

Sentence Production: From Mind to Mouth

Kathryn Bock

Bock, K. (1995). Sentence production: From mind to mouth. In J. L. Miller, & P. D. Eimas (Eds.), *Handbook of perception and cognition. Vol. 11: Speech, language, and communication* (pp. 181-216). Orlando, FL: Academic Press

I. TOWARD AN EXPLANATION OF LANGUAGE PRODUCTION

The study of sentence production is the study of how speakers turn messages into utterances. Messages are communicative intentions, the things a speaker means to convey. Utterances are verbal formulations. They may be complete sentences or fragments of sentences, well formed or errant, fluent or hesitant. To rephrase a famous formulation of Chomsky's (1965, p. 3), a theory of sentence production is concerned with real speakers who are vulnerable to memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying their knowledge of the language.

The goal of a theory of sentence production is to explain how speakers use linguistic knowledge in the production of utterances. This requires a specification of the knowledge that speakers have and a specification of the information processing system in which the knowledge is put to use. Linguistic knowledge is a component of the speaker's long-term memory. As a practical matter, descriptions of language structure from linguistics are often taken as best guesses about the organization of speakers' knowledge of their language (see Frazier, Chapter 1, this volume). The chief issues in production therefore center on information processing, and include how and when the processing system retrieves different kinds of linguistic

knowledge, how the system uses the knowledge once it has been retrieved, how and when the system retrieves and uses nonlinguistic knowledge, and how the system is organized within and constrained by human cognitive capacities.

What is there to explain? Phenomenologically, talking is not hard. In a lecture delivered at the University of Illinois in 1909, Titchener claimed to be able to “read off what I have to say from a memory manuscript” (1909, p. 8), certainly caricaturing the usual experience, but perhaps not by much. Because of this apparent ease, the interesting problems of production are more readily appreciated, and most often studied, in terms of talk’s typical failures.

The failures range widely. A psychologist reported a transient neurological episode in which he was able to form perfectly coherent messages, but could not express them:

The thoughts can only be described in sentence-like form, because they were as complex, detailed, and lengthy as a typical sentence. They were not sentences, however. The experience was not one of merely being unable to articulate a word currently held in consciousness. Instead, it was one of being fully aware of the target *idea* yet totally unable to accomplish what normally feels like the single act of finding-and-saying-the-word. . . . The idea . . . was as complete and full as any idea one might have normally, but was not an unspoken mental sentence. . . . It was the unusual “gap” in this usually seamless process [of sentence production], a process taken completely for granted in normal circumstances, that amazes me. (Ashcraft, 1993, pp. 49, 54)

William James described the tip-of-the-tongue experience, in which a single circumscribed meaning comes to mind but the corresponding word does not:

Suppose we try to recall a forgotten name. The state of our consciousness is peculiar. There is a gap therein: but no mere gap. It is a gap that is intensely active. A sort of wraith of the name is in it, beckoning us in a given direction, making us at moments tingle with the sense of our closeness, and then letting us sink back without the longed-for term. If wrong names are proposed to us, this singularly definite gap acts immediately so as to negate them. They do not fit into its mould. And the gap of one word does not feel like the gap of another, all empty of content as both might seem necessarily to be when described as gaps. (James, 1890, pp. 251–252)¹

A professor on leave from his university, and looking forward to his return, expressed the opposite sentiment:

“I miss being out of touch with academia.”

A radio-talk-show host, querying a guest about a government official who was reportedly having an affair with a female subordinate, questioned the propriety of the official being

“involved with someone who he works for.”

A secretary, discussing the problems of home renovation and their depiction in a movie called “The Money Pit,” said,

“Have you ever seen the money, ‘The Movie Pit’?”

A lecturer, intending to make the point that verbs are inflected for number, said,

“The verb is afflicted for number.”²

And finally, an umbrella-carrying colleague was asked,

“Is that your elevator?”

These cases and many others reveal that a lot can go wrong between the mind and the mouth. What a theory of sentence production has to explain, apart from such errors, is the course of events on those less notable but more frequent occasions when everything goes right.

When set against the backdrop of normal success in speech, failures of production hint at an information handling system of great complexity but considerable efficiency. Measurements of normal speech rates give average values of about 150 words/min (MacLay & Osgood, 1959) or 5.65 syllables/s (Deese, 1984). Yet syntactic errors are surprisingly rare. In a tape-recorded corpus of nearly 15,000 utterances, Deese (1984) counted only 77 syntactic anomalies, roughly one in every 195 utterances. Heeschen (in press) reported a similarly low incidence of syntactic errors in spoken German. Errors of lexical selection are even less common, with attested rates averaging under one per thousand words (Bock & Levelt, 1994).

Dysfluency is more pervasive, as is the jabberwocky most prominently associated with the 41st president of the United States:

No, we’ve taken—we’re taking steps to fund it. Because I think when you create more export market—OPIC—and that’s exactly what it does—you create more jobs in this country. In this sick and anemic country, which has had—incidentally has grown for the last five [audio technical difficulties]—Hey, just a minute. I haven’t finished yet. (remarks by George Bush in Collegeville, Pennsylvania, September 9, 1992)

Notoriety aside, dysfluency and jabberwocky must be distinguished from outright error. First, they can be consequences of changes to or uncertainty in the message, rather than disruptions in production mechanisms themselves (Goldman-Eisler, 1968; Schachter, Christenfeld, Ravina, & Bilous, 1991). Message uncertainty is more akin to a thinking problem than a talking problem. Second, dysfluency and jabberwocky both seem to display local well-formedness: dysfluencies typically interrupt or punctuate syntactically and phonologically well-formed stretches of speech (Levelt, 1983),

and commonplace jabberwocky, like its literary namesake, is often syntactically well-formed at the same time that it is semantically or pragmatically incoherent.

This chapter surveys some contemporary views about the processes that produce the structural and lexical features of utterances, as well as errors and dysfluencies. The first section sketches a modal view of the processes of production as background for the second section, which highlights some current controversies. The third section looks at fluency and how it is linked to the information processing problems of production. The conclusion touches on some residual questions about the relationships between production and comprehension and between performance and competence.

The treatment of these things is too cursory to do them full justice. More comprehensive surveys and alternative analyses of some of the issues can be found in Levelt (1989) and Bock and Levelt (1994). Bock (1995) reviews the observational and experimental methods that are used in production research. For other topics, I offer directions to supplementary literature throughout the chapter.

II. AN ANALYSIS OF SENTENCE PRODUCTION

Behind most current work in language production there is a widely shared view about the broad outlines of the production system (for a divergent perspective, see McNeill, 1987). In this section I sketch those outlines, reserving until later a discussion of some of the disagreements that have developed within this framework.

Figure 1 gives an overview. It separates processing into three broad components, the message component, the grammatical component, and the phonological component. The vertical connections represent the staggered flow of information over time, with the message leading and phonological encoding trailing. The message captures features of the speaker's intended meaning and communicative perspective (see Clark and Bly, Chapter 11, this volume). The phonological component spells out the sound structure of the utterance, both in terms of the phonological segments of word forms and the prosody of larger units (see Cutler, Chapter 4, and Fowler, Chapter 2, this volume). In between lies the grammatical component. This is the heart of language production, serving to build the bridge from meaning to sound.

A. The Grammatical Component: An Overview

The processes within the grammatical component subdivide into *functional processing* and *positional processing*.³ Functional processing associates message elements with grammatical functions on the one hand and with entries in the mental lexicon on the other. Positional processes associate grammatical

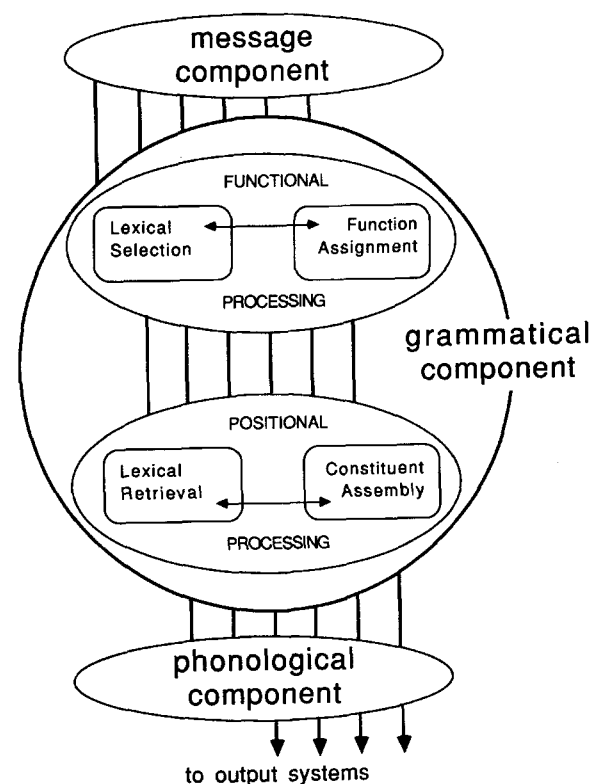


FIGURE 1 The components of the language production system (adapted from Bock & Levelt, 1994). The links among the components are staggered from left to right to represent the flow of information over time.

functions and lexical entries with sentence structures and word forms. Each of these subcomponents thus includes mechanisms that specify words or morphemes (lexical mechanisms) and mechanisms that create word order (structural mechanisms).

In functional processing, the lexical mechanism is lexical selection, which involves the identification of lexical entries that are suitable for conveying the speaker's meaning. These entries are often called *lemmas* in the production literature, and can be likened to the part of a dictionary entry that lists a word (i.e., indicates that an appropriate word exists) and gives its grammatical features. The structural mechanism is function assignment, which assigns and coordinates grammatical roles with features of lexical entries (ensuring, e.g., that the lexical entry destined to serve as the head of the subject phrase is a noun). Positional processing involves lexical retrieval (activating a schematic description of a word's phonological features) and

the creation of a hierarchically structured, ordered set of word and morpheme slots (constituent assembly). In Dell's (1986) model, the lexical work is carried out in the lexical network and the structural work by *tactic frames*.

The grammatical component can be specified more concretely by going through the steps in generating a simple utterance and constructing problems or errors that might arise at each step. Suppose a speaker witnesses an event in which an unknown man gives a stick of licorice to a known boy. Later, in a conversation about the same boy, the speaker describes the event in this way: "A guy offered him some licorice." After the speaker had decided which features of the event to talk about (to encode in a message), what was required in order to convert the remembered event into this utterance?

