Language processing

edited by

Simon Garrod and Martin J. Pickering
University of Glasgow, UK

Psychology Press
a member of the Taylor & Francis group

CHAPTER TWELVE

Language production

Kathryn Bock and John Huitema
Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana, USA

Language production—talking—is a facet of language performance. Its special properties may be set in relief against the backdrop of Noam Chomsky's (1965, p. 3) famous definition of the subject matter of linguistics:

> Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance.

By contrast, psycholinguistic theory 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. Regarding language production in particular, the goal of psycholinguistics is to explain how real speakers in real time retrieve and assemble elements of language from long-term memory in order to communicate their ideas.

The chief issues in language production centre 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 the system interrelates linguistic and non-linguistic

knowledge, and how the system is organised within and constrained by human cognitive capacities. In this chapter we will survey the kinds of phenomena that serve as focal points for research on production, present an overview of the cognitive processes that take place in the course of creating an utterance, and summarise some of the psycholinguistic findings that illuminate the workings of these processes. At the end we will consider how language production fits into the broader framework of psycholinguistic research.

PHENOMENA OF LANGUAGE PRODUCTION

The facts that a theory of language production should explain are not immediately obvious, because—intuitively—talking isn't hard. In a lecture delivered at the University of Illinois in 1909, a famous founder of American psychology claimed to be able to "read off what I have to say from a memory manuscript" (Titchener, 1909, p. 8). This caricatures one's usual experience when speaking, but perhaps not by much. Talking seems just too easy to pose any problems worth explaining.

Because of this, the challenges of production are more readily appreciated in terms of talk's typical failures. The failures range widely. One sort is illustrated by a psychology professor's experience during a transient neurological episode. During the attack, the professor was able to form perfectly coherent messages, but could not express them (Ashcraft, 1993, pp. 49, 54):

> 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.

More than a century earlier, William James (1890, pp. 251–252) described another type of failure during speech, the common 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.

These introspections allude to gaps in the process of putting ideas into words, but gaps of different kinds. Ashcraft experienced an unbridgeable gap between the thought he wanted to convey and the cognitive processes that normally create the linguistic form to express that thought. The tip-of-the-tongue state that James described reveals a gap between a single concept and the word that expresses it.

In addition to such problems of omission, there are problems of commission. Table 12.1 presents a selection, drawn from our own observations, of the many kinds of speech errors that have been studied in research on language production. The first two examples are anticipations—saying too early a word or a sound that is supposed to come later in the utterance. The next two are perseverations—repeating a word or a sound from earlier in the utterance. Sometimes two linguistic elements exchange places in an utterance, as illustrated in examples 5–8. Still

TABLE 12.1
Sample Speech Errors

Type of Error	*Intended Utterance*	*Error*
1. Word anticipation	bury me right with him	bury him right with him
2. Sound anticipation	the lush list	the lust list
3. Word perseveration	evidence brought to bear on representational theories	evidence brought to bear on representational evidence
4. Sound perseveration	President Bush's budget	President Bush's boodget
5. Word exchange	the head of a pin	the pin of a head
6. Sound exchange	occipital activity	accipital octivity
7. Stranding exchange	the dome doesn't have any windows	the window doesn't have any domes
8. Phrase exchange	the death of his son from leukaemia	the death of leukaemia from his son
9. Semantically related word-substitution	I like berries with my cereal	I like berries with my fruit
10. Phonologically related word-substitution	part of a community	part of a committee
11. Sound substitution	the disparity	the disparigy
12. Word blend	it really stood/stuck out	it really stook out
13. Phrase blend	at large/on the loose	at the loose

another type of error, called a substitution, is to say a word or sound other than the one that was intended, as shown in examples 9–11. The final sort of error illustrated in the table is when a speaker blends two words or phrases together, as in the last two examples.

