COGNITIVE PSYCHOLOGY **37,** 201–242 (1998)
ARTICLE NO. CG980693

Repeating Words in Spontaneous Speech

Herbert H. Clark and Thomas Wasow

Stanford University

Speakers often repeat the first word of major constituents, as in, ''I uh I wouldn't be surprised at that.'' Repeats like this divide into four stages: an initial commitment to the constituent (with ''I''); the suspension of speech; a hiatus in speaking (filled with ''uh''); and a restart of the constituent (''I wouldn't . . .''). An analysis of all repeated articles and pronouns in two large corpora of spontaneous speech shows that the four stages reflect different principles. Speakers are more likely to make a premature commitment, immediately suspending their speech, as both the local constituent and the constituent containing it become more complex. They plan some of these suspensions from the start as preliminary commitments to what they are about to say. And they are more likely to restart a constituent the more their stopping has disrupted its delivery. We argue that the principles governing these stages are general and not specific to repeats.

Spontaneous speech is filled with disfluencies—unwanted pauses, elongated segments, fillers (such as *uh* and *um*), editing expressions (such as *I mean* and *you know*), word fragments, self-corrections, and repeated words. Most disfluencies seem to reflect planning problems. When speakers cannot formulate an entire utterance at once, they may suspend their speech and introduce a pause or filler before going on. And when speakers change their minds about what they are saying, they may suspend their speech and then add to, delete, or replace words they have already produced. Disfluencies have long been used as evidence of planning (e.g., Clark, 1996; Goldman-Eisler, 1968; Levelt, 1983, 1989; Maclay & Osgood, 1959; Schegloff, Jefferson, & Sacks, 1977).

In this paper we investigate the origins of repeated words. Consider an utterance in which Reynard is speaking to Sam:

This research was supported by NSF Grants SBR-9309612 and IRI-9314967 and by ATR. We thank Erica Don, Yafeng Li, and Jon Lindsay for their help in the analyses and Eve V. Clark, Gary S. Dell, Jean E. Fox Tree, Antje S. Meyer, Padraig O'Seaghdha, Elizabeth E. Shriberg, and several unnamed reviewers for their suggestions.

Address correspondence and reprint requests to Herbert H. Clark, Department of Psychology, Stanford University, Stanford, CA 94305-2130, or Thomas Wasow, Department of Linguistics, Stanford University, Stanford, CA 94305-2150.

0010-0285/98 $25.00

(1) yes, I uh I wouldn't be surprised at that, – – I really wouldn't, (1.1.278).[1]

After Reynard produces ''I uh,'' he could have continued ''wouldn't be surprised at that,'' but he repeats *I* first. The puzzle is why. Repeating *I* takes extra time and effort. It is redundant. And by most accounts it ought to make the utterance harder to understand, since there is no English clause of the form *I I wouldn't be surprised at that.* Speakers seem to have good reasons for *not* repeating words, yet they often do. Repeated words are one of the most common disfluencies in spontaneous speech (Deese, 1984; Maclay & Osgood, 1959).

Disfluencies have been viewed from two perspectives. In one tradition, they are treated mainly as the outcome of processes that, once initiated, run off without intervention. We will call these *pure* processes. In (1), for example, Reynard might have repeated *I* because it was the most highly activated word when he resumed speaking after *uh,* and he couldn't help but produce it. Accounts in the process tradition tend to eschew appeals to intentions, purposes, or monitoring. In a second tradition, disfluencies are viewed mainly as the result of certain *strategies*—processes with options under a person's control. In (1), Reynard might have repeated *I* because, in the words of Maclay and Osgood (1959), he wanted to ''produce some kind of signal ([m, er], or perhaps a repetition of the immediately preceding unit) which says, in effect, 'I'm still in control—don't interrupt me!' '' (p. 59). Accounts in the strategy tradition generally do appeal to intentions, purposes, and monitoring.

These two traditions, however, offer complementary, not conflicting perspectives on disfluencies. Most pure processes are deployed in the service of strategies—ultimately, what speakers are trying to do by speaking. At the same time, no strategy can work without deploying pure processes. The contrast partly reflects the evidence appealed to. Pure processes have generally been studied in controlled laboratory speech, where speakers have few if any options. Strategies have generally been suggested for spontaneous speech, where speakers have a plethora of options and the opportunities for taking them. We will focus on strategies while taking note of the relevant pure processes.

In this paper we propose a *commit-and-restore model* of repeated words. We first lay out the model, an extension of the model of repairs from Levelt (1983, 1989), and describe three hypotheses that follow from it. We then test the hypotheses as they apply to English articles and personal pronouns.

[1] All of the examples we cite are from one of two corpora, described later. Those from the London-Lund (LL) corpus (Svartvik & Quirk, 1980) are identified by the conversation (1.1) and line (278) they came from. In these examples, the end of a tone unit is marked by a comma (,), a ''brief pause (of one light foot)'' is marked by a period (.), a ''unit pause (of one stress unit)'' is marked by a hyphen (-), and elongated vowels are marked by a colon (:). The rest of the examples come from the switchboard (SW) corpus.

The evidence we use comes from two large corpora of spontaneous speech, one American and one British.

COMMIT-AND-RESTORE MODEL OF REPEATED WORDS

Repeating a word is often treated as an unanalyzable event (e.g., Deese, 1984; Holmes, 1988), but is really a sequence of processes, each with its own options and limitations (Clark, 1996; Levelt, 1983). Here we divide repeats into four stages.

Stage I: Initial Commitment

When speakers produce a word, they are ordinarily committing themselves to one or more constituents containing that word and to meaning something by them. Consider (2), another utterance by Reynard:

(2) I thought it was be*fore* sixty-five, (1.1.244).

When Reynard produces *I,* he is committing himself to producing a larger constituent that begins with *I* and to meaning something by it for Sam, his addressee. Sam can expect him to complete it, unless he is told otherwise. Making such a commitment is both constrained and optional. It is constrained by the *formulation imperative:* Speakers cannot produce an expression until they have formulated it completely (Clark, 1996). On the other hand, Reynard could have delayed his commitment (delaying ''I''), produced a filler (e.g., ''uh''), or made an alternative commitment (''well''). So, even though making a commitment is constrained by the formulation process, it is a strategy speakers can use for particular purposes. When Reynard produces *I* in (1), he makes the same commitment, even if he suspends his speech immediately afterwards.

Stage II: Suspension of Speech

Speakers can in principle suspend their speech at almost any point in an utterance (Levelt, 1989). Consider (3), by Sam:

(3) because you see I {- uh} some of our people, {. (clears throat)} who are doing LEs, {- - u:m} have to consider which paper {.} to do, (1.1.39).

For purposes of exposition, we will label each pair of suspensions and resumptions—each disruption—with left and right curly brackets (Clark, 1996). In (3) Sam suspends his speech four times and apparently for different causes. He stops after *I* to replace it with *some of our people who are doing LEs.* Such a suspension, as Levelt (1989) has argued, is strategic, because it depends on the type of repair the speaker has to make. Sam stops after *people* perhaps to clear his throat. He also stops after *LEs* and *paper,* perhaps because he hasn't yet formulated what he wants to say. Suspending speech isn't specific to repeats. It occurs at many points and for many reasons.

Stage III: Hiatus

The hiatus is the material between the suspension and the resumption of speech—the material between the curly brackets. Speakers may do a variety of things in a hiatus, from nothing to adding fillers or clearing their throat. In (1), Reynard filled the hiatus with ''uh,'' although he had the option of remaining silent for the same length of time. In (4), the hiatus contains nothing, not even a pause:

(4) well I {} I get rather fed up of some of these youngsters, (1.1.768).

Speakers' options in dealing with the hiatus are also not tied to repeats.

Stage IV: Restart of Constituent

When speakers resume speaking after a hiatus, they have many options. Consider (5), another utterance by Sam:

(5) I suppose if I {uh} get more expensive ones, they'll be {.} safer, (1.1.467).

When Sam resumes speaking after ''I uh,'' he simply continues. His choice contrasts with Reynard's in (1), which is to repeat *I*. In the cases like (1), (4), and (5), speakers appear to have two main options: (a) they can *restart* one of the constituents they interrupted; or (b) they have can *continue* where they left off. Repeats arise when speakers take the first option. Speakers, of course, cannot resume speaking until they have something formulated, but they have the option of delaying as long as they wish. Restarting at the beginning of constituents is characteristic of repairs and what are called fresh starts (Levelt, 1983; Maclay & Osgood, 1959). So it, too, is a general process and not tied to repeats.

All four of these processes—initiating constituents, suspending speech, dealing with hiatuses, and restarting constituents—occur in a variety of circumstances. It is their combination that leads to repeated words. If we are to account for repeats, we must account for their combination. We now turn to three hypotheses about the sources of repeats.

Constituent Complexity

In the commit-and-restore model, repeats arise as speakers are trying to produce constituents, especially major ones such as noun phrases (NPs), verb phrases, prepositional phrases, clauses, sentences. Constituents such as these have long been thought to be principal units of planning (Bock & Levelt, 1994; Ford, 1982; Goldman-Eisler, 1968; Holmes, 1988; Levelt, 1989; Maclay & Osgood, 1959). At a conceptual level, speakers choose the message they wish to express, roughly one major constituent at a time. At a syntactic level, they select the functions and arguments needed for that message, including a syntactic framing. At a phonological level, they formulate the phonological words and phrases needed for pronunciation, but for smaller

constituents at a time (Ferreira, 1991; Meyer, 1996; Wheeldon & Lahiri, 1997). These three levels overlap. Speakers generally begin producing larger constituents while they are still formulating the later parts of these constituents.

If speakers find it difficult to plan major constituents, they should have problems starting them up, and they do. They are most likely to pause *before* the first word of such units, next most likely *just after* the first word, and less likely after that (Boomer, 1965; Chafe, 1979, 1980; Ford, 1982; Holmes, 1988; Maclay & Osgood, 1959). According to one account of these findings (Ford, 1982; Holmes, 1988), speakers have difficulty planning so-called basic clauses, those with either a tensed or untensed verb. There is no added difficulty in planning surface clauses or other constituents, so difficulty is categorical. According to another account (Ferreira, 1991; Wheeldon & Lahiri, 1997), it takes speakers longer to initiate complex than simple constituents in part because it takes them longer to create articulatory plans for complex than for simple constituents (see also Meyer, 1996).

Our proposal is that constituents are harder to plan at the conceptual or syntactic level the greater their *grammatical weight.* Grammatical weight is roughly the amount of information expressed in a constituent (Behaghel, 1909/1910; Hawkins, 1994; Wasow, 1997). It can be measured by the number of words, syntactic nodes, or phrasal nodes in the constituent; these numbers correlate with each other at .94 and beyond (Wasow, 1997). Weight has long been known to play a role in production. When speakers have the option, they tend to place lighter constituents before heavier ones. Consider Susan's utterance in (6):

(6) the first European conference on astronomy at Leicester, reported [yesterday morning], - [on overnight observations of the behaviour of the object, - . known as A six uhu two one one zero], (1.11a.28).

Susan produces the lighter of the two bracketed constituents (''yesterday morning'') before the heavier. She would have produced the second one first if it had been of the same weight or lighter. Evidence from spontaneous speech and writing shows that the choice of ordering is based not on absolute weight, but on relative weight—on which of two constituents is heavier (Wasow, 1997). The hypothesis, then, is that many suspensions are prompted by planning difficulties at the conceptual or syntactic level:

The complexity hypothesis. All other things being equal, the more complex a constituent, the more likely speakers are to suspend speaking after an initial commitment to it.[2]

We will refer to grammatical weight simply as complexity. The complexity

[2] Compare Maclay and Osgood (1959, p. 42): ''The larger the unit being 'programmed' . . . the more prolonged the non-speech interval [before the unit] and hence the greater the tendency for an 'ah' or a repetition.''

hypothesis is really a claim about process limitations—about when speakers are likely not to be able to proceed.

The complexity of constituents, in this account, is both graded and hierarchical. Consider Sam's utterance in (7):

(7) this English Language paper, has been bedeviled long enough, by those literature wallahs, (1.1.845).

The word *this* is at the left edge of the NP *this English Language paper*, and that in turn is at the left edge of the clause *this English Language paper has been bedeviled long enough by those literature wallahs.* The word *this* is therefore at the left edge of *both* the NP *and* the clause. In utterances like (7), suspensions after *this* should increase with the complexity of both the NP and the clause (cf. Ferreira, 1991; Wheeldon & Lahiri, 1997). The word *those,* in contrast, is at the left edge of the NP *those literature wallahs,* which is *not* at the left edge of a larger constituent. In utterances like (7), suspensions after *those* should increase with the complexity of the NP even though it is in the middle of the clause. These predictions contrast with the idea that syntactic complexity is categorical and not hierarchical (Ford, 1982; Holmes, 1988).

For speakers to repeat the part of a constituent prior to the suspension, that part must be accessible both at stage I and at stage IV. Let us call this the *accessibility hypothesis.* Important as the hypothesis is, we have only limited ways of testing it in this paper.