1. Functional Processing

A major step in functional processing is lexical selection. The speaker must determine that words with the appropriate meanings exist (i.e., find lemmas corresponding to the intended meanings) and mark the lemmas for inclusion in the utterance. This marking, which constitutes selection, makes available the grammatical information that is associated with the word (e.g., whether it is a noun or verb). Several things can go wrong in this selection process. One is that a word is selected whose meaning does not quite match the intended meaning. This produces a kind of speech error called a semantic substitution (Hotopf, 1980), illustrated by our hypothetical speaker saying "A guy offered him some celery."⁴ Another possibility is that two words with closely related meanings may be selected in tandem, creating a blend (Bolinger, 1961; MacKay, 1972) such as "A my offered him some celery" (*my* blending *man* and *guy*). A third problem is not strictly one of lexical selection, but offers some insight into its nature. Sometimes speakers select a word whose meaning they know but whose pronunciation eludes them. This yields a tip-of-the-tongue state (R. Brown & McNeill, 1966), which can be interpreted as successful lexical selection coupled with failed lexical retrieval or phonological encoding.

Another part of functional processing is function assignment, which determines what grammatical roles different phrases will play. It involves assigning syntactic roles such as nominative, accusative, dative, and so on to the rudiments of eventual phrases.⁵ For example, during the formulation of "A guy offered him some licorice," provision must be made for associating individual referents (the known boy, the unknown man, the stick of licorice) with a nominative noun phrase (*a guy*), an accusative noun phrase (*some licorice*), a dative noun phrase (*him*), and a verb that unites them in the intended way (*offered*). Problems in function assignment manifest themselves in phrase exchange errors such as "I got into this guy with a discus-

sion" when "I got into a discussion with this guy" was intended (Garrett, 1980). In such errors, phrases surface in the wrong syntactic role. So, in the error "He wants us to do something else" ("We want him to do something else" was intended), a plural first-person pronoun seems to have been assigned the accusative instead of the nominative function, and a singular masculine pronoun was assigned to the nominative instead of the accusative function. Notice that these are not simple exchanges of word forms, since the attested error is not "Him want we to do something else." A switch in the function assignments in the example utterance, erroneously assigning the dative to *a guy* and the nominative to the masculine pronoun, could yield "He offered a guy some licorice."

2. Positional Processing

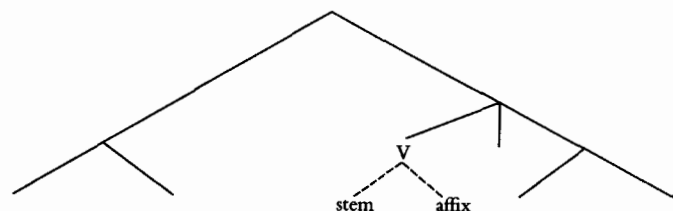
The events of functional processing pave the way for positional processing, which retrieves certain specifications for the elements of the utterance and fixes their locations. Lexical retrieval consists of activating the lexical and grammatical morphemes that fill the slots in this structure. Lexical morphemes are customarily dubbed *lexemes* in the production literature, and they represent such word features as number of syllables, location of primary stress, and segmental (phonological) composition. For example, the lexeme for *licorice* might indicate that it has three syllables with stress on the first, along with a blueprint for its segmental phonology.

Common errors that can be traced to lexical retrieval are phonological word substitutions, in which a word that sounds the same as the intended word substitutes for it (Fay & Cutler, 1977). The intended utterance "make a funny face" turned into "make a fuzzy face" after a phonological substitution. These are sometimes called *malapropisms*. A hypothetical error of this sort in the example "A guy offered him some licorice" might yield "A guy offered him some limerick." A second type of error that is arguably due to a malfunction in lexical retrieval involves the exchange of individual words, as in "You remember . . . our Cambridge in kitchen?" (where the target utterance was "You remember . . . our kitchen in Cambridge?").

Grammatical morphemes consist of inflections and closed-class words. So, to construct "A guy offered him some licorice," the elements *a*, *-ed*, and *some* must be designated. Notably, these elements are rarely involved in errors like the substitutions and exchanges of lexemes (Garrett, 1982). Competing explanations for this apparent invulnerability to error fuel a controversy about whether these elements are retrieved in much the same way as lexemes (e.g., Stemberger, 1984) or are specified in a manner that, in effect, bypasses the need for lexical retrieval (Garrett, 1982).

The other subcomponent of positional processing is constituent assembly. This creates an implicit hierarchy for phrasal constituents and inflectional morphemes. The hierarchy controls the grouping and ordering of

words during production, encoding various dependencies among syntactic functions.⁶ For “A guy offered him some licorice,” the hierarchy can be depicted in this way:



This structure is largely predictable from the types of syntactic functions that have to be represented and from the syntactic features of the selected lemmas. There are few errors that can be unambiguously attributed to the formation of these structures, although certain agreement errors (Bock & Eberhard, 1993; Bock & Miller, 1991) and word deletions are potential candidates. For example, a football broadcaster intoned the commercial message, “Formsby—No one woods as good” instead of “Formsby—No one knows wood as good.” Since the utterance contains a verb slot (its presence can be inferred from the verb inflection *-s*) but lacks a direct object slot, one interpretation of this error (albeit not the only one) is that the direct object occupied the verb’s place. A similar error might yield “A guy himmed some licorice.” This sounds anomalous even as a mistake, in testimony to the fact that structural errors are very rare at all levels. In phonological encoding, as well, the structure of syllables is seldom deformed from acceptable English patterns.

One error feature that is more directly traceable to constituent assembly is known as *stranding* (Garrett, 1982). Stranding occurs because inflections are bound to other words in speech but, in processing, can be divorced from them. So, the past-tense feature on the verb *offered* requires the retrieval or generation of the inflectional morpheme *-ed*. Stranding is illustrated in the utterance of a speaker who intended to say “The dome doesn’t have any windows” and instead said, “The window doesn’t have any domes.” In such errors, bound suffixes (here the plural *-s*) show up in their proper locations in the utterance but affixed to the wrong (often exchanged) word stems.

B. Monitoring

Not shown in Figure 1 is a component that has a major impact on ongoing speech. This is the monitor, which oversees the output of the production system and initiates self-repairs or redirects already-uttered speech. Levelt (1983, 1989) identified the monitor with the comprehension system, and

construed its input as fully formed products of the production system. These products may be either inner speech (the voice in which we silently talk to ourselves) or overt speech. Monitoring should be distinguished from the potential for unconscious modification of the products of processing prior to production (see MacKay, 1987, chap. 9). The questions that have been asked about monitoring have to do with what speakers monitor, when they monitor, and how they repair their speech.

Levelt (1989, chap. 12) offers examples to suggest that speakers monitor every facet of speech for infelicity as well as outright error. Among clearly identifiable errors, there are differences in the likelihood of repair that may reveal something about where attention is directed during speech. Errors of word choice (saying “hot” instead of “cold”) are surprisingly unlikely to be noticed: two studies have shown correction rates hovering around 50% (Levelt, 1983; Nooteboom, 1980). Sound errors may be caught more reliably. In one observational study, they were corrected at a rate of 75% (Nooteboom, 1980).

Blackmer and Mitton (1991) examined the timing of self-repairs in the spontaneous speech of callers to a radio talk-show. These speakers interrupted themselves an average of once every 4.8 s,⁷ sometimes to repair speech errors. For speech-error repairs, the speakers halted 426 ms after the error, overall, and initiated corrections an average of 545 ms after the error. Revision (repairs for style or appropriateness) took significantly longer, averaging 1126 ms between the original utterance and the change. Although they did not break down the production errors, Blackmer and Mitton observed that most of the fastest (under 150 ms) error-to-cutoff times involved the production of aberrant phonemes. This reinforces the suggestion that speakers carefully monitor their articulation as well as the conceptual and pragmatic content of their messages.

How speakers repair their speech depends in part on the kinds of errors they detect. Levelt (1983) found that speakers often acknowledged their errors with the use of editing terms before correcting them (62% of speech errors were so marked vs. only 28% of revisions). The most common editing expression, according to Levelt (1989), is “er”: it exists in many languages with only minor phonetic variations, and in Levelt’s (1983) study of self-repairs, accompanied 30% of all repairs.

III. SOME POINTS OF CONTENTION

The foregoing sketch of the production system can be used to anchor a set of ongoing arguments about how the grammatical component is organized. I will proceed through the questions roughly as they arise within the organization shown in Figure 1, going from top to bottom, although in some cases the issues apply to the architecture of the entire system.

A. From Message to Language

1. Message to Word

The first controversy arises in the transition from the message component to the selection of lemmas. The presupposition behind the transition is that there is some point in processing that is nonverbal, the message, and another that is verbal (cf. Butterworth & Hadar, 1989, vs. McNeill, 1989). One difficulty with this conception arises because languages differ in the kinds of notions that are grammaticized, and therefore in the kinds of message information that will be relevant for grammatical encoding. For example, Chinese lacks the inflectional variation in noun number that is found in English, so there is no need for a Chinese speaker to represent number information in the message unless it is relevant to the communicative situation. English speakers have no such choice, since the vast majority of English nouns indicate number in some way. In the face of these language-conditioned variations in what must be represented in the message, is it defensible to suppose that the message is truly nonlinguistic or, more generally, that there is an abstract thought code ("mentalese") that is completely divorced from language (J. A. Fodor, Bever, & Garrett, 1974, chap. 7)? Put a different way, is there any sense in which an utterance is meaningful on its own (Murphy, 1991) and supposing so, are there limits to that meaning?

One view of these matters is represented in Whorfian claims that the forms of thought are in thrall to the grammatical forms of individual languages (for a wide-ranging review and new evidence, see Lucy, 1992a, 1992b). A compromise can be found in Slobin's (in press) proposal that speakers engage in something that he calls "thinking for speaking," a kind of mental activity whose goal is to package messages in a way that can be verbalized in the target language. So, an English speaker must "think number" in order to speak; Chinese speakers need not. Levelt (1989) endorses a similar view, suggesting that an important component of language learning is the discovery of what features of mentalese must be represented in messages. Levelt observed that once this learning is complete, there is no reason for the representation of the message to interact with the processes of grammatical encoding.