The most common problem in production is disfluency: Speakers commit false starts, they pause silently or noisily (saying "uh" or "um") in the course of an utterance, and they retrace their verbal steps. They do these things very often. Considering only filled pauses ("uh", "um", and "er"), the average rate in the lectures of 45 professors from 10 different disciplines has been clocked at one pause every 18 seconds (Schachter, Christenfeld, Ravina, & Bilous, 1991). Some of these disfluencies reflect simple indecision about what to say next, but others stem from momentary disruptions in specific language processes, such as retrieving a particular word or constructing an expression.

Speech errors and disfluencies both provide clues about the nature of information processing in language production. A schematic view of what they suggest about the processes and their organisation is shown in Fig. 12.1. In the next section we will explain how the components of the figure relate to the errors that people commit when they talk.

THE COGNITIVE COMPONENTS OF LANGUAGE PRODUCTION

Production begins with an intention to communicate an idea. The idea is called a *message*, and we assume that it is a thought, largely unadorned by the trappings of language. At this level, the messages of a French speaker may not be much different from those of an English speaker, or a Japanese speaker, or a speaker of any language. Returning to Ashcraft's description of his transient neurological episode, it appears that he was able to formulate messages but was momentarily unable to express them in language.

In terms of the model in Fig. 12.1, what was missing was the ability to carry out the cognitive work that is needed for finding words and putting them together (grammatical encoding) or for finding sounds and putting them together (phonological encoding). Notice in the figure that these processes are separated into two different components. There are several reasons for this separation. We will briefly survey four of them.

First, consider the frequency with which different kinds of elements are involved in errors, shown in Fig. 12.2. Two kinds of units stand out. Among the meaningful units, words are more frequently implicated than any others, and among the sound units, phonemes (single sounds) are by far the most frequently involved. This would not be surprising if words and phonemes were the most common units in speech overall, but they