Continuity of Delivery

Speakers may initiate a constituent, suspend their speech, and delay (stages I, II, and III) and still not restart the constituent (stage IV). Without a restart there is no repeat. Why, then, do speakers restart (''I uh I wouldn't be . . .'') rather than continue (''I uh wouldn't be . . .'')? According to several accounts (Fox, Hayashi, & Jasperson, 1996; Levelt, 1983, 1989; Schegloff, 1979), it is to make a repair—what Levelt called a covert repair. But a repair of what? ''The repair is called 'covert' because we don't know what was being repaired'' (Levelt, 1989, p. 478); ''in fact the reason for the repair is not obvious'' (Fox et al., 1996). Then why restart? These accounts offer reasons why speakers might suspend their speech (stage II) or delay (stage III), but *not* why they should restart a constituent rather than continue it.

Our proposal is that speakers restart a constituent in order to *restore continuity to its delivery* after the disruption caused by the suspension and hiatus (stages II and III). Note that after a continuation, the final delivery of the constituent has a gaping hole in it (''I {uh} wouldn't be surprised at that''), whereas after a restart it is continuous (''I wouldn't be surprised at that''). Our hypothesis is this:

The continuity hypothesis. All other things being equal, speakers prefer to produce constituents with a continuous delivery.

The continuity hypothesis reflects the notion of *ideal delivery* (Clark &

Clark, 1977). For a phenomenon to be called a disfluency, there must be one way of delivering an utterance that is considered appropriate to the circumstances, and that is the ideal delivery. Repeating a word is an attempt to redo a constituent in its ideal delivery. In this view, repeats are a type of repair, but not of covert or unspecified troubles. They repair the conspicuous disruption that has just occurred to the delivery of the current constituent.

The continuity hypothesis is consistent with many past observations about spontaneous speech. One is that speakers are more likely to pause between than within constituents (Maclay & Osgood, 1959; Boomer, 1965), and the more careful the speech, the fewer pauses and fillers there are within constituents (Goldman-Eisler, 1968). Another observation is that when speakers repair a content word, they often return to a major constituent boundary before that word (Levelt, 1983; Maclay & Osgood, 1959), as here:

(8) I heard his name mentioned by {-} Carter, {I think,} by Darlington, while I was down there, (1.1.585).

Sam doesn't just replace *Carter* by *Darlington.* He adds *by,* which restores continuity to the prepositional phrase *by Darlington.*

Why might speakers prefer a continuous delivery? We can think of at least three reasons. The first is process limitations. When speakers resume speaking after a hiatus, they may find it easier to formulate and produce a constituent from the beginning (''I wouldn't be surprised at that'') than from the middle (''wouldn't be surprised at that''). Producing the complete constituent may help them keep track of where they are. The second and third reasons are strategic. Speakers may be attentive to their addressees. They realize that constituents are easier to parse and understand when they are intact than when they are disrupted. Or speakers may want to present themselves as prepared, thoughtful, and articulate, and disrupted constituents count against that impression. All three reasons may apply, but we won't be able to distinguish among them.

One alternative to the continuity hypothesis is what we will call the *activation hypothesis:* When speakers resume speaking after a hiatus (after ''uh'' in (1)), they tend to repeat the last word produced (*I*) because it is the most highly activated word at that moment. This hypothesis has problems *a priori.* If the last word produced is still the most highly activated word available, why don't speakers *always* repeat it, perhaps forever? To prevent this, many models of the production (e.g., Dell & O'Seaghdha, 1992; MacKay, 1987; Shattuck-Hufnagel, 1979) make the opposite assumption: Once a word has been produced, its activation gets reset to its resting level or below. Let us call this the *deactivation hypothesis. A priori,* this hypothesis has the opposite problem. It predicts that speakers should rarely if ever repeat a word, whereas repeats are common. For the activation hypothesis to work at all, it must follow the Goldilocks principle: The activation cannot be too hot, or too cold, but just right.

The continuity and activation hypotheses make opposite predictions. If speakers use repeats as a remedy for disruptions to discontinuity, the greater the disruption, the more often speakers should apply the remedy. That is, the longer the hiatus, the more often they should repeat the previous word. By the activation hypothesis, in contrast, the longer the hiatus, the *less* active the previous word should be and the *less* often speakers should repeat it.

Repeats themselves go against the continuity hypothesis because they leave behind an incomplete constituent (e.g., *I* in (1)). So the preference for continuity must be viewed alongside preliminary commitments, to which we turn next.

Preliminary Commitments

In stages I and II of repeats, speakers commit to a constituent and then immediately suspend their speech. As outsiders, we would describe these commitments as *premature:* They are made before they should have been if speakers are trying to achieve a continuous delivery. Logically, speakers could be in one of two states when they make these commitments. Either (a) at some level of processing they anticipated the suspension, or (b) they did not. Let us call the first type of commitment *preliminary.* Our hypothesis is that many suspensions are indeed preliminary.

The idea is this. Suspending speech in the middle of a constituent is a violation of continuity, and by the continuity hypothesis, this is something speakers should try to avoid. Yet speakers are also pressed by a *temporal imperative* (Clark, 1996): The time they take in speaking belongs to them and their addressees together, so they must justify to their addressees any extra time they take (Clark, 1996; Goffman, 1981). In (1), if Reynard pauses too long after the ''yes,'' he might be heard as opting out, distracted, or unsure about what he wants to say. He can forestall these attributions by using *I* to make a preliminary commitment to the next constituent.[3] By this logic, speakers are most likely to make preliminary commitments at the start of major constituents, where misattributions are most likely. Our hypothesis is this:

The commitment hypothesis. Some initial commitments to constituents are preliminary, with speakers already expecting, at some level of processing, to suspend speaking immediately afterward.

Premature commitments have long been observed in spontaneous speech. ''Since structural choices typically involve fewer alternatives than lexical choices,'' Maclay and Osgood (1959) argued, ''the speaker will often initiate a constituent before he has completed his lexical decisions—with the result that he may pause slightly in the middle of constituents before such lexical

[3] People may also produce early commitments to keep the floor (Maclay and Osgood's proposal), but that cannot be the whole story. Speakers appear to repeat words as often in monologues as in dialogues, though we have no evidence to offer here.

items'' (p. 42; see also Goldman-Eisler, 1968). The issue is whether any of these commitments are not just premature, but preliminary.

Repeating a word, in short, consists of several processes in succession—initial commitment, suspension of speech, dealing with a hiatus, and restarting a constituent. We have proposed three hypotheses covering one or more of these stages. We will take up each one in turn.

METHODS

The evidence we will consider comes from two large corpora: the Switchboard (or SW) corpus, a collection of spontaneous American telephone conversations; and the London-Lund (or LL) corpus, a collection of spontaneous British face-to-face conversations. We treated the SW corpus as our main source of evidence, because it is 16 times as large as the LL corpus. We identified and classified every repeated word from the SW corpus and every repeated article and pronoun from the LL corpus.

The SW corpus consists of about 2.7 million words from 2340 telephone conversations among about 500 employees of Texas Instruments. People signed up for this project, chose a topic from a list of possible topics, and were matched at random by a mechanical apparatus with other people, strangers, who chose the same topic. On each call, the pair was prompted by a question on the topic they signed up for, e.g., ''Do you believe that the U.S. government should provide universal health insurance, or should at least make it a long term goal? How far in that direction would you be willing to go? What do you see as the most important pros and cons of such a program?'' Each pair typically talked for up to ten minutes. We worked from the computerized transcripts, which coded sentences, words, repeats, *uh*, *um*, and other sounds, but not pauses. See Godfrey, Holliman, and McDaniel (1992) for details.

The LL corpus consists of about 170,000 words of transcripts from 50 face-to-face conversations (numbered S.1.1 through S.3.6) from the corpus of English conversation from Svartrik and Quirk (1980). The conversations were audiotaped between 1961 and 1976 among adult British men and women of various ages in two- to six-person settings. Although some of the speakers knew they were being recorded, most didn't, and we included only those who didn't. The transcripts coded not only words, but word fragments, pauses, tone units (constituents spoken under a single prosodic contour), overlapping speech, stress, and prosodic information such as rising, flat, and falling intonation. They also included the fillers *uh* and *um*, both in short and elongated versions. We worked from the computerized version of the transcript. See Svartvik and Quirk (1980) for details.

In the analyses that follow, we report chi-square statistics. Such statistics can be shown to be appropriate for generalizing over both speakers and items (Clark, 1973) because so few items are sparsely distributed over so many speakers. All the chi-squares we report are significant at the .001 level, except when noted otherwise.

COMPLEXITY OF CONSTITUENTS

It has long been recognized for English that function words are repeated far more often than content words (e.g., Fox & Jasperson, 1995; Holmes, 1988; Maclay & Osgood, 1959; Stenström, 1987). *Content words* refer to entities, events, states, relations, and properties in the world. They are characteristically nouns, verbs, adjectives, or adverbs, and the inventory of such words in the language is rather open-ended, with new ones frequently being

TABLE 1
Five Types of Function Words and the Constituents They Typically Initiate

Word class	Constituent type	Example
Articles	Noun phrases	*the* candidates who pass
Prepositions	Prepositional phrases	*of* our people
Subordinate conjunctions	Clauses	*if* they wanted a man to do medieval literature
Auxiliary verbs	Verb groups	*don't* want
Relative pronouns	Clauses	*who* pass

added. *Function words,* in contrast, are used largely to express the relations among elements of sentences, or to indicate their discourse functions. They include prepositions, conjunctions, articles, auxiliary verbs, and pronouns. The inventory of function words is relatively small and stable. In (1) and (3), the repeated words were the function word *I.* In the commit-and-restore model, function words should be repeated more often than content words for two reasons: (a) they tend to come first in major constituents, and (b) they tend to be more accessible and easier to pronounce.

Function words tend to be the left-most words of major constituents—noun phrases, verb phrases, prepositional phrases, clauses. Table 1 gives a partial list. Content words, in contrast, tend to occur later in these constituents, as illustrated by the nouns *candidates, people, man,* and *literature;* the verbs *pass, want,* and *do;* and the adjective *medieval.* When speakers want to make an initial commitment to a constituent, the word they must use is most likely a function word.

By the complexity hypothesis, then, function words should be repeated more often than content words. In their classic study, Maclay and Osgood (1959) reported the absolute number of repeated content and function words in their corpus, but not the number of unrepeated words, so we cannot test the prediction on their data. Nor can we test it on any other data in the literature. We therefore analyzed all repeated words in the SW corpus. We constructed as complete a list of function words as we could (408 words) and classified every word in the corpus as either a function word or a content word. The repeat rates (per thousand) for seven types of words are presented in Table 2. Overall, function words were repeated more than ten times as often as content words, 25.2 vs 2.4 per thousand, $\chi^2(1) = 17{,}176$.

One reason English function words are repeated so often may be that they are so frequent. In the SW corpus, 37.6% of the function word-types occurred over 1000 times, and only 0.5% of the content word-types did. Conversely, only 5.7% of the function word-types occurred fewer than 10 times, whereas 75% of the content word-types did. The high repeat rate for function words is consistent with the accessibility hypothesis.

Still, these findings cannot be accounted for by frequency alone. Consider

TABLE 2
Repeat Rates for Seven Types of Words

Type of word	Repeats per thousand
Function words	25.2
Auxiliary verbs	8.7
Conjunctions	30.8
Determiners	28.8
Prepositions	14.3
Pronouns	37.7
Miscellaneous	22.3
Content words	2.4

the words *I'm, at, had, because, out,* and *them.* All had frequencies of 10,000 ($\pm$400) in the SW corpus, yet their repeat rates (per thousand) varied by two orders of magnitude: *I'm* 56.1, *at* 16.3, *had* 8.9, *because* 3.1, *out* 1.2, and *them* 0.5, $\chi^2(5) = 1624$. Why do these words vary so much? Intuitively, it is because of their location in larger constituents. *I'm* is common at the left edge of clauses, whereas *them* is likely to occur later on. *At* is common at the left edge of prepositional phrases, whereas *out* is more common at the right edge of verb-particle constructions. Much the same reasoning applies to the other variation in Table 1. But comparisons like these are crude. The cleanest tests of our hypotheses are on the variation of repeat rates *within* categories. In this report we focus on pronouns and articles.

Simple vs Complex NPs

When speakers commit themselves to an NP but don't complete it, they often repeat the articles *a, an,* and *the.* NPs, however, range in complexity (Wasow, 1997). *The mangy dog,* for example, is slightly more complex than *the dog* because of the added modifier, but *the dog down the street* and *the dog my neighbor owns* are much more complex because of the prepositional and clausal modifiers *after* the head noun. To simplify complexity, we divided NPs into *simple NPs,* which don't have anything after the head noun, and *complex NPs,* which do (Erdman, 1988). By the complexity hypothesis, speakers ought to repeat *the* more often in complex NPs than in simple NPs.

We tested this prediction on the SW corpus. We sampled 500 NPs with repeated *the*'s and 500 NPs without and classified each NP as simple or complex (or unclassifiable). We also counted the singleton and repeated *the*'s in the corpus. With a little algebra, these counts allow us to estimate the repeat rates of simple and complex NPs in the entire corpus. As predicted, there was a large difference between them. *The* was repeated 40.9 times per thousand in complex NPs, but only 27.3 times per thousand in simple NPs, $\chi^2(1) = 7.78$, $p < .005$.