A related debate is over the format of mentalese, and specifically whether its elements, concepts, are wholes that are roughly isomorphic to the lexical vocabulary of language (J. A. Fodor, Garrett, Walker, & Parkes, 1980; Roelofs, 1992) or more primitive features that can be independently combined and separated (Bierwisch & Schreuder, 1992; McNamara & Miller, 1989). These *compositional* and *decompositional* views have a variety of implications for lexical access in production. Levelt (1989, 1992) has worked some of them out in what he terms the *hyperonym* problem, which arises most clearly in a decompositional framework. The problem is that the

semantic conditions for a superordinate term (e.g., *dog*) should be satisfied whenever a subordinate (e.g., *poodle*) is conceptually represented. The solution demands something approximating a one-to-one relationship between concepts and lemmas, as in a compositional framework, or a requirement to use the most specific member of a set of satisfied lemmas (Levelt, 1992).

2. Message to Syntax

The mapping from messages to grammatical structures is equally poorly understood. It seems obvious that meanings and communicative functions somehow constrain the structures of utterances: languages have a tendency to group together syntactically things that belong together mentally (E. V. Clark & Clark, 1978), and speakers modify the structural organization of their utterances in response to subtle changes in the communicative context (Bock, 1977). However, it is unclear how message details serve to specify structural details in the course of normal processing and what the limits of these specifications are (for reviews of various perspectives, see Bates & MacWhinney, 1989; Bock, 1990; Newmeyer, 1992). Figure 2 depicts the leading possibilities in a Chinese-menu format.

The columns in the figure contain the basic choices that must be made in a theory of the mapping from messages to grammatical structures. The theory must include a specification of form–function correspondences (select one from Column A) and a specification of the architecture of the processing system (select one from Column B). Regarding form–function correspondence in Column A, the first position is that the grammatical

Column A FORM-FUNCTION CORRESPONDENCES	Column B ARCHITECTURE OF PROCESSING SYSTEM
ICONIC	INTERACTIVE
AUTONOMOUS	MODULAR

FIGURE 2 Theoretical options for specifying the mapping from messages to grammatical forms in production. To build a production theory, combine one option from Column A (form–function correspondences) with one option from Column B (architecture of processing system). Iconicity implies a tight relationship between messages and linguistic forms; autonomy implies a minimal or irrelevant relationship. Modularity implies a production system that imposes its own operating characteristics on its products independent of message specifications; interactivity implies a production system that updates message specifications in light of ongoing processes.

structure of an utterance is heavily constrained by communicative functions and intentions, both developmentally and in ongoing language use. Newmeyer (1992) terms this *iconicity*, and Bates and MacWhinney (1989) call it level 4 functionalism. The alternative is *autonomy*, the position that the relationship between functions and structures is opaque or irrelevant to the explanation of form.

Moving to Column B, processing considerations are then factored into the theory. Because production processing itself creates constraints, over and above message constraints, certain structural details may be modified in the course of utterance formulation. For example, structures may be organized to permit readily selectable words to be produced before inaccessible words (Bock, 1982; 1986a; 1987b) or to allow easily formulated structures to be produced in place of alternatives (Bock, 1986b; Bock & Kroch, 1989). To the degree that a processing system imposes its own unique constraints on its product, it can be said to be *modular*. If the imposition of these constraints involves modifying or updating the message, the system is instead *interactive*. The result in either case is an utterance structure that may be a compromise between the demands of the message, the demands of the grammar, and the demands of efficient encoding in real time.

In practice, theories of language processing tend to draw their options from the same rows in Figure 2. Theories combine iconic correspondences between form and function with interactive processing systems (e.g., Bates & MacWhinney, 1989), or combine autonomous grammatical components with modular processing systems (e.g., Frazier, 1987). In principle, however, a theory could combine iconicity with modularity (e.g., a categorial grammar like the one discussed by Steedman, 1991, could be implemented in a modular processing system), or autonomy with interactivity (MacDonald, Pearlmutter, & Seidenberg, in press, lean in this direction). Although the full range of positions is currently approached only in theories of language comprehension, the same options are available to production theories.

B. Debates about the Grammatical Component

1. Finding Words

Regarding the grammatical component itself, the central debates have to do with the discreteness of processing at each level. A strong force in these debates is the question of modularity (J. A. Fodor, 1983), particularly informational encapsulation. Informational encapsulation implies that specific cognitive processes are heavily restricted in the kinds of representations that are accessible to them. On one side there is the bias that encapsulation is a property that serves to keep production efficient by restricting the passage

of irrelevant noise from one process to the next (Levelt et al., 1991b). On the other side is the bias that the probabilistic sharing of different kinds of information can support efficient encoding through the use of redundancies (Dell & O'Seaghdha, 1991, 1992). Two specific questions will illustrate how these biases play out.

The first question is whether processing is completed at each level of the production system before activity ensues at the next. This point could be argued at every level of the system, but the focus of current attention is on lexical selection (locating a word's entry) and retrieval (getting sound information from the entry). Selection may be completed before retrieval commences (Butterworth, 1989; Levelt et al., 1991a, 1991b; Roelofs, 1992) or retrieval may be affected by partially activated but unselected lemmas (Dell & O'Seaghdha, 1991; Harley, 1993; Peterson, Shim, & Savoy, 1993). Figure 3 illustrates the difference in a simplified way. Discrete activity, shown in (A), activates only the sounds /kæt/ after selection of the lemma for *cat*, whereas cascaded activity (B) results in partial activation of sounds that belong to words related to *cat*, like the /r/ of *rat* and the /d/ of *dog*. The empirical focus of this debate is the *mixed error effect*. Mixed errors are word substitutions that resemble both the meanings and the sounds of an intended words (e.g., saying "rat" instead of "cat"). Mixed errors appear to be more probable than would be expected from a discrete system (Dell & Reich, 1981; Martin, Weisberg, & Saffran, 1989), lending support to the existence of cascaded activation.

The second question has to do with the existence of feedback between levels. A conservative view of production is that encoding proceeds in a strictly top-down fashion with no feedback from lower to higher levels (Garrett, 1976; Schriefers, Meyer, & Levelt, 1990). The alternative, worked out most thoroughly by Dell (1986), is that lower-level activity may influence higher-level processing. In Figure 3B, for example, the bidirectional arrows symbolize a two-way flow of information. At stake empirically is the explanation of a word superiority effect, or *lexical bias*, for sound errors: Sound errors are more likely to result in real words than chance would predict (Dell, 1986; Dell & Reich, 1981). Dell's explanation for this tendency involves the tuning created by interactive activation, which causes activity at lower levels to be modulated by rebounding activation from higher levels. In Figure 3, an erroneous /r/ is more likely to be retrieved than an erroneous /d/, although both sounds receive some activation from words related to *cat*. The /r/ reaps an additional benefit from the activation that rebounds from *rat*. This rebounding activation is a secondary consequence of the activation that feeds back from the /æ/ and /t/ shared by *cat* and *rat*. Because only real words produce this secondary effect, real words are a more likely product of sound errors than chance would predict.

The alternative to the interactive activation account for such errors is an

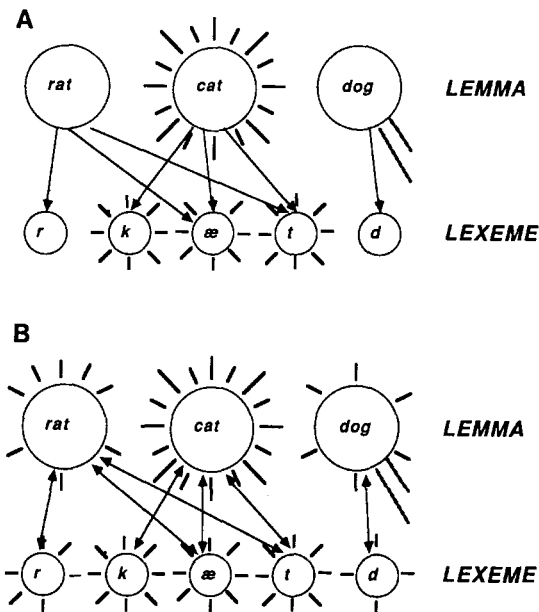


FIGURE 3 Simplified processing models for relating lemmas (word entries) to lexemes (word forms) in word production. Halos around nodes indicate activated information, and denser halos denote more activation. In (A), activation is discrete, so that selection at each level is all or none, and there is no feedback of activation from lower to higher levels. In (B), activation is cascaded, allowing partial activation of information related to unselected information at the level above, and activation feeds back from the lexeme to the lemma level. The probability of selection depends on relative amounts of activation. At the lemma level, *cat* is most likely to be selected, paving the way for selection at the lexeme level, where /k/, /æ/ and /t/ are most likely to be selected. An erroneous selection at the lemma level would tend to involve the *rat* lemma rather than the *dog* lemma, yielding a so-called mixed error. An erroneous selection at the lexeme level (given correct selection at the lemma level) would be more likely to involve the /r/ than the /d/ node, yielding a sound substitution that happens to create a word (*rat*).

editing mechanism that filters the products of processing. There is a wealth of different ways in which editing could work. The editor could be a gatekeeper that hands down only one decision from one level to the next. Alternatively, it could hand down all well-formed candidates, with well-formedness arbitrated by consistency with grammatical knowledge operating (in some unspecified way) outside of the performance system. There are also different loci at which editing could occur. There could be a form of editing at every level (MacKay, 1987; van Wijk & Kempen, 1987), at some subset of levels, or only prior or subsequent to final output (Levelt, 1989).⁸ An editor that serves as a gatekeeper at every level of the production system, passing only one candidate unit or structure from each level to the next,

creates a completely discrete, top-down processing system like the one in Figure 3A. Within this framework, the lexical bias in sound substitutions results from editorial biases. The editor may be more likely to accept real word than nonword errors (see Baars, Motley, & MacKay, 1975, for supporting evidence). Mixed errors in word substitutions constitute a bigger challenge, so far not satisfactorily addressed.

