(s) live in Warsaw now.
Tom enjoys thoroughly the guests who the local grocer(s) think know everything about gardens. | Tom enjoys thoroughly the guests who think the local grocer(s) know everything about herb gardens.
John will hate the inmates who the guard(s) assume sell drugs. | John will hate the inmates who assume the guard(s) sell hard drugs.
Max admires wholeheartedly the artists who the critic(s) argue know nothing. | Max admires wholeheartedly the artists who argue the critic(s) know almost nothing.

NOTES

* The research reported in this paper was supported in part by Grant HD-18708 and by Training Grant HD-07327 to the University of Massachusetts. Because of their length, the materials for Experiments 2, 3, and 5 are not presented in the appendices, but may be obtained from the second author (cec@psych.umass.edu).

[1] Nicol presented data suggesting that gender features limited access to a gender-appropriate antecedent in a grammatically-appropriate position, but her evidence was actually limited to masculine pronouns. Other data suggest that the use of gender information to select an antecedent may be delayed relative to the use of number feature (Di Domenico & Di Vincenzi 1996) or that the use of gender information may be strategic (Garnham, Oakhill & Cruttenden 1992; cf. McDonald & Mac Whinney 1995, for discussion). However, these possibilities do not qualify the arguments we present in this paper, which rest on the failure to find 'decoy' effects from potential antecedents in grammatically-inaccessible positions, regardless of gender (and success in finding structure-dependent effects of passing number features through a tree).

[2] In later, as yet published, work, they report finding effects on the pronoun and the following word as well.

[3] A reviewer points to data reported in this volume which may be considered an example of structure-independent feature matching in processing. Nicol & O'Donnel (this volume) report effects of interference from a structurally inaccessible gender or number mismatching NP in producing a tag question. Neither Straub & Badeker's (1994) linear feature matching account nor our binding theory account of the identification of the antecedents of an anaphor predicts this effect and neither directly tested

for it. We would not expect the effects Nicol observed to carry over to the comprehension of anaphor-antecedent relations since the exact source of the difficulty in the production study is unclear. As was suggested for the Badeker & Straub results, it could be that the demands of the task – in this case, listening to a sentence (presumably building a grammatical representation and interpreting it), repeating it (mapping that interpretation back into a grammatical representation) and producing the tag question based on that sentence – led to an error which does not reflect normal antecedent-anaphor processing.

[4] The results reported in Section 2 might be taken to support the view that a wh-dependency is represented as a path rather than as a chain. For example, in (i) the dependency between who and its trace might be viewed as the path consisting of the nodes dominating the trace up to and including the node dominating who, as illustrated in (ia).

(i) Who_i did John see e_i?
a. {CP,C',TP,T',VP,V',e} path
b. {Who_i, e_i} chain

Alternatively, the wh-dependency in (i) might be represented as the chain consisting of the interrogative constituent and all phrases which it binds, as illustrated in (ib). In fact, we think that chains are necessary for the representation of wh-dependencies (see De Vincenzi 1991, 1995). Further, we suspect that marked syntactic features, not all features of the moved constituent, are passed down the tree. However, these issues go beyond the scope of the current paper. We are pursuing them in related studies that are now in progress.

[5] Further research is needed to securely reject the possibility that simply having an unassigned ('active') filler is enough to distract the processor, creating a possible confusion between the features of the filler and the features of the subject for an agreeing/disagreeing verb. To demonstrate that it is critical that the dual number specification involves at least one shared link on the projection path, we intend to test structures where the violation is in a structure lying off the projection path, eg., an embedded or parenthetical structure.

References

ALBRECHT, Jason & Charles CLIFTON Jr. (1998), "Accessing singular antecedents in conjoined phrases", *Memory & Cognition* 26, 599-610.

BADECKER, William & Kathleen STRAUB (1994), *Evidence that binding principles participate in a constraint satisfaction process*, Poster presented at the 7th Annual CUNY Sentence Processing Conference, New York.

BOCK, Katherine & Kathleen M. EBERHARD (1993), "Meaning, sound and syntax in English number agreement", *Language and Cognitive Processes* 8:57-100.

CHOMSKY, Noam (1981), *Lectures on Government and Binding*, Dordrecht, Foris.

DEEVY, Patricia (1999), *The Role and Representation of Agreement Features in Sentence Processing*, PhD Dissertation, Linguistics Department, University of Massachusetts at Amherst.

DE VINCENZI, Marica (1991), *Syntactic Parsing Strategies in Italian*, Dordrecht, Kluwer Academic Publishers.

DE VINCENZI, Marica (1996), "Syntactic analysis in sentence comprehension: Effects of dependency types and grammatical constraints", *Journal of Psycholinguistic Research* 25:117-134.

DI DOMENICO, Elisa & Marica DE VINCENZI (1996), "Gender and number in the retrieval of pronoun antecedents: Differences in use and representation", in L. NASH, G. TSOULES & A. ZRIBI- HERTZ (eds)., *Actes du deuxième colloque 'langues et grammaire'*, Paris, 8 Juin 1995:95-109.

EBERHARD, Kathleen M. (1997), "The marked effect of number on subject-verb agreement", *Journal of Memory and Language* 36:147-164.

GARNHAM, Alan, Jane OAKHILL & Hannah CRUTTENDEN (1992), "The role of implicit causality and gender cue in the interpretation of pronouns", *Language and Cognitive Processes* 7:231-255.

GORDON, Peter C. & Randall HENDRICK (1997), "Intuitive knowledge of linguistic coreference", *Cognition* 62:325-370.

HEIM, Irene (1982), *The Semantics of Definite and Indefinite Noun Phrases*, Unpublished Doctoral dissertation, Linguistics Department, University of Massachusetts at Amherst.

KIMBALL, John & Judith AISSEN (1971), "I think, you think, he think", *Linguistic Inquiry* 2:241-246.

MCDONALD, Janet & Brian MACWHINNEY (1995), "The time course of anaphor resolution: Effects of implicit verb causality and gender", *Journal of Memory and Language* 34:543-566.

NICOL, Janet (1988), *Coreference Processing during Sentence Comprehension*, MIT doctoral dissertation.

NICOL, Janet & Meghan O'DONNELL (1999), "Pronominal feature distinctions in English", *This volume*.

NICOL, Janet, Kenneth I. FORSTER & C. VERES (1997), "Subject-verb agreement processes in comprehension", *Journal of Memory and Language* 36:569-587.

STRAUB, Kathleen & William BADECKER (1994), *The role of grammar in processing pronouns and reflexives*, Paper presented at Psychonomic Society, St. Louis, MO, November 11-13:1994.