TABLE 3
Roles of NPs in Utterances

Role of noun phrase	Example
Topic	*The van that we've got* the gentleman who owned it had died
Subject	*The dog I have* is a German shepherd
Object of verb	I managed to find *the stereo I wanted* in Austin
Object of preposition	My wife parked her car in *the garage across from our house*

The same prediction applies to *a.* As with *the,* we sampled 500 NPs with repeated *a*'s and 500 without, classified the NPs, and estimated overall repeat rates. *A,* however, often occurs in expressions like *a lot of, a kind of,* and *sort of a.* Although *a lot of people,* for example, may look like a complex NP, it isn't that at all. *A lot of* is a fixed expression that means ''much'' or ''many,'' making the subject of *A lot of people are here* plural and not singular. So we classified the NPs with *a* into three categories: simple NPs, complex NPs, and fixed expressions (and unclassifiable). Again, as predicted, the difference between simple and complex NPs was large. *A* was repeated 54.9 times per thousand in complex NPs, but only 29.4 times in simple NPs, $\chi^2(1) = 17.58$.

What about fixed expressions? The general assumption is that they are formulated and understood as units, or at least not entirely as compositions of their parts (e.g., Nunberg, Sag, & Wasow, 1994). For example, if speakers formulate *a lot of* at one time, and if they want to commit to an NP that begins with *a lot of,* they have three choices: *a, a lot,* and *a lot of.* In fact, we find all three types of repeats (that is, these expressions followed by *a lot of*). They occurred 6, 3, and 6 times per thousand. Overall, then, *a* was repeated much less often in fixed expressions than in simple NPs, 6.3 to 29.4 per thousand, $\chi^2 (1) = 47$.

Articles in Larger Constituents

Constituents are hierarchical. These same NPs may themselves be at the left edge of larger constituents, which also vary in complexity. Consider the four roles of the complex NPs in Table 3. When the NP is in topic position, the article *the* is at the left edge of a very complex constituent. In contrast, when the NP is the object of a preposition, *the* is at the left edge of the NP alone. By the complexity hypothesis, *the* ought to be repeated most often in topic position, next most often in subject position, next most often in object position, and least often in objects of prepositions.

We tested these predictions on the SW corpus. We classified the 1000 NPs with and without repeated *the*'s into the four categories illustrated in Table 3 (plus the category unclassifiable). We counted an NP as a topic whenever it was a left-dislocated NP in a clause; in the example here, *the van that we've got* is coreferential with *it,* which occurs later in the clause.

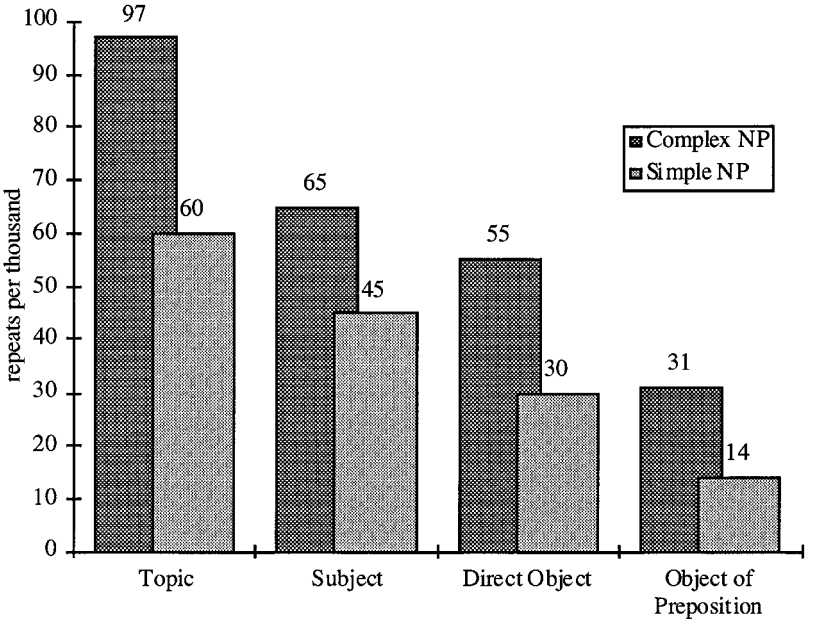

FIG. 1. Repeats of *the* in simple and complex noun phrases in four syntactic roles (SW corpus).

Again we classified each NP as simple or complex. The repeat rates are shown in Fig. 1, and they bear out the predictions. At one extreme, *the* was repeated 97 times per thousand in topics that were complex NPs. At the other extreme, it was repeated only 14 times per thousand in objects of prepositions that were simple NPs. The differences among the four roles are highly significant, $\chi^2(3) = 44$. There were more repeats in complex NPs than in simple NPs in each role. The two factors are roughly independent. In particular, the difference between complex and simple NPs was just as large for NPs in the middle of a clause (in direct objects or objects of prepositions) as for those at the beginning (in topics and subjects).

We repeated the analysis for the article *a*. There we had to add two categories, predicate nominatives as in ''he's a banker,'' and adverbs as in ''he's gone a lot of the time.'' There were so few indefinite NPs in topic and subject position (34 total in our sample) that we had to exclude them, and the indefinite NPs in adverbs were all fixed expressions, which we also excluded. The repeat rates for direct objects, predicate nominatives, and objects of prepositions are plotted in Fig. 2.

By the complexity hypothesis, *a* should be repeated more often in direct objects and predicate nominatives than in objects of prepositions, and it was, though not significantly so, $\chi^2(1) = 1.6$, n.s. As with *the*, *a* was repeated

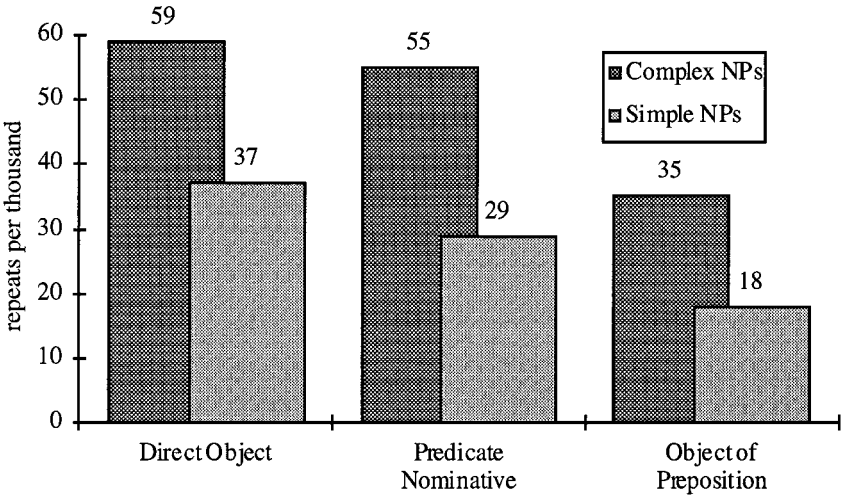

FIG. 2. Repeat rates of *a* in simple and complex noun phrases in three syntactic roles (SW corpus).

more often in complex than in simple NPs for each position separately. The repeat rates for *a* and *the* in the same position were strikingly similar. For direct objects, the rates were 37 and 30 per thousand in simple NPs, and 59 and 50 per thousand in complex NPs. For objects of prepositions, the rates were 18 and 14, and 35 and 31. So *a* and *the* are very much alike in their repeat rates, which are largely determined by the complexity and syntactic location of the NPs they initiate.

Personal Pronouns

Personal pronouns such as *she, our,* and *me* fall into five types, for a total of 34 pronouns:

> *Nominative:* I, you, he, she, it, we, they
> *Dependent possessive:* my, your, his, her, its, our, their
> *Independent possessive:* mine, yours, his, hers, [none for it], ours, theirs
> *Accusative:* me, you, him, her, it, us, them
> *Reflexive:* myself, yourself(ves), himself, herself, itself, ourselves, themselves.

The five types vary in the kind of constituent they can be the left-most word in. Nominative pronouns (e.g., *I*) are typically the left-most words in clauses such as ''*I* just got rid of, uh, a diesel engine Escort.'' Dependent possessives are the left-most words in NPs such as ''*my* wardrobe,'' but these, in turn, may occur anywhere within a clause. In contrast, almost all independent possessive, accusative, and reflexive pronouns (e.g., *mine, me,* and *myself*)

TABLE 4
Typical Constituents Initiated by Five Types of Pronouns

Type of pronoun	Type of constituent	Example of constituent
Nominative	Clause	*I* just got rid of, uh, a diesel engine Escort
Dependent possessive	Noun phrase	*my* wardrobe
Independent possessive	Pronoun	*mine*
Accusative	Pronoun	*me*
Reflexive	Pronoun	*myself*

are the left-most words in constituents of only one word, as in ''But, uh, she is a great comfort to *me.*'' Table 4 lists typical constituents whose left-most word is one of the five forms of *I:* By the complexity hypothesis, a pronoun should be repeated more often the more complex the constituent it is the left-most word of. Nominative pronouns should be repeated most often, dependent possessives next most, and the other three types the least.

To test the prediction, we calculated the individual repeat rates for all 30 pronoun forms in the SW corpus. The rates are shown in Table 5; the totals are the rates summed over the pronouns in that row or column. Note that the word forms *you, his, her,* and *it* are each ambiguous between two types; *you* and *it,* for example, are both nominative and accusative. We have listed these forms as ''ambiguous.''

These data provide solid evidence for the complexity hypothesis. As predicted, nominative pronouns were repeated most (54 times per thousand), whereas independent possessive, accusative, and reflexive pronouns were almost never repeated (0 to 6 times per thousand). The difference is highly significant, $\chi^2(1) = 1235$. By hypothesis, pronouns like *my* should be intermediate. As predicted, they were repeated less often than nominative pronouns, 31 to 54 per thousand, $\chi^2(1) = 258$, but more often than accusative pronouns 31 to 0.4 per thousand, $\chi^2(1) = 613$.

TABLE 5
Repeats per Thousand for Pronouns by Case in the SW Corpus (Total Number of Pronouns in Parentheses)

Type of pronoun	I	You	He	She	It	We	They	Total
Nominative (173,348)	63.9	—	39.9	24.3	—	35.7	41.7	54.0
Dependent possessive (28,062)	31.4	26.7	—	—	31.2	38.8	28.8	31.3
Independent possessive (1140)	3.3	0.0	—	0.0	—	17.0	0.0	6.1
Accusative (19,927)	0.2	—	0.4	—	—	0.5	0.5	0.4
Reflexive (2167)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Ambiguous (129,176)	—	11.0	22.5	14.1	25.0	—	—	11.4
Total	56.9	12.0	29.7	20.4	24.9	33.5	31.0	36.8

Possessive pronouns offer a crucial test of the complexity hypothesis. The two types of possessives are similar in meaning, but contrast in where they occur: *my* is always the left-most word of an NP, but *mine* is a full NP by itself, usually in non-subject position. As predicted, dependent possessives were repeated more often than independent possessives, 31 to 6 per thousand, $\chi^2(1) = 24$. But the test has an added twist. Although most independent possessives were not subjects (e.g., ''A friend of *mine* had a, uh, had a Buick''), some were (''*Mine* is sitting in the driveway''), and by the complexity hypothesis, these are more likely to be repeated. Indeed, although independent possessives were repeated only seven times in our corpus, all seven repeats were in subject position, as in, ''Well, mine, mine have always slept outdoors'' and, ''Ours, ours is the smallest one on the road.''

The ambiguous pronouns *her, you,* and *it* allow still another test. When speakers repeat these words, it should be because of their syntactic locations, not their phonological shapes. *Her,* for example, should be repeated as a possessive (*her wardrobe*), but not as an accusative pronoun (*And I feed her indoors*). Instead of classifying every *her, you,* and *it* in the SW corpus (over 125,000 tokens), we again used random samples. For example, we sampled 500 non-repeated instances of *her* and classified each as either possessive or accusative. The split was 49 to 51%. Next, we coded all 46 repeats of *her* as either possessive or accusative. The split was 100 to 0%. We then used these numbers, along with the total number of singleton and repeated *her*s in the corpus, to estimate the repeat rates for the two forms in the entire corpus.

As predicted, possessive *her* was repeated far more often than accusative *her,* 28.2 vs 0 per thousand, $\chi^2(1) = 36$. These rates are in line with the rates for other dependent possessives and accusatives. Likewise, the repeat rates for nominative and accusative *it* were 34.6 and 1.7 per thousand, $\chi^2(1) = 144$, and for nominative and accusative *you,* 9.5 and 0.7, $\chi^2(1) = 23$. These also fall in line with the rest of the pronouns. Plainly, the repeat of a pronoun is determined mostly by its location, not its phonological form or weight.

Even nominative pronouns aren't always in a position to be repeated. They should get repeated when they are at the left edge of a major constituent (as in, ''Well, uh, I, I can tell by your accent that you're a Texan,''), but not when they are not (as in, ''Oh, wha-, what else can I tell you about painting?''). We tested this notion on the SW corpus by identifying (a) every nominative pronoun (*I, he, she, we,* and *they*) that came just before the verbs *can, can't, could, couldn't, was,* or *wasn't* (chosen because they do not contract); and (b) every nominative pronoun that came just *after* these verbs as their subjects. The pronouns just before the verbs (about 15,000 of them) were repeated 31 times per thousand, but those just after (about 500) were repeated 0 times. The difference is reliable, $\chi^2(1) = 9.38$, $p < .002$. So when nominative

TABLE 6
Repeats (per Thousand) of Nominative Pronouns That Were Subjects of Positive and Negative Verbs

Verbs	Positive	Negative	Ratio 3 to 2
Can/can't	31	54	1.7
Could/couldn't	22	58	2.6
Was/wasn't	26	49	1.9
Total	27	54	2.0

pronouns are repeated, it isn't because they are nominative or the subject of a clause. It is because they are located at the left edge of major constituents.