2. Finding Places for Words

The preceding questions about grammatical encoding concern the “vertical” relationships in Figure 1. There is another set of questions, less well developed in the literature, that have to do with “horizontal” relationships between the lexical and structural processes at each level. Broadly stated, the problem is that words must find places within structures, and not just any place will do. Bock (1987a) called this the *coordination problem*.⁹ A common theoretical solution is a slot and filler mechanism (Dell, 1986; Garrett, 1988; Shattuck-Hufnagel, 1979) in which a structural frame provides slots into which words or sounds must be inserted, and the words and sounds carry information about the kinds of slots they require. Coordination is then a simple matter of fitting pieces into place. After the manner of Dell (1986), Figure 4 shows a simple slot and filler arrangement accompanying a lexical network.

Constraints on speech errors make this kind of mechanism plausible, because exchange errors of all kinds usually involve units of the same lin-

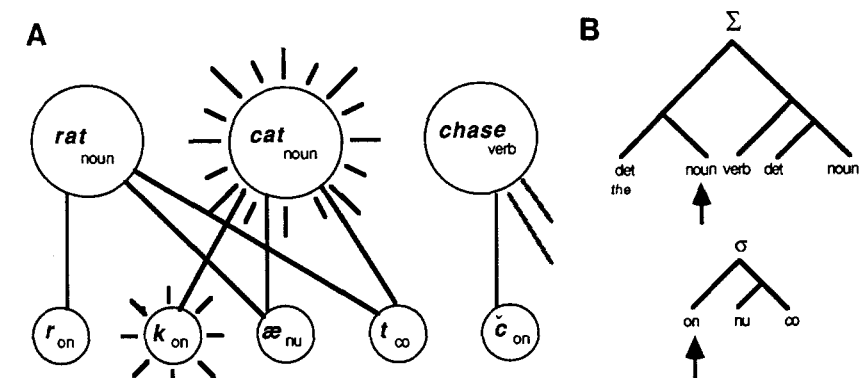


FIGURE 4 A slot and filler scheme for utterance (Σ) and syllable (σ) structures. The units within the lexical network (A) are labeled according to filler type (noun, verb, on = syllabic onset, nu = syllabic nucleus, co = syllabic coda). The structural frames (B) have matching labels to constrain the types of fillers that may be inserted into each terminal position or slot. The pointers beneath each structure indicate the slot currently being filled from the network. See Dell (1986) for detailed discussion of a model of this kind.

guistic type. However, the mechanism can be notoriously ungainly. The features of the frames are stipulated rather than generated, information is multiply represented, and representations are sometimes taken apart only to be put back together again. For example, in Shattuck-Hufnagel's (1979) model, a word is retrieved from the lexicon with its sounds properly ordered, but the sounds are then copied, one by one, into an independently generated frame. The emergence of more elegant theoretical alternatives to the slot and filler approach (e.g., Dell, Juliano, & Govindjee, 1993) should serve to fuel interest in alternative processing mechanisms for solving various coordination problems in production. So far, however, these mechanisms do not go far enough on their own to solve the coordination problem for syntax (see Dell & O'Seaghdha, 1993, for discussion).

a. Incrementality and Parallelism in Coordination

One issue that is latent in the coordination problem concerns incrementality (V. Ferreira, 1993) and, intimately related to it, the scope of planning units in encoding. Incrementality refers to the word-by-word character of speech. This creates a challenge for speakers, because there is a certain amount of parallelism in thinking and in messages: one can seemingly hold in mind all at once a complex idea or a rich visual image. To communicate the information, it may have to be dissected into propositionlike elements (Levelt, 1981) and known concepts (Levelt, 1992) and the pieces arrayed in time. The tension between parallelism and incrementality shows up in various speech errors. Certain kinds of blends, exchanges, and anticipation errors might be taken as reflections of simultaneity in the preparation of elements that are asynchronous in the spoken utterance.

Where is linearity imposed? The most common solution puts a hierarchical frame in control of the sequencing of terminal elements, and calls in some fashion on slot and filler mechanisms. The frames impose an order on elements (words, sounds, etc.) that may be available in parallel (Dell, 1986; Garrett, 1988). Because the scope of the frame becomes smaller as an utterance develops from a message to an articulatory plan, the sizes of the planning units change across the levels of Figure 1. Although there is no requirement in such models that the planning of high-level units be completed before work on lower-level units can begin (Garrett, 1976), there is typically no account of when or how the frames themselves are constructed.

A somewhat different solution ties incrementality to the order in which words (technically, lemmas) become accessible in the mental lexicon (Bock, 1982, 1987a; Bock & Warren, 1985; Kelly, Bock, & Keil, 1986; Levelt, 1989; McDonald, Bock, & Kelly, 1993) and builds sentence structures around the words as they are selected (V. Ferreira, 1993; Kempen & Hoenkamp, 1987). The structures that are created are drawn from the possibilities licensed by the language, but their assembly is opportunistic and carried out in ongoing

time. For instance, a speaker might tend to say "The bird fed a worm to the nestling" if the word *worm* is selected earlier than *nestling*, but say "The bird fed the nestling a worm" if the lemma for *nestling* is selected first.

The attraction of this solution is that it accounts at once for the seriality of speech and for the intuition that words normally emerge as they are retrieved, often without logjams or temporal chasms. Still, there must be mechanisms that implement grammatical constraints at the same time that they juggle words, and these mechanisms must be explicit within a production theory.¹⁰ The theory must also explain how word order in an utterance is negotiated with respect to the elements of a message, a problem with a long history: Levelt (1989, p. 26) attributes to Wundt the view that "word order follows the successive apperception of the parts of a total conception," and notes the difficulties that this hypothesis confronts when it is faced with word-order constraints in languages. Any general solution to the problem will require accounts of whether and how the dissection of messages is constrained so as to deliver chunks for grammatical encoding on a manageable schedule, whether and how choices among competing realizations are made in the (possibly frequent) event that several words are available at the same time, and whether and how information is buffered when it is delivered ahead of schedule. This leads directly to the issue of minimal units of preparation.

b. The Scope of Coordination

The contrast between incrementality and parallelism in coordination has obvious implications for utterance preparation. The larger the minimal units of coordination (e.g., word, phrase, clause), the wider the scope of planning must be, and the greater the simultaneity in processing. In slot and filler approaches, differences in the sizes of planning units at different processing levels are explicitly linked to differences in linguistic units. For example, Garrett (1988) suggested that the scope of functional planning is roughly two clauses, and that of positional planning roughly a phrase, as reflected, respectively, in the spans of word exchange errors and sound exchange errors. Other work has also pointed toward clauselike units of higher-level planning (Bock & Cutting, 1992; Ford, 1982; Ford & Holmes, 1978).

However, the scope of preparation need not be tied strictly to linguistically defined units. An alternative, one that is more consistent with lexically driven incrementalism, envisions a flexible deployment of information processing capacity across linguistic levels (Bock, 1982; Huitema, 1993; Levelt, 1989). The presumption is that variations in the construction and use of structural frames reflect changes in cognitive capacity and general constraints on the size or complexity of individual structural realizations (e.g., the length of an individual phrase or word), as well as hierarchically

defined variations in structural units. So, the scope of preparation might sometimes encompass an entire simple utterance (e.g., "here it is") and sometimes no more than one complicated word (e.g., "eponymously").

3. Positional Processing

a. Lexical Retrieval

There are two hot spots in work on lexical retrieval in production, both of them related to the role of word frequency. One has to do with the precise locus of word-frequency effects in production, mirroring a related controversy in the study of reading (Balota & Chumbley, 1985, 1990; Monsell, 1990, in press; Monsell, Doyle, & Haggard, 1989; Savage, Bradley, & Forster, 1990), but with a different set of concerns. There is evidence that the production of words in the course of naming pictures is affected by word frequency (Huttenlocher & Kubicek, 1983; Oldfield & Wingfield, 1965), recently confirmed in detail by Jescheniak & Levelt (1994). Frequency also affects the likelihood of speech errors in words and sounds (Dell, 1990; del Viso, Igoa, & García-Albea, 1991; Kelly, 1986; Levitt & Healy, 1985; Stemberger & MacWhinney, 1986b), word order (Fenk-Oczlon, 1989), the spoken durations of words (Wright, 1979), and word length (Landauer & Streeter, 1973).

The question is where the effects of frequency arise. One theoretical view links the impact of frequent use most tightly to those components of processing that require the retrieval of stored information rather than inference, construction, or application of grammatical rules (Pinker, 1991). A different answer, implicit in connectionist views of language processing, is that frequency affects everything (Seidenberg & McClelland, 1989). Although the question is far from settled, some observations (del Viso et al., 1991; Kelly, 1986) and experimental results (Jescheniak & Levelt, in press) favor lexeme retrieval as the locus of frequency effects. This makes sense from the standpoint that lexeme retrieval is, in fact, retrieval, whereas lemma selection may demand inference in order to abstract a message from the communicative context, dissect a message into meanings that can be lexicalized, and map meanings to lemmas (as in a meaning-postulates framework; J. D. Fodor, Fodor, & Garrett, 1975).

Other evidence, however, points indirectly toward the lemma as a source of frequency effects. Using an error-eliciting procedure, Dell (1990) found that sound errors were more likely to occur on low-frequency than on high-frequency words, and that the occurrence of errors involving homophones like *by* and *buy* was related to the combined frequency of the words. The homophone result showed that low-frequency homophones inherit the benefits of having a high-frequency twin. This implies a lexeme locus for the frequency effect, since homophones share the same lexeme but have differ-

ent lemmas. Paradoxically, however, frequency did not reliably affect the strength of lexical bias: high-frequency words were not significantly more likely to be the byproducts of sound error than low-frequency words. The absence of this frequency difference is inconsistent with the lexeme hypothesis, if lexical bias results from rebounding activation between lexemes and segments. Simulating all of these effects, Dell found that the best account was given by an interactive activation model in which frequency is represented at the lemma level. The same model readily explained homophone inheritance in terms of interactions between lemmas and lexemes.