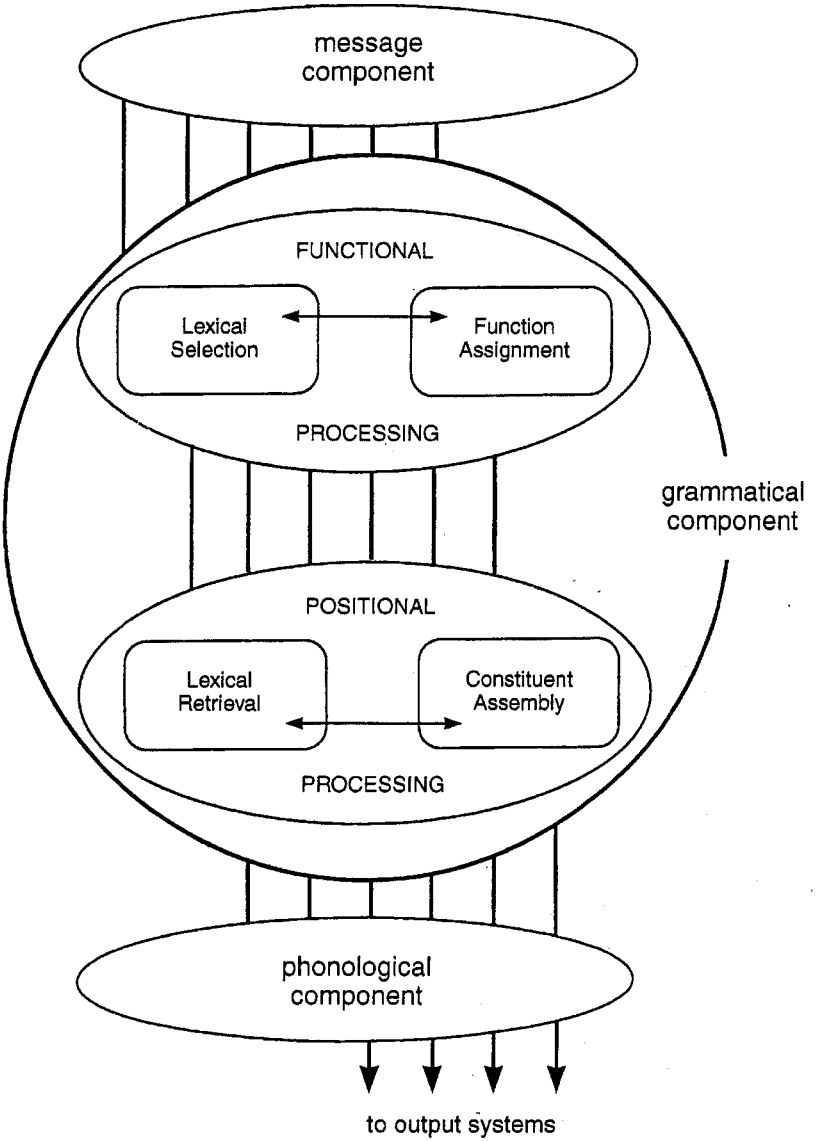

FIG. 12.1 The organisation of processing components in normal language production (from Bock, 1995). Copyright © (1995) Academic Press. Reprinted with permission.

are not: Words are less common than morphemes (units of meaning like *un-* and *-happy* in the word *unhappy*), and phonemes are less common than features (articulatory components of sounds, such as voicing). The implication is that there is a set of processes that deal mainly with finding

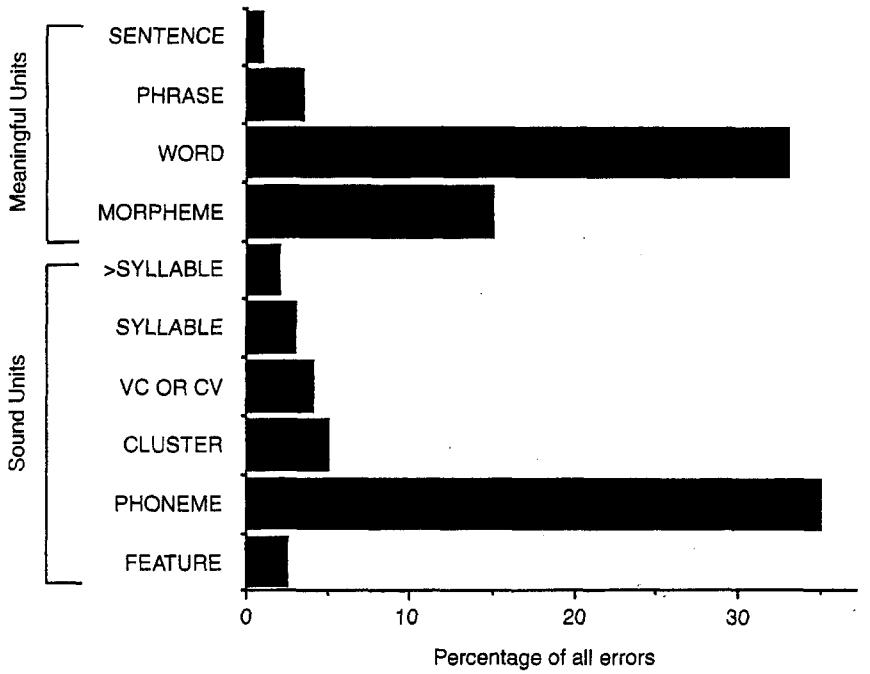

FIG. 12.2 The frequency of different types of linguistic units in exchange errors (from Dell, 1995). Copyright © (1995) MIT Press. Reprinted with permission.

and arranging words (grammatical encoding) and a set of processes that deal mainly with finding and arranging phoneme segments (phonological encoding).

A second reason for separating the grammatical and phonological components can be seen in another feature of exchange errors. The words in exchanges usually represent the same grammatical category: Nouns exchange with nouns, verbs with verbs, and so on. By contrast, when a phoneme exchanges with another phoneme, there is no obvious grammatical similarity between the words in which the exchanging phonemes originated. So, counter to what one would expect if syntactic categories constrained all production processes, the syntactic categories of the containing words are irrelevant to sound exchanges. However, the exchanging sounds themselves tend to come from similar phonological categories: Consonants exchange with other consonants, whereas vowels exchange with other vowels. The implication again is that one component of the production system attends to the syntactic category of words in order to arrange them grammatically, while another component attends to the sounds of words and is oblivious to their syntactic functions.