The clauses initiated by nominative pronouns can also be positive or negative, as in ''I was . . .'' vs ''I wasn't. . . .'' Negative clauses are more complex, or heavier, than their positive counterparts both theoretically and empirically (Clark & Clark, 1977; Horn, 1989), so they should lead to more repeated pronouns. They did, as shown in Table 6. Nominative pronouns (*I, he, she, we, they*) were repeated 54 times per thousand before negative *can't, couldn't,* and *wasn't,* but only 27 times per thousand before positive *can, could,* and *was,* a ratio of 2 to 1. The difference is highly reliable, $\chi^2(1) = 48$.

Contractions

Nominative pronouns are often produced as parts of contractions such as *I'm, you've,* and *he'd.* Contractions of this type occur in mostly the same positions as nominative pronouns (e.g., *I'd like to see the mayor*), so by the complexity hypothesis, they should be repeated about as often as nominative pronouns, and they were. Table 7 lists the repeat rates for 17 contractions, categorized by type of contraction and by pronoun. The row labeled ''am/are'' contains *I'm, you're, we're,* and *they're;* ''have'' contains *I've, you've,*

TABLE 7
Repeats per Thousand for Pronoun Contractions by Type of Contraction (Total Number of Contractions in Parentheses)

Type of contraction	I	You	He	She	It	We	They	Total
Am, are (24,001)	56.1	31.8	—	—	—	53.1	60.1	52.8
Have (10,319)	55.5	19.4	—	—	—	33.1	35.2	44.9
Will (3753)	29.7	54.5	63.7	26.6	55.8	33.2	39.2	37.2
Had/would (2995)	37.2	17.9	6.2	10.1	0.0	20.9	15.6	26.0
Is/has (25,795)	—	—	45.4	31.4	81.8	—	—	76.2
Total (44,299)	52.3	29.8	44.5	29.5	81.2	43.0	52.8	58.5

etc.; ''will'' contains *I'll, you'll,* etc.; ''had/would'' contains *I'd, you'd,* etc.; and ''is/has'' contains *he's, she's,* and *it's.* Contractions were repeated 59 times per thousand, which is roughly the same as nominative pronouns, which were repeated 54 per thousand, but only roughly. By the complexity hypothesis, contractions should be repeated slightly more often than nominative pronouns. As just noted, some nominative pronouns occurred after such verbs as *can, can't, could, couldn't, was,* and *wasn't,* where they shouldn't be repeated, and weren't. Contractions, in contrast, can never occur in these locations, but only in locations where they should often be repeated. Indeed, the repeat rate for contractions (59 per thousand) was reliably greater than the rate for nominative pronouns (54 per thousand), $\chi^2(1) = 19$.

These findings cannot be accounted for by raw frequency alone. Nominative pronouns alone occurred far more often than contractions (4.5 times as often), yet they were repeated slightly less often. To take a dramatic example, *they're* occurred two-thirds as often as *them,* 6394 to 9763 times, but was repeated 353 times as often, 60 to 0.17 times per thousand. It is crucial where in the constituent the word is located.

Still, by the accessibility hypothesis, the more frequent the word, the more often it should be repeated—all other things being equal. To test the hypothesis, we compared words that occupied the same location in constituents, but differed only in frequency. Among the contractions, *is/has* was repeated most often, *am* or *are* next most, *have* next, *will* next, and *had/would* least. These rates are correlated .85 with the overall frequency of occurrence of these five types, $F(1, 4) = 12.81$, $p < .02$. Likewise, contractions fell in an order of frequency from *it* and *I* at the top end to *you* and *she* at the bottom. This order correlates .90 with the overall frequency of occurrence of these contractions, $F(1, 3) = 12.82$, $p < .05$. With *it* and *you* excluded, the more often a nominative pronoun was repeated, the more often its contractions were repeated as well ($r = .81$, n.s.). So within a structural category, the more frequent the word, the more likely it is to be repeated.

The complexity hypothesis, in brief, has excellent support. The articles *the* and *a* are repeated more often the more complex the local NP, and also the more complex the constituent of which the NP is the left-most constituent. These findings go counter to models in which syntactic complexity is categorical and not hierarchical (Ford, 1982; Holmes, 1988). Likewise, pronouns are repeated more often the more complex the constituent they are the left-most words of. Nominative pronouns are the most apt to be repeated, dependent possessives the next most, and accusative, reflexive, and independent possessives the least.

CONTINUITY OF DELIVERY

In the commit-and-restore model, repeating a word is really about restoring continuity to a constituent. According to the continuity hypothesis,

TABLE 8
Rates (per Thousand) of Fillers (Uh or Um) before and after Singleton Words of Four Types in the SW Corpus

Word type	Filler before	Filler after	Ratio 2 to 3
The	24.3	19.1	1.3
A/an	16.0	8.6	1.9
Nominative pronouns	53.7	8.1	6.6
Dependent possessives	34.7	17.1	2.0

speakers prefer the continuous delivery of a constituent, all other things being equal. The hypothesis leads to several testable predictions. (a) When speakers have trouble planning a constituent, they should prefer to delay *before* rather than *after* initiating it. (b) Yet once they have disrupted the constituent, the more serious the disruption, the more likely they should be to restart it, repeating the first word. (c) Once they have restarted a constituent, they should try to produce it with continuity. We take up the three predictions in order.

Starting up Constituents

By the first prediction, speakers would rather introduce a delay before than after initiating a constituent. Compare (9) and (10):

(9) and {uh} I got on to Hart again about it, (1.1.880)
(10) I {u:m -.} thought, this is not sensible, (2.4b.1179).

Clauses take time to plan, so speakers should sometimes pause near the beginning. If they are trying to produce the entire clause fluently, they should pause more often just before the word *I,* as Sam does in (9), than just after it, as Alex does in (10). The delay in (10) disrupts the continuity of the clause. The delay in (9) does not.

We tested the prediction in the SW corpus. Consider the article *the.* We identified each singleton *the* in the corpus (all 82,000 of them), noted whether it had a filler (*uh* or *um*) before or after it, and then computed the rates of fillers before and after it. Table 8 lists the rates of fillers before and after *the, a,* nominative pronouns, and dependent possessives. In each case, there were more fillers before than after the singleton words, $\chi^2(1) = 54, 166, 5848, 165$. Here, then, is strong evidence for the continuity hypothesis.

The rates of fillers in Table 8 are also evidence for the complexity hypothesis. As noted earlier, nominative pronouns tend to be at the left edge of clauses, whereas dependent possessives and articles are at the left edge of NPs. Clauses tend to be more complex than NPs because most NPs are constituents of clauses. If so, speakers should delay more often before clauses than before NPs, and they did. The ratio of fillers before to after nominative pronouns was 6.6, whereas the other three ratios ranged from 1.3 to 2.0.

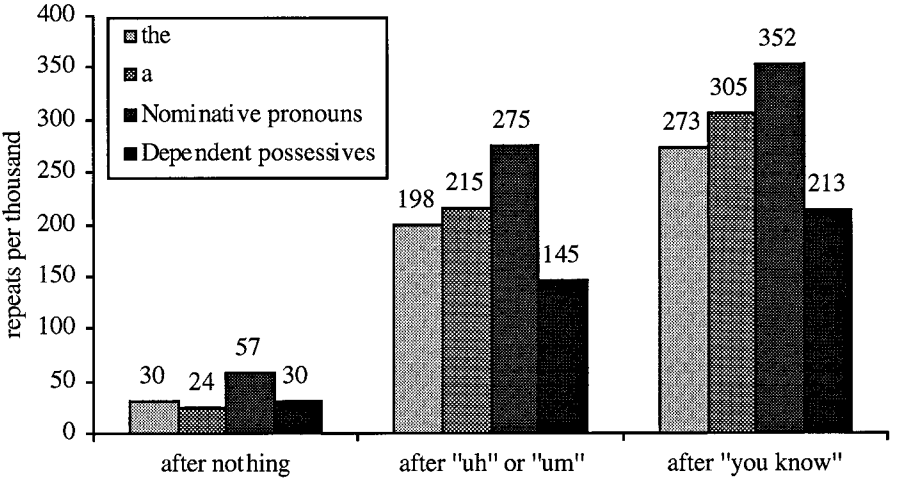

FIG. 3. Repeat rates after three types of interruptions (SW corpus).

Disruptions in Delivery

As we have noted, speakers often suspend their speech immediately after initiating a constituent. The disruption they create, however, may range from barely noticeable to very serious. Compare the disruptions in these four hiatuses between repeated *I*'s:

Case 1. well I {} I got rather fed up of some of these youngsters, (1.1.768)

Case 2. and I {-} I didn't want him to tell them, (1.9.1217)

Case 3. yes, I {uh} I wouldn't be surprised at that, (1.1.278)

Case 4. but I {. you know,} I recognize, it was equally difficult for her, (2.7.508).

Case 1 has no pause; case 2 has a pause, but nothing else; case 3 has a filler; and case 4 has the editing phrase *you know*. The disruptions range from minimal in case 1 to major in case 4. By the continuity hypothesis, speakers should be more likely to restart a constituent as the disruptions increase from case 1 to case 4.

We compared cases 1, 3, and 4 in the SW corpus. (Case 2 is swallowed up in case 1, because the corpus doesn't mark pauses.) Overall, *the* was repeated 31 times per thousand, which is an average of cases 1 and 3. So we divided the data into cases 1 and 3 and added case 4. The repeat rates for these three cases were 30, 198, and 273 per thousand. Figure 3 plots the three repeat rates for *the*, *a*, nominative pronouns, and dependent possessives. All show an increase from case 1 to case 3 to case 4. The differences from case 1 to case 3 are each significant, $\chi^2(1) > 88$, and so are those from case 3 to case 4, except for the possessives, $\chi^2(1) = 10$, 7.0 ($p < .01$), 16, and 2.37 (n.s).

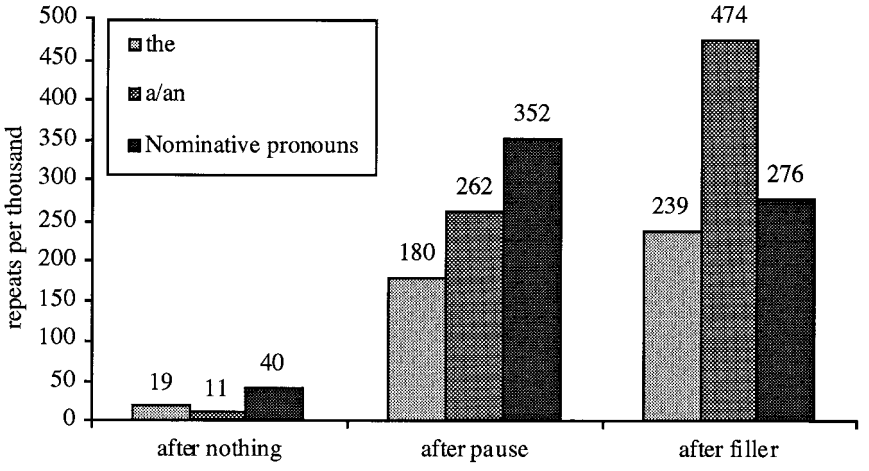

FIG. 4. Repeat rates after three types of disruptions (LL corpus).

By the continuity hypothesis, disruptions should lead to repeats only when they arise immediately *after* the to-be-repeated item. In the SW corpus, for example, speakers went on to repeat *the* 30 times per thousand after ''the {}'' (case 1), but 198 times per thousand after ''the {uh/um}'' (case 3). A devil's advocate might argue that the reason speakers repeated *the* so much after ''the {uh/um}'' was because they were already in a disfluent phase of their speech. If they were, they should repeat *the* just as often after ''{uh/um} the,'' but they didn't. The repeat rate for ''{uh/um} the,'' was 64 per thousand, which is about a third of the rate for ''the {uh/um}'' (case 3), which was 198 per thousand. The corresponding rates for *a/an,* nominative pronouns, and dependent possessives were 17 to 215, 67 to 275, and 53 to 142 (per thousand). All four differences are significant, $\chi^2(1) = 143, 184, 593, 21$. The devil's advocate cannot be correct. For a disruption to lead to a repeat, it must come just *after* the to-be-repeated item, just as the continuity hypothesis predicts.

We also compared cases 1, 2, and 3 in the LL corpus. (There weren't enough instances of ''you know'' to include case 4.) *The,* for example, was repeated 19 times per thousand for case 1, 180 times for case 2, and 239 times for case 3 (a filler with or without other pauses). Figure 4 lists the analogous statistics for *a/an,* and nominative pronouns. Cases 2 and 3 led to more repeats than case 1 for all three types of words, $\chi^2(1) > 63$. Case 3, in turn, led to more repeats than case 2 for *the* and *a/an,* $\chi^2(1) = 1.24$ (n.s.), 3.79 ($p < .05$), but to fewer repeats for nominative pronouns, $\chi^2(1) = 1.54$ (n.s.). So there were more repeats in case 2 than in case 1 for all three types of items, and more in case 3 than in case 2 in the one reliable difference.