Frequency is also at the center of a controversy over the closed-class vocabulary. As noted in Section IIA2, the closed-class elements of utterances are sometimes accorded a special status in generation, with the view being that they are inserted by grammatical mechanisms rather than lexical ones. One aspect of this view is a claim for syntactic immanence (see Bock, 1989): the constituent tree carries features that may uniquely specify a particular element (e.g., if a slot calls for a determiner that is indefinite and singular, English offers only one possibility¹¹). A second aspect is a claim for special access: rather than being retrieved from the general lexicon, these elements come from a special store. A variant of this is Levelt's (1989) indirect access hypothesis: closed-class words are recruited by way of the syntactic information stored with the lemmas for open-class words. The evidence for a special encoding route comes from several sources, prominently including the vulnerability of the closed class to disruption in aphasia and its relative invulnerability to normal error (Garrett, 1982, 1992; Lapointe & Dell, 1989). Although the supposed invulnerability to error has been challenged by evidence that error probabilities are highly sensitive to frequency (Dell, 1990; Stemberger, 1984), the evidence from aphasia remains troublesome. As Garrett (1992) notes, many aphasics disproportionately lose the elements of the closed class, which are among the most frequent elements in the language. The paradox is that in the open-class vocabulary, the elements most likely to be lost are the infrequent ones.

b. Constituent Assembly

A different question about the effects of frequency arises with respect to constituent assembly. Bock (1986b, 1990; Bock, Loebell, & Morey, 1992) marshaled evidence that speakers tend to repeat constituent structures across successive utterances. For example, a speaker who employs a double-object dative in one utterance (perhaps, "The governess made the princess a pot of tea") is subsequently more likely to use a double-object dative than a prepositional dative, even in a completely unrelated utterance (so, "The woman is showing a man a dress" becomes relatively more likely than "The woman is showing a dress to a man"). There is suggestive but still inconclusive evidence that this effect strengthens with the repeated use of a structure (Bock

& Kroch, 1989), raising the possibility that the creation of sentence structures is a frequency-sensitive process. But since structural repetition does not appear to depend on the repetition of particular words (Bock, 1989) or thematic roles (Bock & Loebell, 1990) or features of word meaning (Bock et al., 1992), its explanation rests on the transient enhancement of general structural procedures rather than the enhanced retrieval or use of specific stored information. If this is so, it becomes difficult to use frequency effects as symptoms of simple retrieval processes. More broadly, it points to mechanisms of syntactic performance that undergo continuous modification with use, counter to the standard view of syntax as something whose use is fixed beyond childhood.

IV. FLUENCY

One issue that has yet to attain a sure foothold in the contemporary study of production is the question of what supports fluency. Although considerable attention is paid to the absence of fluency in errors and hesitations (Garrett, 1982, reviews these literatures together), the mere absence of trouble does not explain fluent speech. H. H. Clark and Clark (1977, pp. 261–262) characterized a hypothetical *ideal delivery* as one in which “people know what they want to say and say it fluently,” with “most types of clauses . . . executed in a single fluent speech train under one smooth intonation contour” and pauses restricted to grammatical junctures. This characterization emphasizes a regular speech rate with fluid prosody. The grounds for this ideal might be placed either in the wants and needs of the listener (H. H. Clark & Clark, 1977) or in the speaker’s own desire for proficient action. In either case, fluency is an ideal to which speakers aspire.

Two different explanatory traditions offer cognitive accounts of fluency. One emphasizes the salutary effects of an efficient mechanism, with efficiency achieved either by dint of an elegantly tailored architecture or by dint of practicing cognitive tasks. I will call these *mental skill* explanations. The other tradition stresses the role of attention or working memory resources in the creation of speech. I will call this a *mental energy* explanation. After reviewing some of the production work that falls into these two traditions, I briefly consider two other components of fluency that have so far gone mostly unmentioned, prosody and preparation.

A. Mental Skill and Mental Energy

Most current models of production offer a skill-based explanation of fluency that emphasizes efficiencies of architecture. Although the models differ over what creates an efficient architecture, whether it is something like informational encapsulation or the exploitation of informational redundan-

cies, there has been little consideration from either standpoint of how the system comes to have the form that it does, or becomes as efficient as it is. Conceivably, the parser and production system are innate (Pinker, 1984), but just as conceivably, they organize themselves as the grammar is learned. There has likewise been little consideration of how the parameters of fluency (speech rate, prosody, preparation, etc.) change with experience, particularly across complete utterances (though see Wijnen, 1990). With the emergence of interest in developmental language performance problems (Gerken, 1991; Plunkett, 1993; Stemberger, 1989; Wijnen, 1990, 1992) and in production models that can change features of their organization with training, more emphasis may be laid on these issues.

A different, sometimes complementary tradition in cognitive explanation stresses mental capacity over the organization of mental skills. Some of the extensive research on hesitations in speech (filled and silent pauses) has been carried out in this framework. Garrett (1982; also see Goldman-Eisler, 1968) observed that hesitations may reflect normal advance planning and retrieval for an upcoming structural unit (Beattie, 1983; Boomer, 1965; Butterworth, 1980; Ford, 1982; Goldman-Eisler, 1972; Holmes, 1988; see Bock & Cutting, 1992, for review), or delay created by the momentary inaccessibility of a needed piece of information (Lounsbury, 1965; Smith & Clark, 1993), or transient increases in processing load. Garrett dubbed these three culprits “wait till the boat’s loaded,” “it’s in the mail,” and “don’t bother me, I’m busy.” The last of these implies that changes in fluency can be traced to transient variations in cognitive capacity that may affect production globally or locally. Experimental efforts to create variations of this kind have examined the effects on speech and writing of situational anxiety (Mahl, 1987; Reynolds & Paivio, 1968), secondary task demands (J. S. Brown, McDonald, Brown, & Carr, 1988; Power, 1985, 1986), and individual differences in working memory capacity (Daneman & Green, 1986). Errors can be analyzed from a similar perspective (Fayol, Largy, & Lemaire, 1994).

One intriguing relationship that has emerged within the capacity-limitations framework is a link between the duration of traditionally defined short-term memory traces and speech rate (Baddeley & Hitch, 1974; Schweickert & Boruff, 1986). Beginning with Baddeley, Thomson, and Buchanan (1975), evidence began to appear that the measured capacity of short-term memory is sensitive to the amount of time required to articulate words. Generally, people can immediately recall about as many items as they can pronounce in two seconds. This interacts with differences among languages in the normal sizes of the lexical units employed in short-term memory tasks (Cheung & Kemper, 1993; N. C. Ellis & Hennelly, 1980; Hoosain, 1982; Stigler, Lee, & Stevenson, 1986; Zhang & Simon, 1985) and with developmental differences in speech rate (Hulme, Thomson, Muir, &

Lawrence, 1984; Kynette, Kemper, Norman, & Cheung, 1990). The appearance is that various components of phonological encoding in normal language production may be closely identified with the components of verbal short-term memory, and are limited in similar ways (A. W. Ellis, 1980a, 1980b; see Cowan, 1992, and Schweickert, Guentert, & Hersberger, 1990, for qualifications).

B. Prosody

Integral to the ideal delivery is a smooth intonation contour. To achieve this, English speakers must control their speech rate, the relative timing of stressed and unstressed syllables, changes in amplitude, and changes in fundamental frequency. The determinants of these factors are of many kinds. According to Cutler and Isard (1980), they fall into four main categories: the stress patterns of individual words, the location of the main accent in an utterance (sentence accent), syntactic structure, and various pragmatic factors (including discourse structure, speech acts, speaker's attitude, and so on).

Interest in the production of prosody has centered on two of these factors. The first is the role of syntactic structure in determining prosodic boundaries. There is obviously a relationship between syntactic structure and prosodic structure, with pauses, slowing of speech, and drops in amplitude and fundamental frequency tending to occur at major syntactic breaks (Cooper & Paccia-Cooper, 1980; Cooper & Sorenson, 1981; Steedman, 1991). A question has been raised about whether these prosodic features are under the direct control of syntactic structures, or are instead created by an intermediate representation that makes its own contribution to prosody. F. Ferreira (1993) argued for the operation of an abstract prosodic structure, showing experimentally that changes in the durations of words and any pauses that follow them can reflect prosodic boundaries that are different from syntactic boundaries. Ferreira also reported striking evidence that the prosodic representation is not dependent on an utterance's specific speech sounds. She found that the combined duration of a word and any pauses that followed it was not predictable from the segmental contents of the words: longer words were followed by correspondingly shorter pauses and vice versa. This suggests that a prosodic structure can be created without reference to the sounds that comprise individual words (though data from Meyer, *in press*, dispute this possibility).

A second, related factor in the production of prosody is lexical stress and how it is distributed throughout an utterance. In English, words differ regarding whether they typically bear stress and, if they bear it, where. Function words commonly lack stress, but among content words in natural speech, nearly 90% are monosyllabic (*i.e.*, stress bearing) or have stress on

the first syllable (Cutler & Carter, 1987). There is a long-standing view (*e.g.*, Pike, 1945) that English speakers try to space these stresses relatively evenly over time. Although this appears to be far from accurate (see Levelt, 1989, chap. 10), a grosser kind of rhythmicity is created by a tendency to alternate strong and weakly stressed syllables. To achieve this, speakers subtly shift the stress pattern of words (Kelly, 1988; Kelly & Bock, 1988) or arrange words to avoid long sequences of weak or strongly stressed syllables (Kelly & Rubin, 1988).

Errors involving stress and intonation sometimes occur, and these, too, have been examined for what they reveal about the production of prosody. In line with the evidence for rhythmic alternation, Cutler (1980b) showed that syllable deletion errors (*e.g.*, "Next we have this biCENTial rug" instead of "bicenTENnial") tend to improve the rhythm of utterances, as do erroneous movements of lexical stress ("we do think in SPEcific terms," instead of "speCIfic"). Perhaps the best-known observation about prosody in speech errors is that sentence intonation tends to be preserved in the face of other errors: Garrett (1975) gives as an example the error, "stop beating your BRICK against a head wall" (produced in place of "stop beating your HEAD against a brick wall"). However, Cutler (1980a) contended that this holds only for open-class (content) words. When closed-class words move, they tend to carry sentence accent with them, if they bear it, as in, "you can turn it ON back now" (instead of "you can turn it back ON now"). Levelt (1989) interpreted these contrasting patterns as reflecting differences in when the accent-bearing elements actually move. On this view, open-class movement errors occur during grammatical encoding, and do not change the syntactic structure over which prosody is calculated, whereas closed-class elements move after their metrical patterns are specified, during phonological encoding.