A third piece of evidence for the separation of grammatical and phonological processes comes from yet another property of word and sound exchanges. Words that exchange are typically separated by a phrase or two, whereas sounds that exchange usually come from adjacent words in the same phrase (Garrett, 1980a). This observation suggests that grammatical processes and phonological processes differ in the range over which they operate, with grammatical encoding having a longer view of the eventual utterance than phonological processes.

The model's division into grammatical and phonological processes can also help to explain the features of certain complex errors, such as "The skreaky gwease gets the wheel". The speaker intended to say "The squeaky wheel gets the grease", implying that someone who whines and complains is more likely to get attention than someone who suffers in silence. According to the model in Fig. 12.1, the erroneous utterance occurred because of two distinct disruptions. First, in arranging the words, the grammatical component misplaced the nouns *wheel* and *grease*, setting the stage for the utterance "The squeaky grease gets the wheel". But then something else went wrong. Since *squeaky* and *grease* became next-door neighbours as a result of the word-exchange error, they were close enough for their sounds to exchange, and two of them did. So, while ordering the words' phonemes, the /r/ and /w/ sounds exchanged, leading to "skreaky gwease".

GRAMMATICAL ENCODING

Now we will look inside the grammatical component at the processes that retrieve and arrange words. There are two sets of operations, divided into functional processing and positional processing. In describing these operations, we will make use of an analogy to a mental dictionary with entries arranged like the one in the left panel of Fig. 12.3, and a mental sentence skeleton constructed like the one in the right panel. Notice that the lexical entry works like one in a reverse dictionary, or Roget's (1852) original thesaurus, a dictionary in which entries must be consulted according to their meanings rather than their letters or sounds.

Functional processing

Functional processing is concerned with selecting words from the mental lexicon (lexical selection) and assigning syntactic functions to them (function assignment). Lexical selection can be likened to locating an entry with the right meaning in the mental reverse dictionary, prior to finding a pronunciation for it (unlike a real dictionary entry). Function assignment is deciding which message element is going to be the grammatical subject, which the direct object, and so on.