TABLE 9
Rates (per Thousand) of Fillers (Uh or Um) before, between, and after Repeated Articles Pronouns (W1, W2) (SW Corpus)

Pattern	uh W1 W2	W1 uh W2	W1 W2 uh	Ratio 2 to 3
The	66.4	153.7	54.9	2.8
A/an	18.5	90.7	20.7	4.4
Nominative pronouns	74.2	54.8	16.1	3.4
Dependent possessives	66	79.7	49.3	1.6

When speakers pronounce *the* as ''thiy'' (rhyming with *see*) as opposed to ''thuh'' (rhyming with the second syllable of *sofa*), they normally signal that they are likely to suspend their speech at the end of the word (Fox Tree & Clark, 1997). In this way, *thiy* signals a disruption, so it should be repeated (as *thiy* or *thuh*) more often than *thuh.* The contrast between *thiy* and *thuh* is coded in the LL corpus, and indeed, *thiy* is repeated 155 times per thousand, and *thuh,* only 15, $\chi^2(1) = 342$. That is, *thiy,* which disrupts continuity, is repeated more than ten times as often as *thuh,* which does not.

So speakers are more likely to restart a constituent the more severe the disruption—from a null hiatus, to a non-reduced vowel (as in *thiy*), to a pause, to a filler, to the editing expression *you know.* These increases aren't trivial. The articles and pronouns we are looking at were repeated in the LL corpus from 1 to 4% of the time when there was no serious disruption. With just a pause, those percentages shot up to 18 to 35%, and with a filler, they went up further to 24 to 47%. Repeats aren't a consequence of general planning difficulty. They increase only when the disruption comes immediately *after* the to-be-repeated word. All this is evidence for the continuity hypothesis.

Continuity after Restarting

Once speakers have suspended their speech, they should prefer to restart the constituent and deliver it as a continuous whole. If W1 and W2 are the two tokens of a repeat, speakers should try to present W2 as a fluent piece of the constituent it is part of, so they should be more likely to delay before W2 than after W2. To test this prediction, we counted the fillers (*uh* and *um*) around repeated nominative pronouns, dependent possessives, *the,* and *a/an* in the SW corpus. For example, we tallied how often *he* plus *he* had an *uh* or *um* just preceding it (''uh he he''), in the middle (''he uh he''), and just following it (''he he uh''); ''uh he uh he'' was counted in both the first and second patterns, and ''he uh he uh'' in both the second and third patterns. Table 9 shows the result for *the, a/an,* nominative pronouns (*I, he, she, we,* and *they*), and dependent possessives (*my, your, its, our,* and *their*).

The fillers around repeated pronouns and articles follow the predictions of the continuity hypothesis. Take the *we*'s in, ''uh, we, uh, we had r-, our

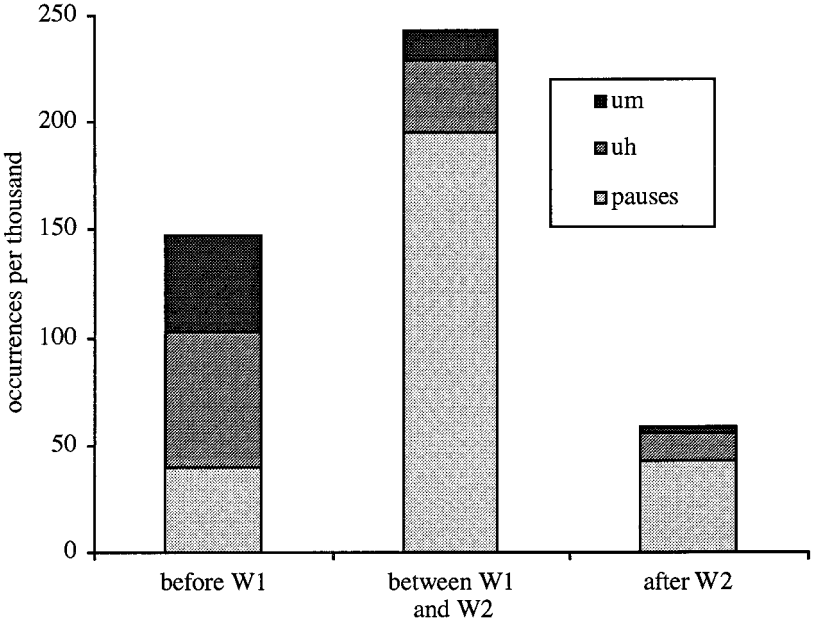

FIG. 5. Pauses and fillers before, between, and after repeated pronouns (W1 and W2) (LL corpus).

requirements change.'' Because entire clauses take time to plan, speakers should often pause before *we1* and between *we1* and *we2,* as in this example. But once they resume speaking, they should try to produce the entire clause fluently and should pause less often after *we2.* As predicted, there were more fillers in the first two positions than in the third position for all four types of words. For nominative pronouns, there were 74, 55, and 16 fillers per thousand in the three positions. In particular, the ratio of fillers before W2 to after W2 was 55 to 16, or about 3 to 1, $\chi^2(1) = 209$. The same arguments apply to possessives, *the,* and *a/an.* For them, the ratios were 80 to 49, 154 to 55, and 90 to 20, $\chi^2(1) = 6.84$ ($p < .01$), 146, and 146.

Pauses—silences—should also provide evidence for the continuity hypothesis. Speakers use *uh* and *um* mainly when they anticipate longish delays (Clark, 1994; Smith & Clark, 1993). When they use pauses alone, that should reflect shorter delays. To compare pauses and fillers, we turned to the LL corpus. The pauses and fillers around the 370 repeats of nominative pronouns are shown in Fig. 5. The portion labeled *um* includes ''um'' or ''u:m'' plus any accompanying pauses, and the portion labeled *uh* includes ''uh'' or ''u:h'' plus any accompanying pauses. The portion labeled *pauses* includes all combinations of brief and unit pauses without fillers.

The pattern of pauses in Fig. 5 supports the continuity hypothesis. Speak-

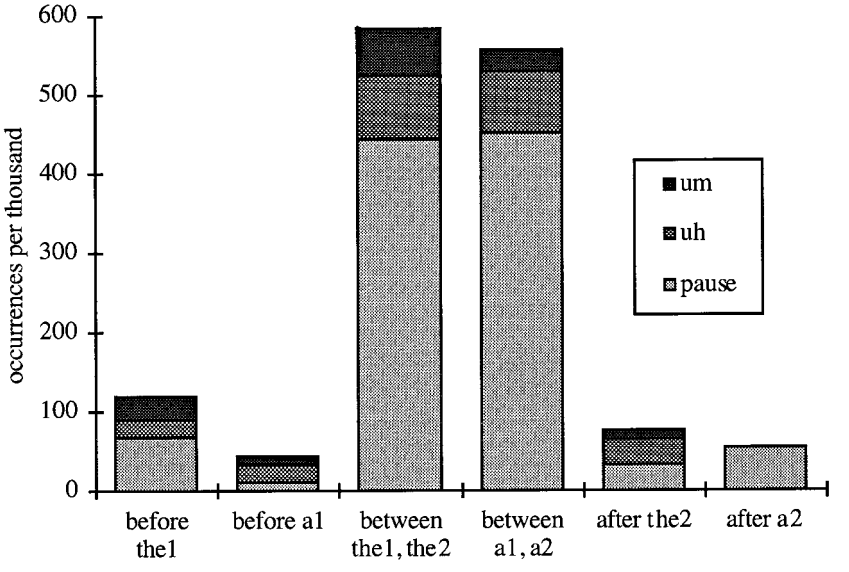

FIG. 6. Pauses and fillers before, between, and after repeated *the* and *a/an* (LL corpus).

ers paused 43, 218, 43 times per thousand before W1, between W1 and W2, and after W2. The ratio of pauses before W2 to after W2 was about 5 to 1, $\chi^2(1) = 48$. If we combine fillers and pauses for a measure of total delay, the rates become 148, 244, and 59. The ratio of delays before W2 to after W2 is 4 to 1, $\chi^2(1) = 56$. Once again, speakers are most likely to delay *between* W1 and W2.

The pauses and fillers around *the* and *a/an,* shown in Fig. 6, give much the same picture. (Although *a* and *uh* are both pronounced as schwas, they are distinguished by their fundamental frequency, F_0 (Shriberg & Lickley, 1993).) There were delays before W2 over 55% of the time, but after W2 less than 8% of the time. For *the,* there were 9 times as many delays before W2 as after W2, $\chi^2(1) = 93$, and for *a/an,* over 10 times as many, $\chi^2(1) = 41$. Furthermore, as we will note in the next section, speakers generally pronounce W1 in a special way, lengthened, but they pronounce W2 as they would pronounce it in any other intact constituent. All three of these findings are strong support for the continuity hypothesis.

Speakers, then, try to avoid disrupting the delivery of major constituents. First, they are more likely to delay before than after initiating a constituent—even without repeating any words. Second, they are more likely to repeat a word, restoring continuity to the constituent, as the disruption after that word becomes more serious—from a null hiatus to a pause to a filler to the editing expression *you know.* Finally, they are more likely to pause before than after restarting a constituent. These three points hold for all four word types we have looked at, *the, a/an,* nominative pronouns, and dependent possessives.

PRELIMINARY COMMITMENTS

In repeats, speakers make what outsiders would call premature commitments: they produce a word and then immediately suspend their speech. The commitment hypothesis is that many of these commitments are not merely premature, but preliminary: At the point at which speakers make the commitment, they are already expecting, at some level of processing, to suspend their speech. To test this hypothesis, we must look at the prosody of repeats.

Phonological Words

The prosodic structure of ideal utterances is taken by many to be hierarchical (e.g., Hayes, 1989; Nespor & Vogel, 1986; Selkirk, 1984, 1995). It divides into intonational phrases, phonological phrases, phonological (or prosodic) words, feet, and finally syllables. Most actual utterances, however, deviate from the ideal. We will focus on deviations in phonological words.

Phonological words are often formed from two or more words at the syntactic level (hereafter simply *words*). So in an ideal delivery, the five words in (11) might combine to form the two phonological words in (11′):

(11) I'm employed as an artist.
(11′) I.mem.ployed a.sa.nar.tist.

In our notation, phonological words are separated by spaces, and syllables by periods; otherwise, we use ordinary orthography (see Kenstowicz, 1994). The phonological words formed this way are often subject to prosodic adjustments. One such adjustment is *resyllabification.* In (11), for example, the first two words, *I'm employed,* get resyllabified in (11′) to form the phonological word ''I.mem.ployed,'' and the next three words, *as an art-ist,* get resyllabified as ''a.sa.nar.tist.'' In contrast, *I'm happy for you* requires no resyllabification, and it comes out ''Im.hap.py for.you.'' Words get realized differently at the phonological level depending on what other words they are combined with.

Prosodic adjustment places certain constraints on planning (e.g., Ferreira, 1991; Wheeldon & Lahiri, 1997). Speakers cannot formulate the first phonological word of (11′) (''I.mem.ployed'') until they have formulated the first *two* words of (11) (''I'm employed''). Prosodic adjustment is especially important for articles and pronouns. In ideal utterances, according to Selkirk (1995), function words of one syllable always get attached to, or *cliticized* onto, the following content words at the phonological level—unless they are spoken in isolation, carry stress, or come at the end of a construction. As (11) and (11′) show, such attachment often requires major prosodic adjustments.

If speakers prefer a continuous delivery, they should prefer to combine words into the longer standard phonological words, as in (11′). But sometimes they do not produce the expected phonological words. Consider Ken's ''I'm employed'' in (12):

(12) well, {.} I'm {.} employed as a mathematician, (1.6.14.B)
(12′) Im (pause) em.ployed

If Ken had delivered (12) in the expected way, he would have produced ''I.mem.ployed'' as a single phonological word. Instead, he produced the *two* phonological words in (12′), pronouncing ''Im'' as phonological word on its own. This is non-standard because it was *not* attached to the following word (*employed*) and it was *not* resyllabified as it would have been in the standard case. Ideally, it should *not* be pronounced as a separate syllable. We will call such a case a *phonological orphan.* Orphans are different both from *intact words,* the corresponding portions of standard phonological words (''I.m'' in ''I.mem.ployed'' in (11′)), and from *word fragments* such as William's ''th-'' in (13):

(13) and even if promises {.} are given, th- {} the actual workers, {you know, I mean,} (3.4.410.A).

We will argue that orphans, word fragments, and intact words have different origins.

Two Types of Interruption

How do speakers suspend their speech? At first, the answer seems obvious. They just stop. But stop what? Speakers can interrupt production at several points in the process. We will argue that orphans arise from interruptions at the syntactic level, and fragments, from interruptions at the phonological or articulatory level.