C. Preparation

Speakers engage in a certain amount of preparation before they begin to talk. They take longer to initiate long utterances than short ones, and fundamental frequency is higher at the beginning of long utterances than short ones (Cooper, Soares, & Reagan, 1985). Oddly, despite the considerable attention given to hesitation and other forms of dysfluency in connection with questions about speech planning, comparatively little is known about changes in fluency as a consequence of variations in preparation. An utterance that is fully prepared (phonologically encoded into inner speech) can probably be uttered more fluently than an utterance whose preparation falls short of this (see Sternberg, Knoll, Monsell, & Wright, 1988). However, most utterances undergo less than full preparation because of the limit on the amount of fully formulated material that can be held in immediate

memory, and because utterances can be launched before this limit is reached. Speakers often begin talking not too long after they formulate the first content word in the subject phrase (Huitema, 1993; Lindsley, 1975).

Between the extremes of maximal and minimal preparation, differences in the amount and type of preparation may have consequences for fluency. Huitema (1993) showed that about 95% of a set of simple descriptive utterances were produced fluently when speakers had a minimum of 1550 ms of pure planning time, versus 60% when the minimum planning time was 600 ms. Huitema also found that complex subject noun phrases consume planning resources to the point that speakers hesitate or slow their speech after producing them, presumably to prepare the rest of the utterance. With regard to lexical planning, Butterworth (1980, 1989) argued that content words are produced more fluently (with fewer pauses) when their lemmas have already been selected than when both the lemma and the lexeme must be prepared. Similarly, Deese (1980) suggested that speech whose content (but not form) is planned in advance displays less disruptive hesitation patterns than completely extemporaneous speech.

Less explicit kinds of preparation come in the form of contextual priming and interference, caused either by one's own prior speech or that of an interlocuter. Relationships of similarity or identity across the elements of sentences may affect the speed with which utterances are initiated (Dell & O'Seaghdha, 1992; Schriefers, 1992, 1993), the fluency with which they are produced (Bock & Loebell, 1990; Sevald & Dell, 1992; Sevald, Dell, & Cole, 1993), their word order (Bock, 1986a, 1987a; Bock & Irwin, 1980), their syntactic structure (Bock, 1986b, 1990), the ease of lexical selection and phonological encoding (Meyer & Schriefers, 1991; Roelofs, 1992, 1993; Schriefers et al., 1990; Wheeldon & Monsell, 1994), and the likelihood of error (Baars et al., 1975; Dell, 1984; Harley, 1990; Martin et al., 1989; Stemmer, 1985; Stemmer & MacWhinney, 1986a).

D. Summary

Because of the traditional emphasis on errors and dysfluency in production research, the positive components of fluency have received little direct attention. In some cases, explanations of error and dysfluency will help to shed light on their fluent speech counterparts. This seems particularly likely for hesitant speech, where fluency is the background against which hesitations are assessed. Likewise, the study of speech errors may help to disclose the joints in the architecture of normal language production. In other ways, however, fluency is more than the absence of dysfluency, and its explanation depends on characterizing how the disparate elements of language can be integrated into a smooth stream of speech.

V. CONCLUSION: THE PC PROBLEMS

From the broad standpoint of psycholinguistic theory, two overarching questions about language production have to do with how the processes of production (P_1) are related to the processes of comprehension (C_1), and how the processing theory for production (a performance theory, P_2) is related to a linguist's theory of language knowledge (a competence theory, C_2). These are the PC problems.

A. Production and Comprehension

Views of the relationship between production and comprehension run the gamut from close identification of the performance systems (MacKay, 1987; MacKay, Wulf, Yin, & Abrams, 1993) to substantial dissociations of them (Frazier, 1982; Shallice, McLeod, & Lewis, 1985). Although it may turn out that production is simply "comprehension turned upside down," in what follows I give some reasons for skepticism about this possibility. The main point is that differences in what a processing system has to do create differences in how the processing system works (Marr, 1982).

First, differences in the goals of the two systems make it reasonable to suppose that there may be differences in their normal functioning. The goal in comprehension is to create an interpretation of an utterance. In production, the speaker's immediate goal is simply to create an utterance. The utterance should be adequate to convey the speaker's meaning, but to do so, it must meet a range of constraints that are specified in the grammar of the language. This necessity gives the production problem a spin that Garrett put like this: "The production system must get the details of form 'right' in every instance, whether those details are germane to sentence meaning or not" (1980, p. 216). As a result, a general issue in a theory of production is to explain how speakers create linguistic structures at all levels. There can be no argument about whether syntax, for example, is "important" in production, because the speaker as a matter of course creates those features of utterances that we call their syntax. It is considerably less clear whether language comprehension requires that listeners reconstruct these features to the same level of detail.

Beyond this, differences in the starting points for production and comprehension also lead one to expect differences between them. Listeners must piece interpretations together from degraded and ambiguous input that arrives one syllable at a time. Speakers begin with an interpretation, in the form of a message. The message may be apprehended all at once, and without ambiguity. The hurdle is to assemble the linguistic pieces that convey the communicative intention.

Despite these differences, it is unlikely that the linguistic resources employed in comprehension are entirely independent of the resources for production. Communication is successful because words and their arrangements mean the same thing regardless of whether one's role is that of speaker or listener. One implication is that there is an important role in performance theories for general descriptions of the language knowledge that is called upon in comprehension and production. More controversial is the role that performance data can play in informing theories of linguistic knowledge.

B. Performance and Competence

One of the weightiest questions about the relationship between language knowledge and the systems that put that knowledge to use is whether the latter are in any important sense specialized for the implementation of the former (see Frazier, Chapter 1, this volume). The view that performance systems are linguistically neutral (and therefore linguistically uninteresting) is well entrenched in linguistics, inaugurated by Chomsky's (1965) famous formulation of the distinction between competence and performance. J. A. Fodor (1983) put forward a very different view, according to which the performance systems that subserve language are dedicated, largely segregated from other cognitive and perceptual systems, and innate. Dedicated performance systems, in turn, may be distantly related to the grammar of the language or may transparently encode it (J. D. Fodor, 1981, 1989).

Underlying the view that performance systems transparently encode the grammar is the *strong competence assumption*, which says that a linguistic theory captures knowledge as the speaker actually represents it. In some quarters (e.g., Bresnan & Kaplan, 1984; J. D. Fodor, 1989), this motivates an interest in language performance as one way of testing hypotheses about knowledge representation. On this view, when performance data diverge from the predictions of a linguistic theory, the data count against the theory.

In part because most psycholinguistic data reflect language comprehension, comprehension research has played a more substantial role than production research in informing competence theories. The undeniably linguistic nature of the input to comprehension also helps to drive interest: there is no need to grapple directly with the unknown commerce between language and thought. Moreover, the study of comprehension is closely tied to linguists' interest in language acquisition, insofar as the acquisition problem is in the first instance seen as a problem of making sense of linguistic input.

For several reasons, production data merit comparable weight in developing theories of linguistic knowledge. First, speakers are, in a sense, the

proximate cause of the features of language. They create linguistic structure whenever they talk. Second, although variation in a speaker's use of structures is often ascribed to a concern for making comprehension easier, it is likely that some variations serve only to make production easier (P. M. Brown & Dell, 1987; Dell & Brown, 1991; Keysar, 1994; Schober, 1993). This means that linguistic explanations that incorporate performance constraints (e.g., Hawkins, 1990) cannot overlook the problems faced by speakers, which may be different from those faced by comprehenders and just as decisive in shaping the grammar. Third, questions about language acquisition motivate attention to production at least as much as attention to comprehension. Children's speech is the main source of current data about the course of language development, and because knowledge acquisition and skill acquisition are inextricably intertwined in this development, it is impossible to understand one without the other. Fourth, speaking is a cardinal facet of language ability: the commonsense test of skill in a language is not the ability to understand it, but the ability to speak it. And finally, production data can provide converging validation for competence explanations of comprehension evidence. This convergence is essential if linguistic theory is to be a theory of the knowledge that is used by all of the cognitive systems that perform language.

The problem of characterizing the conceptual input to production remains. However, with the advent of experimental methods that tap ongoing processes of production without changing a speaker's communicative intention (see Bock, 1995), it has become possible to temporarily bypass some of the thornier questions about relationships between language and thought. This makes the study of production more tractable and the prospects for satisfactory theoretical tests brighter.

One last cautionary point concerns the nature of the relationship between a production theory and linguistic theory (see Frazier, Chapter 1, this volume). There remains in some quarters a tendency to regard language production as the animation of a linguistic theory. This is a vestige of the early days of transformational generative grammar, when the mathematical term "generative" was sometimes confused with the processes of generation that production requires. The basic constructs of transformational theories (rewrite rules, transformations), linking abstract structures to concrete linguistic details, also appeared to trace a route for production. Efforts to use linguistic theory as a performance theory encouraged these misleading equations (see J. A. Fodor & Garrett, 1966, for review). It should be clear by now that production theory is a theory of using language, not an account of how language is represented, and that linguistic frameworks whose province is the static organization of language knowledge are unlikely to provide the theoretical machinery for explaining how that knowledge is deployed in time.

Acknowledgments

Several people aided in the preparation of this chapter, and I would like to acknowledge their help, with gratitude. In chronological order, they are Gary Dell, John Huitema, William F. Brewer, and Gregory Murphy. While writing the chapter, I received support from grants awarded by the National Science Foundation (BNS 90-09611) and the University of Illinois Research Board.

Endnotes

1. There are important differences between the "gap" that James describes and the "gap" that Ashcraft describes. Ashcraft's gaps seem to contain propositions, the elements of what I am here calling messages. James's gaps contain concepts, the meanings behind single words.
2. This error was reported in a scrawled note along with the annotation "poor verb."
3. The division into functional and positional processing generally corresponds to the *functional* and *positional* levels in Garrett's theory (1975, 1988), to *functional* and *constituent integration* in Bock (1978a), and to the top two levels of the lexical network with their associated tactic frames in Dell's (1986) model. Collectively, Levelt (1989) terms the processes *grammatical encoding*.
4. A genuine error of this sort occurred in a request for "one of those celery things" when the speaker meant to say "one of those licorice things."
5. These syntactic roles correspond roughly to the traditional roles of subject, direct object, and indirect object.
6. This conflates the constituent assembly and inflection processes discussed in Bock and Levelt (1994).
7. Because Blackmer and Mitton (1991) included filled pauses and repetitions in their repair classifications, not all of these interruptions represented actual corrections. Only a small minority were repairs of speech errors (3% of the sample), suggesting that many of the dysfluencies may have been something else (e.g., planning delays).
8. An editor that operates only on the final output is equivalent to the monitor discussed above.
9. The coordination problem is a species of the *binding* problem that afflicts connectionist approaches to language (J. A. Fodor & Pylyshyn, 1988).
10. DeSmedt (1990) and Kempen and Hoenkamp (1987) have developed computational models that satisfy these requirements, but as yet there have been few tests of their psycholinguistic implications.
11. The fact that this choice comes in two allomorphic variants (*a* and *an*) is not relevant here, since the variants are phonologically conditioned.