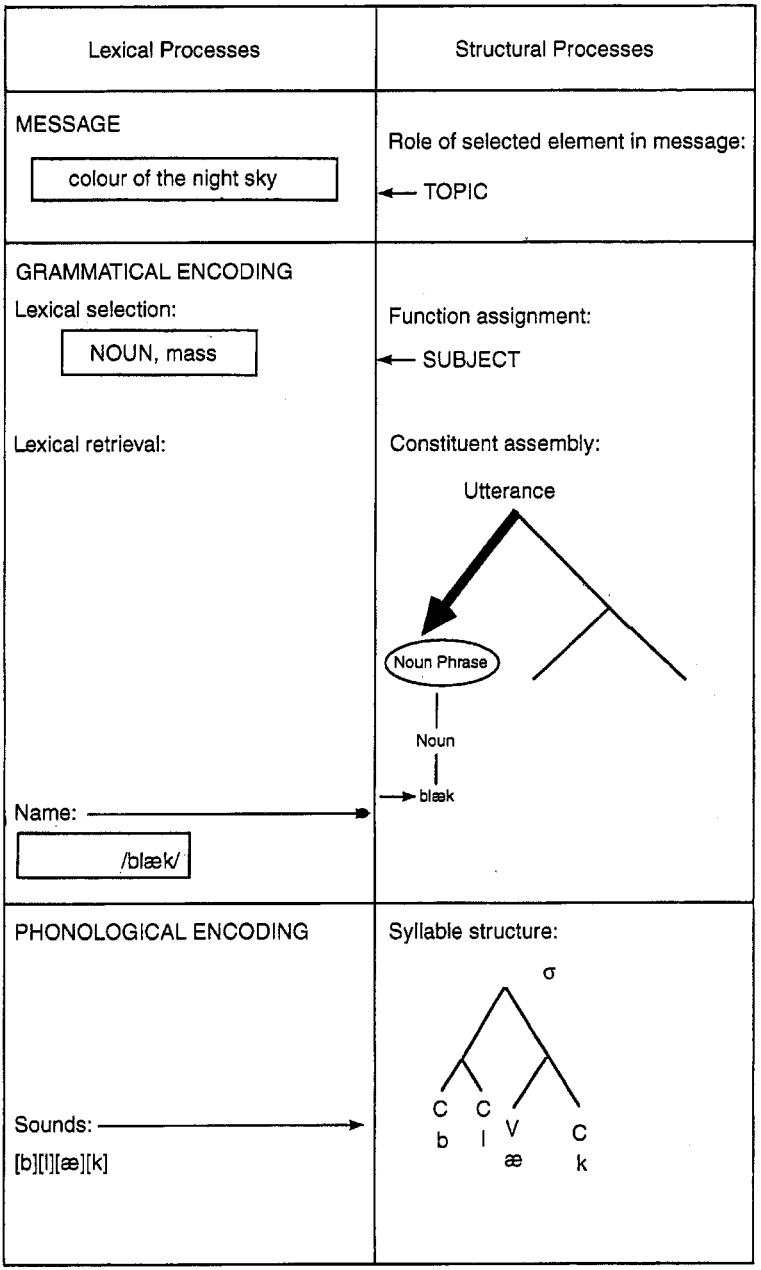

FIG. 12.3 Retrieving a word (*black*) for production in the utterance *Black is my favourite colour.*

The operation of lexical selection is discernible in the tip-of-the-tongue (TOT) state. This is the annoying condition that William James described, in which one is quite sure of knowing a word to express a particular meaning, while being unable to retrieve the word's sounds. Some evidence that speakers in this state have in fact selected a particular word, rather than just a concept, comes from a speaker of Italian who suffered brain damage, making it very difficult for him to name pictures of everyday objects. Even when he could only guess at what the first or last sound of the name was, he was almost perfect at choosing the appropriate masculine or feminine article (Badecker, Miozzo, & Zanuttini, 1995). The concepts expressed by these words were not inherently male or female: For example, the Italian word for dessert, *dolce*, is masculine, and the word for hand, *mano*, is feminine. The concept alone is not enough to determine the gender of the word. Therefore, the speaker must have been using information about the word itself to make his judgement, showing that he had indeed selected a particular word, even though he was not able to retrieve the sounds of that word.

Lexical selection can go astray in a different way, as in the error "I like berries with my fruit" (produced instead of "I like berries with my cereal", see in Table 12.1). In such semantic substitutions, an incorrect but related word is selected to express a concept. The problem is analogous to mistakenly picking an entry next to the intended word in the reverse dictionary. The result is a substitution that is similar in meaning to the intended word but not usually similar in sound.

The other component of functional processing is function assignment. Function assignment determines the syntactic role that message elements will play in an utterance. Errors involving the exchange of pronouns are particularly informative about the process. This is because pronouns in English overtly mark syntactic functions like subject (nominative case), object (objective case), possessive (genitive case), and so on, as seen in the respective forms of the masculine singular pronoun in "*He* liked *him* and *his* family". An illustrative mistake was reported by Garrett (1980b). A speaker intended to say "She offends his sense of how the world should be", but what came out was "He offends her sense of how the world should be". This is an exchange of pronouns, but the thing to notice is that it is not a simple exchange of the pronoun forms: The error is not "His offends she sense of how the world should be". What happened? Apparently, the syntactic function of subject was erroneously assigned to the masculine player in the event, while the possessive function was assigned to the feminine player. In other words, something went wrong during function assignment, causing the message elements to be mapped to the wrong syntactic roles.

The product of functional processing is a representation that indicates for each message element its syntactic role and the words to be used for expressing it. Two important properties of utterances are yet to emerge: The actual order of the phrases and specifications for the sounds of words. These are part of positional processing.