Speakers sometimes suspend their speech between words, as after ''I'm'' in (12), and sometimes within words, as after ''th-'' in (13). They don't take these options at random (Brédart, 1991; Levelt, 1983, 1989; Nooteboom, 1980). In self-repairs, they rarely interrupt a word when it is correct on its own, but they often do so when it isn't (3% vs 23% of the time in Brédart's corpus). As Levelt (1989, p. 481) put it, ''Words that are not errors themselves tend to be completed before interruption.'' Levelt proposed a strategy to account for this finding: ''By interrupting a word, a speaker signals to the addressee that that word is an error. If the word is complete, the speaker intends the listener to interpret it as correctly delivered.'' So whenever speakers complete a word, they are committing themselves to its correctness—at that moment at least. Let us call this the *completed word strategy.*

To apply such a strategy, speakers must *plan* their interruptions. In (12), Ken may have formulated *I'm* or *I'm employed* or even *I'm employed as a mathematician* at the syntactic level. Yet all he sent on to the phonological level at that moment was *I'm,* which therefore got formulated as the orphan ''Im.'' Producing an orphan, in this view, reflects an interruption at the syntactic level. In (13), by contrast, let us suppose that William had formulated *the new* at the syntactic level and had sent it on to the phonological level, where it became ''the.new.'' But in monitoring what he was about to articu-

late (see Wheeldon & Levelt, 1995), he discovered that ''the.new'' was incorrect, so he interrupted the articulation, leaving ''th-'' as a fragment. Fragments, in this view, reflect an interruption at the phonological or articulatory level.

On this account, orphans reflect preliminary *syntactic* commitments. In (12), when Ken sent on *I'm* by itself to the phonological level, he was guaranteed to produce the orphan ''Im'' plus an immediate suspension. At some level of processing, he *expected* to produce the orphan, and by the completed word strategy, he was committing himself to its correctness. He was, therefore, *marking* a preliminary commitment to the word *I'm* and to a clause beginning with *I'm.*

Fragments, in contrast, reflect *articulatory* commitments. In (13), when William hypothetically sent *the new* on to the phonological level, he was expecting to produce a standard phonological word (''the.new''), and at that point he was committed to its correctness. Then, based on monitoring, he interrupted his articulation to indicate that it was *not* correct. Some articulatory commitments may be preliminary, and others may not be.

Syntactic Commitments

By these arguments, when speakers interrupt the syntactic level after certain function words (e.g., after *I'm* in (12)), they are making preliminary commitments both to the words themselves (*I'm*) and to the constituents they initiate (a clause beginning with *I'm*). The result is a phonological orphan (''Im''). The pronunciations of orphans (''Im'') are longer and fuller than those of both fragments (''th-'') and intact words (''I.m''). If these arguments are correct, the first tokens of many repeats should have these pronunciations, and they do.

Thiy. Thiy is a nonstandard pronunciation of *the,* and in spontaneous speech, it is almost always an orphan. Speakers in the LL corpus suspended their speech immediately after *thiy* 81% of the time and after *thuh* only 7% of the time (Fox Tree & Clark, 1997). That is, when speakers formulate *the* as the non-standard *thiy,* they expect to suspend their speech immediately afterwards. They are using *thiy* to make a preliminary commitment to the constituents it initiates. As we noted earlier, *thiy* is repeated as *thiy* or *thuh* 155 times per thousand, whereas *thuh* is repeated only 15 times per thousand.

Thuh. Thuh can also be an orphan. Record yourself saying, ''I see the cat on the mat'' first fluently and then one word per second to the beat a metronome. Try as you might, you cannot edit the metronome version to get the fluent version. The metronome *thuh*'s are longer than the fluent *thuh*'s, and their prosody doesn't fit with *cat* and *mat.* First tokens of repeated *thuh*'s are much like the metronome *thuh*'s. According to measurements by Shriberg (1994) on the SW corpus, they averaged 2.5 times as long as the second tokens, which had the same average length as the unrepeated *the*'s. If these first tokens had been fragments, they would have been as short as or shorter

than unrepeated *the*'s. By these measurements, then, many first tokens of repeated *the*'s are orphans, marking preliminary commitments.

Contractions. In (12), when Ken produced ''I'm {.} employed,'' we argued that his ''Im'' was an orphan because *I'm employed* would normally resyllabify as ''I.mem.ployed.'' *I'm* should resyllabify whenever the following word begins with a vowel, and the same holds for every other contraction in Table 7. By this argument, the first token of every repeated contraction before a vowel must be an orphan and not a fragment.

To see how often these orphans arise, we examined *I'm* in the SW corpus. We sampled 545 singletons and 565 repeats and classified each by the initial sound of the following word (e.g., ''I'm, I'm planning on it'' vs ''I'm, I'm intrigued by it''). Repeats of *I'm* were followed by a vowel 21% of the time. The first tokens in these repeats are orphans. But *I'm* was repeated just as often before consonants as before vowels; singletons of *I'm* were followed by vowels 25% of the time, which is not significantly different from 21%. If so, it follows that the first *I'm* in most of the remaining repeats were probably orphans too. The same argument holds for the other repeated contractions.

An. An always occurs before vowels, and like *I'm,* it resyllabifies when combined with the following word. *An egg* is normally pronounced as ''a.negg.'' So whenever speakers interrupt the syntactic level after *an,* they produce an orphan. As a result, the first token of a repeated *an* is always an orphan, and speakers in the SW corpus repeated it 8.4 times per thousand. But if the first tokens of all these *an*'s are orphans, then the first tokens of many repeated *a*'s are surely orphans too.

Pronouns. Whenever the first tokens of repeats are orphans, they should have pronunciations that are distinct from the second tokens. We have already seen such a difference in *thiy* repeated as *thuh.* The LL corpus provides further evidence for such differences. It marks words for several prosodic features with respect to ''tone units'' (intonational phrases). (a) It marks the ''onset'' of each tone unit, the word with the first prominent syllable. (b) It marks the ''nucleus'' of each tone unit, the peak of greatest prominence. (c) It marks the intonation on that nucleus as rising, falling, level, or some combination. (d) It marks four levels of relative pitch on particular words. (e) It marks normal vs heavy stress. If the two tokens of a repeat differ in pronunciation in any of these ways, that will turn up in their markings. In ''I I get rather fed up,'' for example, the first *I* is marked for a ''higher pitch than the preceding pitch-prominent syllable,'' whereas the second *I* is unmarked. We examined all 506 repeats of *I, he, she, we,* and *they* in the LL corpus. The two tokens differed in their markings 37% of the time. The remaining 63% of the repeats may have differed as well, of course, but on other features.

In many repeats, therefore, the first tokens have the pronunciation of orphans and not of fragments or intact words. Their pronunciations tend to be

longer (as with *thuh*) and fuller (as with *thiy, I'm, an*), and their prosody tends to be different from second tokens. All this is evidence that many first tokens are phonological orphans being used for preliminary commitments to constituents.

Syntactic Commitments with Fillers

The first tokens of some repeats are followed by *uh* or *um*, as in this utterance by Duncan:

(14) we {um .} we learn something, {.} not just about Hamlet, but about {.} human nature in general, (3.5a.334).

For repeated nominative pronouns, first tokens were followed by fillers 55 times per thousand in the SW corpus and 48 times per thousand in the LL corpus (Table 9 and Figure 4). In the LL corpus, they were followed by the filler without a pause in 95% of the cases, as in (1) and (14). Most fillers in repeated pronouns, therefore, are what we will call *instant fillers.*

Fillers, like any word, take planning. In (14), Duncan was monitoring what he was about to produce, and when he anticipated a major delay after *we,* he formulated *um* to insert immediately after it. If he had anticipated a minor delay, he would have formulated *uh* instead (Clark, 1994; Smith & Clark, 1993). In (14), the time interval between the first token and the filler was effectively zero, which was too little time for Duncan to plan and produce *um* only after he had produced *we.* He must have planned the two items at the same time. That in turn implies that he planned *we* as an orphan, as a preliminary commitment to the clause it initiated. This logic holds for all instant fillers.

This analysis is supported by surprising evidence—the existence of phonological words like ''I.muh.'' We taped three hours of a local radio interview program in which a moderator interviewed experts and callers about several political topics. Except for the introductions, the speech was spontaneous, and it contained many instant fillers. Many of these fillers were *enclitic* on the previous word to form a single phonological word. The italicized parts of (15) and (16), for example, were produced as (15′) and (16′):

(15) but it might also lead *to uh to all sorts of* bogus representations
(15′) to.wuh tall.sort.sof
(16) th- there is a (0.2) *a uh* (0.5) *a potential* problem
(16′) ei.yuh (0.5) a.po.ten.tial.

In (15), the first token of *to* was pronounced ''tuw'' (with an unreduced vowel) and resyllabified with *uh* to form ''to.wuh''. The second *to* was cliticized onto *all* to produce the syllable ''tall.'' The two tokens were 280 and 50 ms long, much like the difference between the first and second tokens of *the.* In (16), the first token of *a* was pronounced ''ei'' (rhyming with *day*) and resyllabified with *uh* to form ''ei.yuh.'' The second *a* was pronounced as a schwa and cliticized onto *potential* to form ''a.po.ten.tial.'' The first

token had an unreduced vowel, just as *the* and *to* were pronounced as ''thiy'' and ''tuw,'' and it too was longer than the second token, 184 to 100 ms.

To produce ''to.wuh'' and ''ei.yuh'' as phonological words, the speakers must have formulated both the word (*to* or *a*) and the filler (*uh*) at syntactic level and ordered them before sending them on to the phonological level. If so, they must have expected at that point to produce ''to.wuh'' and ''ei.yuh'' as non-standard phonological words—as orphans. They couldn't have cliticized *uh* onto *to* or *a* without pronouncing them with the unreduced vowels in ''tuw'' and ''ei,'' which also marks them as orphans.[4] That makes ''tuw'' and ''ei'' preliminary commitments to the phrases they initiate.

Articulatory Commitments

If fragments come from interruptions at the phonological or articulatory level, they should be different from orphans and intact words. By the completed word strategy, speakers should interrupt a word only when it is an error. By definition, the first tokens in repeats are not errors—at least on the surface—so few of them should be fragments. As predicted, the first tokens of repeats were fragments only 12% of the time in the corpus of Maclay and Osgood (1959) and even less often in the SW corpus. In the SW corpus, fragments of *the,* for example, were coded as ''th-'' and ''the-''. Once we include these among the repeats, then first tokens were fragments only 3% of the time for *the,* 2% of the time for nominative pronouns, and 2% of the time for contractions. Let us call these *fragment repeats.*

Fragment repeats are unlike word repeats in several ways. By hypothesis, what speakers are interrupting is a phonological word, not a syntactic word. So, for example, the coders of the SW corpus heard the first element in ''I-, I'm'' in (17) as a fragment and, therefore, as different from the first element in ''I, I'm'' in (18):

(17) I-, I'm only twenty-five, so I've never actually been through a period of war—
(18) But, uh, I guess, I, I'm a little encouraged with home prices.

''I-'' in (17) could have been a fragment of ''Ive'' or ''Ill'' or ''I.yam'' or even ''I.'' This is further evidence that the first tokens of word repeats (as in (1)) have a source that is different from the first tokens of fragment repeats.

Also unlike word repeats, fragment repeats are almost always instantaneous. We compared ''the the'' vs ''th- the'' in the SW corpus. There were fillers 155 times per thousand within the word repeats, but only 9 times per thousand within the fragment repeats, $\chi^2(1) = 17.89$. The corresponding rates for the nominative pronouns, *I, he, she, we,* and *they* were 65 vs 10 times per thousand, $\chi^2(1) = 4.87$, $p < .05$, and for the contractions with *I, he, she, we,* and *they,* 51 vs 16 times per thousand, $\chi^2(1) = 1.49$, n.s. Most

[4] We know of no model of production able to account for the processes needed for ''tu.wuh'' and ''ei.yuh.''

fragment repeats, in short, arise from the instant replacement within one phonological word (''th-'') by another phonological word (''the.yactual'').

Instant replacements are also part of *instant repairs.* In an example recorded by Blackmer and Mitton (1991), a man calling into a radio talk show was recorded as saying, ''The Lord says that and eventually you'll have to re- {} answer to him'' (p. 188). In this utterance the caller began to produce ''respond,'' suspended his speech midword, and produced ''answer'' to replace it. The hiatus, as Blackmer and Mitton noted, was 0 ms, which is too little time for the caller to have decided on a replacement, formulated it, and begun articulating it. The caller must have detected the problem earlier and interrupted articulation only when he had the repair formulated and ready to articulate.

Fragment repeats are analogous. When William says, ''th- the actual workers,'' he produces ''th-,'' suspends his speech, and then produces the entire phrase ''the actual workers.'' In (16), where we do have measurements, there was only 85 ms between ''th-'' and ''there'' (175 ms from the beginning of ''th-'' to the beginning of ''there''). As with instant repairs, a hiatus of 85 ms is too little time for Sam to decide to repeat *the,* prepare it again, and then begin articulating it. He must have prepared the phonological word ''the.yactual'' before he produced ''th-.''

Why, then, do people initiate words like ''the'' or ''respond'' when they have already formulated replacements for them? One hypothesis is that they are making articulatory commitments: They are committing themselves to going on with their speech. When the talk-show caller continues ''you'll have to respond'' despite his plans to replace *respond* with *answer,* he creates the illusion of a continuous delivery.

Default Commitments

Speakers often produce what we will call *near repeats,* as when a speaker named Adam replaces ''a'' with ''an'':

(19) we're in a {.} an intolerable position, (3.4.949).

Near repeats fall into at least three possible categories. The first type we will call *substitutions.* Suppose Adam's original syntactic plan was to say *a terrible.* He sends both words on to the phonological level, where they are formulated as ''a.ter.ri.ble.'' But after committing himself, he changes his mind and decides to say ''an intolerable.'' He therefore interrupts ''a.ter.ri.-ble'' after its first syllable, formulates ''a.nin.tol.e.ra.ble,'' and articulates it. The result is the *fragment* ''a-'' followed by ''an intolerable position.''