References

- Ashcraft, M. H. (1993). A personal case history of transient anomia. *Brain and Language*, 44, 47–57.
- Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, 14, 382–391.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.
- Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 575–589.
- Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, 24, 89–106.
- Balota, D. A., & Chumbley, J. I. (1990). Where are the effects of frequency in visual word recognition tasks? Right where we said they were! Comment on Monsell, Doyle, and Haggard (1989). *Journal of Experimental Psychology: General*, 119, 231–237.
- Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–73). Cambridge, UK: Cambridge University Press.
- Beattie, G. (1983). *Talk: An analysis of speech and non-verbal behaviour in conversation*. Milton Keynes, England: Open University Press.
- Bierwisch, M., & Schreuder, R. (1992). From concepts to lexical items. *Cognition*, 42, 23–60.
- Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition*, 39, 173–194.
- Bock, J. K. (1977). The effect of a pragmatic presupposition on syntactic structure in question answering. *Journal of Verbal Learning and Verbal Behavior*, 16, 723–734.
- Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89, 1–47.
- Bock, J. K. (1986a). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 575–586.
- Bock, J. K. (1986b). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355–387.
- Bock, J. K. (1987a). Coordinating words and syntax in speech plans. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 3, pp. 337–390). London: Erlbaum.
- Bock, J. K. (1987b). An effect of the accessibility of word forms on sentence structures. *Journal of Memory and Language*, 26, 119–137.
- Bock, J. K. (1989). Closed-class immanence in sentence production. *Cognition*, 31, 163–186.
- Bock, J. K. (1990). Structure in language: Creating form in talk. *American Psychologist*, 45, 1221–1236.
- Bock, J. K. (1995). *Language production: The methods and their theories*. Manuscript submitted for publication.
- Bock, J. K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, 31, 99–127.
- Bock, J. K., & Eberhard, K. M. (1993). Meaning, sound, and syntax in English number agreement. *Language and Cognitive Processes*, 8, 57–99.
- Bock, J. K., & Irwin, D. E. (1980). Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior*, 19, 467–484.
- Bock, J. K., & Kroch, A. S. (1989). The isolability of syntactic processing. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 157–196). Dordrecht: Kluwer.
- Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA: Academic Press.
- Bock, J. K., & Loebell, H. (1990). Framing sentences. *Cognition*, 35, 1–39.
- Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, 99, 150–171.
- Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, 23, 45–93.
- Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21, 47–67.

- Bolinger, D. L. (1961). Verbal evocation. *Lingua*, 10, 113–127.
- Boomer, D. S. (1965). Hesitation and grammatical encoding. *Language & Speech*, 8, 148–158.
- Bresnan, J., & Kaplan, R. M. (1984). Grammars as mental representations of language. In W. Kintsch, J. R. Miller, & P. G. Polson (Eds.), *Method and tactics in cognitive science* (pp. 103–135). Hillsdale, NJ: Erlbaum.
- Brown, J. S., McDonald, J. L., Brown, T. L., & Carr, T. H. (1988). Adapting to processing demands in discourse production: The case of handwriting. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 45–59.
- Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology*, 19, 441–472.
- Brown, R., & McNeill, D. (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 5, 325–337.
- Butterworth, B. (1980). Evidence from pauses in speech. In B. Butterworth (Ed.), *Language production* (Vol. 1, pp. 155–176). London: Academic Press.
- Butterworth, B. (1989). Lexical access in speech production. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 108–135). Cambridge, MA: MIT Press.
- Butterworth, B., & Hadar, U. (1989). Gesture, speech, and computational stages: A reply to McNeill. *Psychological Review*, 96, 168–174.
- Cheung, H., & Kemper, S. (1993). Recall and articulation of English and Chinese words by Chinese-English bilinguals. *Memory & Cognition*, 21, 666–670.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Clark, E. V., & Clark, H. H. (1978). Universals, relativity, and language processing. In J. H. Greenberg (Ed.), *Universals of human language: Vol. 1. Method and theory* (pp. 225–277). Stanford, CA: Stanford University Press.
- Clark, H. H. & Clark, E. V. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.
- Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech*. Cambridge, MA: Harvard University Press.
- Cooper, W. E., Soares, C., & Reagan, R. T. (1985). Planning speech: A picture’s word’s worth. *Acta Psychologica*, 58, 107–114.
- Cooper, W. E., & Sorensen, J. M. (1981). *Fundamental frequency in sentence production*. New York: Springer-Verlag.
- Cowan, N. (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory and Language*, 31, 668–684.
- Cutler, A. (1980a). Errors of stress and intonation. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 67–80). New York: Academic Press.
- Cutler, A. (1980b). Syllable omission errors and isochrony. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 183–190). The Hague: Mouton.
- Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, 2, 133–142.
- Cutler, A., & Isard, S. D. (1980). The production of prosody. In B. Butterworth (Ed.), *Language production* (Vol. 1, pp. 245–269). London: Academic Press.
- Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, 25, 1–18.
- Deese, J. (1980). Pauses, prosody, and the demands of production in language. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 69–84). The Hague: Mouton.
- Deese, J. (1984). *Thought into speech: The psychology of a language*. Englewood Cliffs, NJ: Prentice-Hall.
- Dell, G. S. (1984). Representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 222–233.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.
- Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, 5, 313–349.
- Dell, G. S., & Brown, P. M. (1991). Mechanisms for listener-adaptation in language production: Limiting the role of the “model of the listener.” In D. J. Napoli & J. A. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore festschrift for Lila Gleitman* (pp. 105–129). Hillsdale, NJ: Erlbaum.
- Dell, G. S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science*, 17, 149–195.
- Dell, G. S., & O’Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991b). *Psychological Review*, 98, 604–614.
- Dell, G. S., & O’Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, 42, 287–314.
- Dell, G. S., & O’Seaghdha, P. G. (1994). Inhibition in interactive activation models of linguistic selection and sequencing. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 409–453). San Diego, CA: Academic Press.
- Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, 20, 611–629.
- del Viso, S., Igoa, J. M., & García-Albea, J. E. (1991). On the autonomy of phonological encoding: Evidence from slips of the tongue in Spanish. *Journal of Psycholinguistic Research*, 20, 161–185.
- deSmedt, K. J. M. J. (1990). *Incremental sentence generation*. Unpublished doctoral dissertation, University of Nijmegen, The Netherlands.
- Ellis, A. W. (1980a). Errors in speech and short-term memory: The effects of phonemic similarity and syllable position. *Journal of Verbal Learning and Verbal Behavior*, 19, 624–634.
- Ellis, A. W. (1980b). Speech production and short-term memory. In J. Morton & J. C. Marshall (Eds.), *Psycholinguistics: 2. Structures and processes* (pp. 157–187). Cambridge, MA: MIT Press.
- Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology*, 71, 43–51.
- Fay, D., & Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry*, 8, 505–520.
- Fayol, M., Largy, P., & Lemaire, P. (1994). When cognitive overload enhances subject-verb agreement errors: A study of French written language. *Quarterly Journal of Experimental Psychology*, 47A, 437–464.
- Fenk-Oczlon, G. (1989). Word frequency and word order in freezes. *Linguistics*, 27, 517–556.
- Ferreira, F. (1993). The creation of prosody during sentence production. *Psychological Review*, 100, 233–253.
- Ferreira, V. (1994). *Is it better to give than donate? The consequences of syntactic flexibility in language production*. Manuscript submitted for publication.
- Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language*. New York: McGraw-Hill.
- Fodor, J. A., & Garrett, M. (1966). Some reflections on competence and performance. In J. Lyons & R. J. Wales (Eds.), *Psycholinguistics papers* (pp. 135–154). Edinburgh: Edinburgh University Press.

- Fodor, J. A., Garrett, M. F., Walker, E. C. T., & Parkes, C. H. (1980). Against definitions. *Cognition*, 8, 263-367.
- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3-71.
- Fodor, J. D. (1981). Does performance shape competence? *Philosophical Transactions of the Royal Society of London, Series B*, 295, 285-295.
- Fodor, J. D. (1989). Empty categories in sentence processing. *Language and Cognitive Processes*, 4, 155-209.
- Fodor, J. D., Fodor, J. A., & Garrett, M. F. (1975). The psychological unreality of semantic representations. *Linguistic Inquiry*, 6, 515-531.
- 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* (pp. 797-827). Cambridge, MA: MIT Press.
- Ford, M., & Holmes, V. M. (1978). Planning units and syntax in sentence production. *Cognition*, 6, 35-53.
- Frazier, L. (1982). Shared components of production and perception. In D. Caplan & J. C. Marshall (Eds.), *Neural models of language processes* (pp. 225-236). New York: Academic Press.
- Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559-586). Hillsdale, NJ: Erlbaum.
- Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 133-177). New York: Academic Press.
- Garrett, M. F. (1976). Syntactic processes in sentence production. In R. J. Wales & E. C. T. Walker (Eds.), *New approaches to language mechanisms* (pp. 231-256). Amsterdam: North-Holland.
- Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (pp. 177-220). London: Academic Press.
- Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19-76). London: Academic Press.
- Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey: Vol. 3. Language: Psychological and biological aspects* (pp. 69-96). Cambridge, UK: Cambridge University Press.
- Garrett, M. F. (1992). Disorders of lexical selection. *Cognition*, 42, 143-180.
- Gerken, L. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language*, 30, 431-451.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. London: Academic Press.
- Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language & Speech*, 15, 103-113.
- Harley, T. A. (1990). Environmental contamination of normal speech. *Applied Psycholinguistics*, 11, 45-72.
- Harley, T. A. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Language and Cognitive Processes*, 8, 291-309.
- Hawkins, J. A. (1990). A parsing theory of word order universals. *Linguistic Inquiry*, 21, 223-261.
- Heeschen, C. (in press). Morphosyntactic characteristics of spoken language. In G. Blanken, J. Dittman, H. Grim, J. Marshall, & C. Wallesch (Eds.), *Linguistic disorders and pathologies*. Berlin: de Gruyter.
- Holmes, V. M. (1988). Hesitations and sentence planning. *Language and Cognitive Processes*, 3, 323-361.
- Hoosain, R. (1982). Correlation between pronunciation speed and digit span size. *Perceptual and Motor Skills*, 55, 1128.
- Hotopf, W. H. N. (1980). Semantic similarity as a factor in whole-word slips of the tongue. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 97-109). New York: Academic Press.
- Huitema, J. S. (1993). *Planning referential expressions in speech production*. Unpublished doctoral dissertation, University of Massachusetts, Amherst.
- Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology*, 38, 241-253.
- Huttenlocher, J., & Kubicek, L. F. (1983). The source of relatedness effects on naming latency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 486-496.
- James, W. (1890). *The principles of psychology* (Vol. 1). New York: Dover.
- Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20, 824-843.
- Kelly, M. H. (1986). *On the selection of linguistic options*. Unpublished doctoral dissertation, Cornell University, Ithaca, NY.
- Kelly, M. H. (1988). Rhythmic alternation and lexical stress differences in English. *Cognition*, 30, 107-137.
- Kelly, M. H., & Bock, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 389-403.
- Kelly, M. H., Bock, J. K., & Keil, F. C. (1986). Prototypicality in a linguistic context: Effects on sentence structure. *Journal of Memory and Language*, 25, 59-74.
- Kelly, M. H., & Rubin, D. J. (1988). Natural rhythmic patterns in English verse: Evidence from child counting-out rhymes. *Journal of Memory and Language*, 27, 718-740.
- Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11, 201-258.
- Keysar, B. (1994). The illusory transparency of intention: Linguistic perspective taking in text. *Cognitive Psychology*, 26, 165-208.
- Kynette, D., Kemper, S., Norman, S., & Cheung, H. (1990). Adults' word recall and word repetition. *Experimental Aging Research*, 16, 117-121.
- Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior*, 12, 119-131.
- Lapointe, S. G., & Dell, G. S. (1989). A synthesis of some recent work in sentence production. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 107-156). Dordrecht: Kluwer.
- Levelt, W. J. M. (1981). The speaker's linearization problem. *Philosophical Transactions of the Royal Society of London, Series B*, 295, 305-315.
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14, 41-104.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, 42, 1-22.
- Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991a). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, 98, 122-142.
- Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991b). Normal and deviant lexical processing: Reply to Dell and O'Seaghdha (1991). *Psychological Review*, 98, 615-618.
- Levitt, A. G., & Healy, A. F. (1985). The roles of phoneme frequency, similarity, and availability

- ty in the experimental elicitation of speech errors. *Journal of Memory and Language*, 24, 717-733.
- Lindsay, J. R. (1975). Producing simple utterances: How far ahead do we plan? *Cognitive Psychology*, 7, 1-19.
- Lounsbury, F. G. (1965). Transitional probability, linguistic structure, and systems of habit-family hierarchies. In C. E. Osgood & T. A. Sebeok (Eds.), *Psycholinguistics: A survey of theory and research problems* (pp. 93-101). Bloomington: Indiana University Press.
- Lucy, J. A. (1992a). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge, UK: Cambridge University Press.
- Lucy, J. A. (1992b). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. Cambridge, UK: Cambridge University Press.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (in press). The lexical nature of syntactic ambiguity resolution. *Psychological Review*.
- MacKay, D. G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology*, 3, 210-227.
- MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer-Verlag.
- MacKay, D. G., Wulf, G., Yin, C., & Abrams, L. (1993). Relations between word perception and production: New theory and data on the verbal transformation effect. *Journal of Memory and Language*, 32, 624-646.
- Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word*, 15, 19-44.
- Mahl, G. F. (1987). *Explorations in nonverbal and vocal behavior*. Hillsdale, NJ: Erlbaum.
- Marr, D. (1982). *Vision*. New York: Freeman.
- Martin, N., Weisberg, R. W., & Saffran, E. M. (1989). Variables influencing the occurrence of naming errors: Implications for models of lexical retrieval. *Journal of Memory and Language*, 28, 462-485.
- McDonald, J. L., Bock, K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology*, 25, 188-230.
- McNamara, T. P., & Miller, D. L. (1989). Attributes of theories of meaning. *Psychological Bulletin*, 106, 355-376.
- McNeill, D. (1987). *Psycholinguistics: A new approach*. New York: Harper & Row.
- McNeill, D. (1989). A straight path—to where? Reply to Butterworth and Hadar (1989). *Psychological Review*, 96, 175-179.
- Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 1146-1160.
- Monsell, S. (1990). Frequency effects in lexical tasks: Reply to Balota and Chumbley (1990). *Journal of Experimental Psychology: General*, 119, 335-339.
- Monsell, S. (in press). The nature and locus of word frequency effects in reading. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition*. Hillsdale, NJ: Erlbaum.
- Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, 118, 43-71.
- Murphy, G. L. (1991). Meaning and concepts. In P. Schwanenflugel (Ed.), *The psychology of word meaning* (pp. 11-35). Hillsdale, NJ: Erlbaum.
- Newmeyer, F. J. (1992). Iconicity and generative grammar. *Language*, 68, 756-796.
- 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: Slips of the tongue, ear, pen, and hand* (pp. 887-95). New York: Academic Press.
- Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 17, 273-281.
- Peterson, R. R., Shim, M., & Savoy, P. (1993, November). *Time course of lexicalization in language production*. Paper presented at the meeting of the Psychonomic Society, Washington, DC.
- Pike, K. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Pinker, S. (1991). Rules of language. *Science*, 253, 530-535.
- Plunkett, K. (1993). Lexical segmentation and vocabulary growth in early language acquisition. *Journal of Child Language*, 20, 43-60.
- Power, M. J. (1985). Sentence production and working memory. *Quarterly Journal of Experimental Psychology*, 37A, 367-385.
- Power, M. J. (1986). A technique for measuring processing load during speech production. *Journal of Psycholinguistic Research*, 15, 371-382.
- Reynolds, A., & Paivio, A. (1968). Cognitive and emotional determinants of speech. *Canadian Journal of Psychology*, 22, 164-175.
- Roelofs, A. (1992). A spreading activation theory of lemma retrieval in speaking. *Cognition*, 42, 107-142.
- Roelofs, A. (1993). Testing a non-decompositional theory of lemma retrieval in speaking: Retrieval of verbs. *Cognition*, 47, 59-87.
- Savage, G. R., Bradley, D. C., & Forster, K. I. (1990). Word frequency and the pronunciation task: The contribution of articulatory fluency. *Language and Cognitive Processes*, 5, 203-236.
- Schachter, S., Christenfeld, N., Ravina, B., & Bilous, F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology*, 60, 362-367.
- Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition*, 47, 1-24.
- Schriefers, H. (1992). Lexical access in the production of noun phrases. *Cognition*, 45, 33-54.
- Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 841-850.
- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86-102.
- Schweickert, R., & Boruff, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 419-425.
- Schweickert, R., Guentert, L., & Hersberger, L. (1990). Phonological similarity, pronunciation rate, and memory span. *Psychological Science*, 1, 74-77.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523-568.
- Sevold, C. A., & Dell, G. S. (1994). The sequential cuing effect in speech production. *Cognition*, 53, 91-127.
- Sevold, C. A., Dell, G. S., & Cole, J. (1993, November). *Syllable structure in speech production*. Paper presented at the meeting of the Psychonomic Society, Washington, DC.
- Shallice, T., McLeod, P., & Lewis, K. (1985). Isolating cognitive modules with the dual-task paradigm: Are speech perception and production separate processes? *Quarterly Journal of Experimental Psychology*, 37A, 507-532.
- Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial-ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 295-342). Hillsdale, NJ: Erlbaum.
- Slobin, D. I. (in press). From "thought and language" to "thinking for speaking." In J. Gumperz

- & S. C. Levinson (Eds.), *Rethinking linguistic relativity*. Cambridge, UK: Cambridge University Press.
- Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language*, 32, 25–38.
- Steedman, M. (1991). Structure and intonation. *Language*, 67, 260–296.
- Stemberger, J. P. (1984). Structural errors in normal and agrammatic speech. *Cognitive Neuropsychology*, 1, 218–313.
- Stemberger, J. P. (1985). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the psychology of language* (pp. 143–186). London: Erlbaum.
- Stemberger, J. P. (1989). Speech errors in early child language production. *Journal of Memory and Language*, 28, 164–188.
- Stemberger, J. P., & MacWhinney, B. (1986a). Form-oriented inflectional errors in language processing. *Cognitive Psychology*, 18, 329–354.
- Stemberger, J. P., & MacWhinney, B. (1986b). Frequency and the lexical storage of regularly inflected forms. *Memory & Cognition*, 14, 17–26.
- Sternberg, S., Knoll, R. L., Monsell, S., & Wright, C. E. (1988). Motor programs and hierarchical organization in the control of rapid speech. *Phonetica*, 45, 175–197.
- Stigler, J. W., Lee, S.-Y., & Stevenson, H. W. (1986). Digit memory in Chinese and English: Evidence for a temporally limited store. *Cognition*, 23, 1–20.
- Titchener, E. B. (1909). *Lectures on the experimental psychology of the thought-processes*. New York: Macmillan.
- van Wijk, C., & Kempen, G. (1987). A dual system for producing self-repairs in spontaneous speech: Evidence from experimentally elicited corrections. *Cognitive Psychology*, 19, 403–440.
- Wheeldon, L. R., & Monsell, S. (1994). Inhibition of spoken word production by priming a semantic competitor. *Journal of Memory and Language*, 33, 332–356.
- Wijnen, F. (1990). The development of sentence planning. *Journal of Child Language*, 17, 651–675.
- Wijnen, F. (1992). Incidental word and sound errors in young speakers. *Journal of Memory and Language*, 31, 734–755.
- Wright, C. E. (1979). Duration differences between rare and common words and their implications for the interpretation of word frequency effects. *Memory & Cognition*, 7, 411–419.
- Zhang, G., & Simon, H. A. (1985). STM capacity for Chinese words and idioms: Chunking and the acoustical loop hypothesis. *Memory & Cognition*, 13, 193–201.