The second type we will call *word repairs.* As before, Adam's original plan was to say *a terrible,* but this time he sends *a* by itself to the phonological level, where it is formulated and articulated as the orphan ''a.'' As before, Adam changes his mind and decides to say ''an intolerable.'' He therefore sends the two words on to the phonological level, formulates ''a.nin.tol.e.ra.-

TABLE 10
Occurrences of Four Sequences of Indefinite Articles

Sequence	Occurrences
1. a a	1800
2. an an	42
3. a an	126
4. an a	16

ble,'' and articulates it. The result is the *orphan* ''a'' followed by ''an intolerable position.''

The third type we will call *restarts.* Suppose that, because of time pressure, Adam feels he must start speaking even though he hasn't yet formulated the next lexeme. So he formulates a default *a* at the syntactic level and sends it on *by itself* to the phonological level, where it gets formulated and articulated as the orphan ''a.'' Then, once he has chosen *intolerable* as his next lexeme, he sends *an intolerable* on to the phonological level, formulates it as ''a.nin.tol.e.ra.ble,'' and articulates it. As with word repairs, the result is the *orphan* ''a'' followed by ''an intolerable position.''

The first token of Adam's near repeat, then, could reflect three distinct commitments. In a substitution, it would reflect an articulatory commitment, not necessarily preliminary. In a word repair, it would reflect a preliminary syntactic commitment to the word *a.* In a restart, it would reflect a preliminary syntactic commitment to an indefinite NP. Happily, we can distinguish these three cases in the SW corpus. We already know that substitutions are rare because fragment repeats are rare. The question is: How many repeats reflect preliminary commitments to particular words and how many to constituents? We will take up near repeats for *a* and *an,* and for pronouns and contractions.

A vs an. The article *a* was repeated more than three times as often as *an,* 29.6 vs 8.4 times per thousand, $\chi^2(1) = 76$. This is expected if *a* is used by default in making preliminary commitments. Let us see how.

The choice of *a* vs *an* depends on the word that follows it. In the SW corpus, the word after *a/an* began with a consonant 92.9% of the time and with a vowel 7.1% of the time. It would be surprising if *a* were repeated more often than *an* simply because it was followed by a consonant. That point is confirmed in the data. Table 10 lists the counts not only for ''a a'' and ''an an'' but also for ''a an'' and ''an a'' (with and without fillers between them). There were also 9 near repeats beginning with the fragments *a-* or *an-*, but they do not bear on this analysis. If we treat all four sequences as repeats, the rate of repeating *a/an* was the same before consonants (28 per thousand) as before vowels (29 per thousand). Yet speakers go from *a* to *an* (line 3) eight times as often as they go from *an* to *a* (line 4). Why?

Suppose speakers have a default rule for committing to an indefinite count NP, and it is to say ''a.'' That is, speakers proceed according to this model:

To make an initial commitment to an NP:

(1) If all you know is that the NP is indefinite and count, produce *a* as the default.

(2) If you know that the NP is indefinite and count, and if you know the next lexeme, produce *a* or *an* depending on the initial sound of that lexeme.

When speakers haven't yet accessed the first lexeme after *a/an,* they choose default option (1) and produce ''a.'' But when they have, they choose option (2), producing ''a'' before consonants and ''an'' before vowels.

The model enables us to estimate how often speakers take the default option when they repeat or nearly repeat *a* or *an.* In the model, the probability that speakers start with *a* is the sum of two probabilities. The first is the probability that speakers take default option (1), which automatically leads to *a.* Let us call this probability *p,* and it is the probability we want to estimate. The second is the probability that speakers take option (2), and because the next lexeme begins with a consonant, they also say *a.* If the probability that they take option (1) is *p,* then the probability that they take option (2) is $(1 - p)$. Given they have taken option (2), it is easy to compute how often they will say *a.* According to Table 10, 1816 of the 1984 near repeats ended in *a,* indicating that the next lexeme began with a consonant 1816/1984, or .915, of the time. So the probability of producing *a* as the first token, according to the model, is $p + .915\ (1 - p)$. Now the actual proportion of times speakers produced ''a'' as the first token, also found in Table 10, is 1926/1984, or .971. That gives the equation $p + .915\ (1 - p) = .971$. Solving for *p,* we get a value of .655.

According to this analysis, when speakers make premature commitments to NPs and repeat *a* or *an,* they produce *a* by default 66% of the time. They commit themselves to an NP that begins with *a* even though they haven't yet formulated the next lexeme. When they produce *a* in option (1), they are publicly committing themselves to an NP whose next word begins with a consonant. This, of course, is an excellent strategy when they don't know what the next word will be, because they will be correct 92% of the time. But if they had known that the next word began with a vowel, they would *not* have produced *a* by default (by option (1)). They would have produced *an* (by option (2)). The default *a* is, therefore, a genuinely preliminary commitment to an indefinite NP.

Contractions. Sequences like ''I I'm'' and ''I'm I'' are another type of near repeat. Table 11 shows the counts of exact repeats and near repeats for seven pronouns. (There were also 160 fragments of *I, he, she, we,* and *they* followed by the same pronoun or a contraction, but as before, these do not enter into this analysis.) As with *a* and *an,* pronoun + contraction (line 3)

TABLE 11
Frequencies of Four Types of Pronoun Sequences

Type of sequence	I	He	She	We	They	You	It	Total
1. Pronoun + pronoun	6823	344	123	825	1243	—	—	—
2. Contraction + contraction	1004	121	41	282	466	167	1833	3914
3. Pronoun + contraction	857	75	18	125	148	93	642	1958
4. Contraction + pronoun	272	35	9	45	63	53	187	664
Ratio of line 3 to line 4	3.15	2.14	2.00	2.78	2.35	1.75	3.43	2.95

was more frequent than contraction + pronoun (line 4) for all seven pronouns. The overall ratio was almost 3 to 1, $\chi^2(1) = 664$. Also as with *a* and *an*, near repeats were less frequent than exact repeats.

Suppose speakers have a default rule for committing to clauses that begin with a pronoun or contraction, and it is to produce the pronoun alone. That is, speakers proceed according to this model (analogous to the model just described for *a* and *an*).

To make an initial commitment to a clause beginning with a pronoun:

(1) If all you have formulated so far is the pronoun, produce it as the default.

(2) If you have already formulated the pronoun and an auxiliary (if there is one), produce the appropriate pronoun or contraction.

When speakers have formulated an initial pronoun but don't yet know whether there is a following auxiliary, they choose default option (1) and produce the pronoun alone. But when they have formulated an auxiliary, they choose option (2), producing whichever form is appropriate, the pronoun alone or the contraction. So the man who said, ''I, I'm more worried about the ozone,'' produced *I* before he had formulated the auxiliary '''m.'' But the man who said, ''I'm, I'm thinking about like Cinderella,'' produced *I'm* only after formulating the auxiliary as well.

As before, this model enables us to estimate how often speakers take the default option in clauses that begin with repeated pronouns or contractions. The probability of taking that option was .31, .20, .15, .20, and .14 for *I, he, she, we,* and *they.* The probability for the five of them together is .25. So when speakers make premature commitments to clauses and repeat a pronoun or contraction, they produce the first pronoun by default 25% of the time. They commit to a clause that begins with the bare pronoun even though they haven't yet formulated the next element. By the same logic as for default *a*, then, speakers here too make genuinely preliminary syntactic commitments.

GENERAL DISCUSSION

Repeating a word, we have argued, arises from strategies that result in the four stages in Table 12. One strategy is to make an *early* commitment

TABLE 12
Four Stages in Repeating a Word

Stage	Speaker S's action	Spoken example
I. Initial commitment	S commits to a constituent	I
II. Suspension of speech	S stops vocalizing	{
III. Hiatus	S deals with potential delay	uh
IV. Restart of constituent	S restarts the constituent, restoring continuity to it	} I wouldn't be surprised at that

at stage I. When people are under pressure to speak, they can opt to commit to a constituent by producing its first word or words (e.g., ''I'') as early as possible. But if they start *too* early, they may have to suspend speaking immediately afterwards (stage II). A common form of this strategy is to make a *preliminary* commitment, producing the first word or words already expecting to suspend speaking immediately. In either case, once people resume speaking (stage IV), they can opt to continue the constituent (''wouldn't be surprised at that'') or to restart it (''I wouldn't be surprised at that''). In restarting, they repeat its first word or words (''I uh I wouldn't be surprised at that''). People are more likely to take this option when the break at stages II and III was especially disruptive. By restarting, they restore continuity to the constituent.

The commit-and-restore model has strong empirical support. We first review the evidence for the complexity, continuity, and commitment hypotheses and then return to the strategies and their rationale.

Complexity

Speakers often suspend speaking (stage II) after an initial commitment to a constituent (stage I). The complexity hypothesis is that they are more likely to *need* to do that the more complex the constituent. Our data give broad support for this hypothesis.

Take noun phrases (NPs). Speakers were more likely to repeat an initial *the* or *a* of an NP when the NP was complex (e.g., ''the, the time we were there at the warehouse'') than when it was simple (''the, the diesel''). They were also more likely to repeat *the* or *a* when the NP was at the left edge of a larger constituent (a clause) than when it was at the left edge of a smaller one (the object of a verb or preposition). That is, repeats increased with both short- and long-range complexity—both with the local NP and with the larger constituent, if any, initiated by it.

Next take pronouns. Speakers were more likely to repeat pronouns at the left edge of clauses (''we, we planted three pecan trees out front'') than of NPs (''our, our victims''), and they were more likely to repeat these than pronouns that were not at the left edge of larger constituents (*ours*, *us*, *ourselves*). Speakers were also more likely to repeat nominative pronouns at the

left edge of negative clauses (''we, we can't clean it up'') than of positive clauses (''we, we can, uh, be recorded''). They were even more likely to repeat contractions (e.g., *we've*), which always occur at the left edge of a clause, than nominative pronouns (*we*), which sometimes do not. They never repeated nominative pronouns in inverted subject positions after auxiliaries, as in, ''What can we do to solve that?''

Repeats of a word increase, therefore, with the grammatical weight of the *hierarchy* of constituents initiated by that word. This immediately rules out theories in which complexity is categorical and not hierarchical (Ford, 1982; Holmes, 1988). Such theories cannot account for the increase in repeats with *both* short- and long-range complexity, or for the increase in negative clauses. It also rules out local phonological complexity (Ferreira, 1991; Wheeldon & Lahiri, 1997) as the sole source of repeats. Such a proposal cannot account for the increase in repeats with long-range complexity.

For speakers to suspend speaking immediately after initiating a complex constituent, they need to be able to estimate its complexity—its grammatical weight. How might they do that? It is logically impossible for speakers to count the phonetic segments, or words, or syntactic nodes of a constituent they have yet to formulate, so they must be estimating how much information at a *conceptual* level the constituent is to express. This proposal is consistent with the long-established observation that when speakers have the option, they prefer to place more complex constituents after less complex ones (Behaghel, 1909/1910; Hawkins, 1994; Wasow, 1997). They prefer, ''Nobody reported [to the police] [where the accident took place]'' over, ''Nobody reported [where the accident took place] [to the police].'' For speakers to produce constituents in order of increasing complexity, they must be able to estimate the relative complexity of constituents they haven't yet formulated. To do that, they must be able to estimate the amount of conceptual information to be expressed.

Continuity

Why do speakers take the trouble to repeat a word they have just produced—as in, ''I uh I wouldn't be surprised at that''? Our proposal is that they do that to restore continuity to their delivery of the constituent they were initiating (''I wouldn't be surprised at that'').

The continuity hypothesis is supported by three types of evidence. Speakers are much more likely to add a delay just *before* initiating a constituent than after its first word. They are also more likely to add a pause or filler just *before* the restart of the constituent than just after. The most dramatic evidence is that speakers are more likely to repeat a word the more disruptive the action in the following hiatus. For the words we examined, repeats occurred 1 to 6% of the time when there was no pause in the hiatus. They jumped to 14 to 35% when there was a pause and to 21 to 47% when there was a filler. Note that these are percentages, not occurrences per thousand.

Overall, nominative pronouns are repeated about 5% of the time, which is remarkable enough. But when they are followed by a filler, that number jumps to 28%, and when they are followed by ''you know,'' the number becomes 35%—one out of every three times. Repeats increase sharply as the disruption gets larger.

These findings go directly against the activation hypothesis—that speakers repeat a word because it is the most highly activated word at the moment a speaker resumes speaking. By that hypothesis, activation should decay over long hiatuses, so the longer the hiatus, the less often speakers should repeat. The data show precisely the opposite.

The activation hypothesis, and many related hypotheses, are also at odds with the facts of spontaneous Japanese. English is pre-positional, with function words coming mainly at the left edge of local constituents, but Japanese is post-positional, with particles and auxiliaries coming mainly at the right edge of local constituents (pronouns are rare and articles non-existent in Japanese). English ''in Tokyo,'' for example, corresponds to Japanese ''Tokyo de.'' When there is an interruption after *in* in ''in Tokyo,'' repeating *in* will restore continuity to the constituent. But when there is an interruption after *de* in ''Tokyo de,'' repeating *de* will not restore continuity. So by the continuity hypothesis, Japanese speakers have no reason to repeat *de.* By the activation hypothesis, in contrast, Japanese speakers should repeat *de* after interruptions about as often as English speakers repeat *in* after interruptions. But Japanese speakers simply do not repeat post-positional function words like *de.* They deal with delays by other strategies (Fox et al., 1996).

Repeats help restore continuity to a constituent, but so do other techniques as well. Two of these are fillers and elongated syllables. When speakers add *uh* or *um* to their utterances, they create the illusion of greater continuity. A delay containing a filler is subjectively shorter than the same delay with dead silence (Brennan & Williams, 1995). And when speakers elongate words midutterance, they also create an illusion of continuity. Replace the elongation by a pause equal to the added time, and the result sounds more disrupted. The continuity principle appears to lie behind many features of spontaneous speech.

Commitment

Speakers commit themselves to producing one or more constituents—and to meaning something by them—every time they produce a word (stage I). One problem is *when* to commit. They can initiate a new constituent (''I'') immediately after completing the last one (''yes''), as in ''yes, I,'' or they can wait a moment first, as in ''yes, um, I.'' If they commit too early, they may have to suspend their speech immediately, as in ''I uh.'' We have described such a commitment as premature: Speakers commit themselves to a constituent that they are unable to continue. The hypothesis is that some of these commitments are preliminary: At the point at which speakers make

the commitment, they are already expecting, at some level of processing, to suspend their speech immediately. We have offered two types of evidence for this hypothesis.

The first type is phonological orphans. Take the NP ''thiy . the literature,'' in which the first *the* is a separate phonological word, ''thiy,'' pronounced with an unreduced vowel. To produce this, the speaker must have interrupted the syntactic level after the word *the* and sent the word on to the phonological and articulatory levels, expecting to produce an orphan immediately followed by a suspension of speaking. Or take ''a uh (pause) a potential problem,'' in which *a* and *uh* are combined in a single phonological word and resyllabified as ''ei.yuh.'' To produce this, the speaker must have anticipated that he was suspending his speech immediately after *a* and planned to produce it with an unreduced vowel as ''ei'' combined with an enclitic ''uh.'' Note that in both ''thiy . the literature'' and ''ei.yuh (pause) a potential problem,'' the second tokens were cliticized onto the following words, restoring continuity to the constituents *the literature* and *a potential problem.* The evidence we presented shows that the first tokens of many repeats are indeed phonological orphans.

The second type of evidence is near repeats, as in, ''we're in a . an intolerable position.'' With ''a'' and ''an,'' speakers often commit themselves to ''a'' by default, before they know whether it should be ''a'' or ''an.'' We estimated that they did this in 66% of their full and near repeats. With pronouns and contractions such as ''I'' and ''I'm,'' the percentage was 25%. With near repeats, the preliminary commitment isn't to a word but to a constituent type, such as indefinite count NP.

Premature commitments have been observed for a long time (e.g., Maclay & Osgood, 1959). The puzzle is why speakers make them. Why don't they wait until they are confident of proceeding fluently—at least with the second word? We have assumed that they are pressed by a temporal imperative: They must justify any excessive time they take in speaking (Clark, 1996; Goffman, 1981). If they delay too long, they may be heard as opting out, as confused or distracted, as uncertain about what they want to say, or as having nothing immediately to contribute. They can forestall these attributions by producing the first word of the next constituent (even if prematurely) to show that they are engaged in planning the constituent. Alternatively, they could add fillers like *uh* and *um* (Clark 1994; Goffman, 1981; Smith & Clark, 1993), or editing expressions like *you know, I mean, rather,* and *like* to describe why they are delaying (Erman, 1987; James, 1972, 1973; Levelt, 1983, 1989; Schiffrin, 1987). As Jefferson (1989) found, speakers won't tolerate a pause mid-utterance that is more than about one second long. When speakers anticipate too long a pause, they need to deal with it. They can use a filler, editing expression, or preliminary commitment to the next constituent.

Strategies and Processes

One hallmark of strategies is that they are optional, and that is true of the strategies in the commit-and-restore model. When speakers commit to speaking, for example, they have options about both when and how to start. In (1), Reynard could have paused before *I* as long as he wanted, following the strategy, ''Pause until you have formulated the entire next major constituent.'' Or he could have added ''uh'' or ''um'' in that delay, following the strategy, ''Add *uh* or *um* if you anticipate a minor or major delay.'' With preliminary commitments, his strategy might be, ''Initiate speaking with the first word you formulate.'' Good public speakers modulate these strategies depending on the setting.

These strategies depend on pure processes. Each of the strategies is really a high-level decision with low-level consequences. The strategy ''Initiate speaking with the first word you formulate'' leads to a syntactic interruption after the first word (e.g., ''I'') and to its pronunciation as an orphan. The processes at the syntactic and phonological levels may follow without intervention from the high level decision. At the same time, each of the strategies is constrained by certain processes. Speakers cannot apply the strategy ''Initiate speaking with the first word you formulate'' until they have formulated a first word, and that depends on retrieving the word, a process they may have no control over. Or take the strategy, ''After a major disruption restart the current constituent.'' Speakers may find it easier to start articulating from the beginning of a constituent (''I . . .'') than from the middle (''wouldn't . . .''), and that would favor the strategy of restarting. Still, when necessary, good public speakers can follow the high level strategy, ''After a disruption continue the current constituent.'' The choice isn't forced.

Many of these strategies are communicative. Consider the high-level decision to refer to an object—say a dog—that is already part of the speaker and addressee's common ground. This results in the low-level processes of formulating and articulating ''*the* dog'' as opposed to ''*a* dog.'' Speakers are usually unaware of this choice, yet they mean something by it. Their production of *the* is a classic signal, an act by which they mean something, and not merely a symptom, an automatic consequence of some process (Grice, 1991). The same logic applies to the choice of orphans over intact words, such as ''thiy'' over ''thuh.'' That also follows from a high level decision (see Fox Tree & Clark, 1997), and speakers are usually unaware of their choice. And yet they mean something by it—among other things, that they are making a preliminary commitment to a definite NP.

Repeating a word is often viewed as an error, but it is not itself an error. It is a tidy solution to a pair of common problems: how to speak in a timely fashion and yet how to speak smoothly. Repeating a word deserves our respect as an efficient and effective way of dealing with these problems.

REFERENCES

Behaghel, O. (1909/1910). Beziehungen zwischen Umfang and Reihenfolge von Satzgliedern. *Indogermanische Forschungen,* **25,** 110–142.

Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition,* **39,** 173–194.

Bock, K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego: Academic Press.

Boomer, D. S. (1965). Hesitation and grammatical encoding. *Language and Speech,* **8,** 148–158.

Brédart, S. (1991). Word interruption in self-repairing. *Journal of Psycholinguistic Research,* **20**(2), 123–138.

Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language,* **34,** 383–398.

Chafe, W. (1979). The flow of thought and the flow of language. In T. Givon (Ed.), *Syntax and semantics: 12. Discourse and syntax* (pp. 159–181). New York: Academic Press.

Chafe, W. (1980). The deployment of consciousness in the production of a narrative. In W. Chafe (Ed.), *The pear stories* (pp. 9–55). Norwood, NJ: Ablex.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior,* **12,** 335–359.

Clark, H. H. (1994). Managing problems in speaking. *Speech Communication,* **15,** 243–250.

Clark, H. H. (1996). *Using language.* Cambridge: Cambridge Univ. Press.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language: An introduction to psycholinguistics.* New York: Harcourt Brace Jovanovich.

Deese, J. (1984). *Thought into speech: The psychology of a language.* Englewood Cliffs, NJ: Prentice Hall.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition,* **42,** 287–314.

Erdmann, P. (1988). On the principle of 'weight' in English. In C. Duncan-Rose & T. Vennemann (Eds.), *On language: rhetorica, phonological, syntactica: A festschrift for Robert P. Stockwell* (pp. 325–339). London: Routledge.

Erman, B. (1987). *Pragmatic expressions in English: A study of you know, you see, and I mean in face-to-face conversation.* Stockholm, Sweden: Almqvist & Wiksell.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language,* **30,** 210–233.

Ford, M. (1982). Sentence planning units: Implications for the speaker's representation of meaningful relations underlying sentences. In J. Bresnan (Ed.), *The mental representation of grammatical relations.* Cambridge MA: MIT Press.

Fox, B. A., Hayashi, M., & Jasperson, M. (1996). A cross-linguistic study of syntax and repair. In E. Ochs, E. Schegloff, & S. Thompson (Eds.), *Interaction and grammar.* Cambridge: Cambridge Univ. Press.

Fox, B. A., & Jasperson, R. (1995). A syntactic exploration of repair in English conversation. In P. W. Davis (Ed.), *Alternative linguistics: Descriptive and theoretical modes* (pp. 77–134). Amsterdam: Benjamins.

Fox Tree, J. E., & Clark, H. H. (1997). Pronouncing ''the'' as ''thee'' to signal problems in speaking. *Cognition,* **62**(2), 151–167.

Godfrey, J. J., Holliman, E. G., & McDaniel, J. (1992). SWITCHBOARD: Telephone speech corpus for research and development. *Proceedings of the IEEE conference on acoustics, speech, and signal processing.* San Francisco: IEEE.

Goffman, E. (1981). Radio talk. In E. Goffman (Ed.), *Forms of talk* (pp. 197–327). Philadelphia: Univ. of Pennsylvania Press.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech.* New York: Academic Press.

Grice, H. P. (1991). *In the way of words.* Cambridge MA: Harvard Univ. Press.

Hawkins, J. A. (1994). *A performance theory of order and constituency.* Cambridge: Cambridge Univ. Press.

Hayes, B. (1989). The prosodic hierarchy in meter. In P. Kiparsky & G. Youmans (Eds.), *Rhythm and meter* (pp. 201–260). New York: Academic Press.

Holmes, V. M. (1988). Hesitations and sentence planning. *Language and Cognitive Processes,* **3**(4), 323–361.

Horn, L. R. (1989). *A natural history of negation.* Chicago: Univ. of Chicago Press.

James, D. (1972). Some aspects of the syntax and semantics of interjections. *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society* (pp. 162–172). Chicago: Chicago Linguistic Institute.

James, D. (1973). Another look at, say, some grammatical constraints on, oh, interjections and hesitations. *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society* (pp. 242–251). Chicago: Chicago Linguistic Institute.

Jefferson, G. (1989). Preliminary notes on a possible metric which provides for a ''standard maximum'' silence of approximately one second in conversation. In D. Roger & P. Bull (Eds.), *Conversation* (pp. 166–196). Clevedon: Multilingual Matters.

Kenstowicz, M. (1993). *Phonology in generative grammar.* Oxford: Blackwell Sci.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition,* **14,** 41–104.

Levelt, W. J. M. (1989). *Speaking.* Cambridge, MA: MIT Press.

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills.* New York: Springer.

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word,* **15,** 19–44.

Meyer, A. S. (1994). Timing in sentence production. *Journal of Memory and Language,* **33,** 471–492.

Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture-word interference experiments. *Journal of Memory and Language,* **35,** 477–496.

Meyer, A. S. (1997). Conceptual influences on grammatical planning units. *Language and Cognitive Processes,* **12,** 859–864.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology.* Dordrecht: Foris.

Nooteboom, S. G. (1980). Speaking and unspeaking: Detection and correction of phonological and lexical errors in spontaneous speech. In V. A. Fromkin (Ed.), *Errors in linguistic performance* (pp. 87–95). New York: Academic Press.

Nunberg, G., Sag, I. A., & Wasow, T. (1994). Idioms. *Language,* **70,** 491–538.

Schegloff, E. A. (1979). The relevance of repair to syntax-for-conversation. In T. Givón (Ed.), *Syntax and semantics* (Vol. 12). New York: Academic Press.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language,* **53,** 361–382.

Schiffrin, D. (1987). *Discourse markers.* Cambridge: Cambridge Univ. Press.

Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure.* Cambridge MA: MIT Press.

Selkirk, E. O. (1995). The prosodic structure of function words. In K. Demuth & J. Morgan (Eds.), *Signal to syntax.* Hillsdale, NJ: Erlbaum.

Shattuck-Huffnagel, S. (1979). Speech errors as evidence for a serial-order mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett.* Hillsdale, NJ: Erlbaum.

Shriberg, E. E. (1994). *Preliminaries to a theory of speech disfluencies.* Unpublished Ph.D. dissertation, University of California, Berkeley.

Shriberg, E. E., & Lickley, R. J. (1993). Intonation of clause-internal filled pauses. *Phonetica,* **50,** 172–179.

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language,* **32,** 25–38.

Stenström, A.-B. (1987). Carry-on signals in English conversation. In W. Meijs (Ed.), *Corpus linguistics and beyond* (pp. 87–120). Amsterdam: Rodopi.

Svartvik, J., & Quirk, R. (Eds.). (1980). *A corpus of English conversation.* Lund, Sweden: Gleerup.

Wasow, T. (1997). Remarks on grammatical weight. *Language Variation and Change,* **9,** 81–105.

Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language,* **37,** 356–381.

Wheeldon, L., & Levelt, W. J. M. (1995). Monitoring the time course of spoken word production. *Journal of Memory and Language,* **34,** 311–334.

Accepted July 20, 1998