Positional processing

Like functional processing, positional processing involves both a syntactic and a lexical sub-component. The syntactic sub-component is called *constituent assembly*. It puts phrases, words, and grammatical inflections in order, arranging them in accordance with the grammatical patterns of the language. One error feature that is associated with constituent assembly is termed *stranding* (see Table 12.1), which reliably accompanies any exchange of inflected words. The occurrence of stranding implies that the stems and affixes of words are positioned separately during processing, even though they eventually surface together in speech. This 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". Despite the exchange of the word stems *dome* and *window*, the plural *-s* suffix stayed put: It was stranded in the direct object position and affixed itself to *dome*. The error suggests that function assignment was correctly carried out: The intended subject was singular and the subject remains singular in the error, though the subject noun is the wrong one. Similarly, the intended direct object was plural and the direct object remains plural in the error, though the noun is wrong. Evidently, the problem arose when the processes of constituent assembly positioned the retrieved word forms, putting them into phrase slots like the one shown on the right in Fig. 12.3.

The lexical sub-component of positional processing, lexical retrieval, is concerned with retrieving abstract word forms, like the one shown on the left in Fig. 12.3. More precisely, the outcome of lexical retrieval is a description of a word's morphology to be filled out in more detail during phonological encoding. Continuing the analogy to the reverse-dictionary entry, lexical retrieval involves finding the part of the entry that indicates the word's structure and alphabetic spelling (which is likewise a description of a word's sounds, albeit a different sort of description than one envisions for the mental lexicon). One type of error that may be attributable to disruptions in lexical retrieval is the phonological word-substitution, sometimes termed a malapropism. The substituted word sounds similar to the intended word (e.g. "committee" instead of "community"; see Table 12.1) but need not be related to it in meaning. This indicates that similar-sounding word forms can interfere with one another during

retrieval, despite being different in meaning, and can be explained as another consequence of the separation of lexical retrieval from lexical selection.

PHONOLOGICAL ENCODING

Whereas grammatical encoding manipulates words or morphemes as wholes, phonological encoding manipulates the components of words, the speech sounds. It is responsible for putting phonemes and syllables in order, within representations that carry the rhythmic and intonational qualities of the language. It determines how the individual sounds of a word are pronounced, how they should be ordered, where syllable boundaries are, and so on. At the bottom of Fig. 12.3, the sounds of the word *black* are spelled out in preparation for assignment to slots in a syllabic frame that helps to control the articulation of speech.

Failures of phonological encoding are revealed in such errors as sound exchanges, perseverations, and anticipations (see Table 12.1). These errors show that phonological encoding is, like grammatical encoding, tightly constrained. Phoneme exchanges almost always involve sounds from the same class (consonant or vowel; MacKay, 1970). Elements are more likely to exchange when the sounds that precede or follow them are phonetically similar (Garrett, 1975). There is a positional constraint, as well, such that errors tend to implicate the same parts of different syllables. For example, in the error "lust list" (in Table 12.1), the cluster of syllable-final consonants (or *coda*) in the intended word "lush" was replaced by the coda of the following word, "list". Errors in which different parts of successive syllables interact (e.g. "stush list") are rare.

The product of phonological encoding serves as input to the articulatory processes that actually give voice to the utterance. Although not the focus of this chapter, articulation is itself a complex skill: Producing the roughly 15 sounds per second that make up fluent speech requires the rapid co-ordination of more groups of muscles than are involved in any other mechanical performance of the human body (Fink, 1986).

AN ALTERNATIVE PERSPECTIVE ON SPEECH ERRORS?

The information-processing approach exemplified by the model in Fig. 12.1 stands in stark contrast to a more famous theory of the origin of speech errors, that of Sigmund Freud. Freud formulated his views in reaction to the work of an Austrian linguist of the late 1800s named Rudolf Meringer (Meringer & Mayer, 1895/1978). Meringer recorded the speech errors he heard, and noticed that errors often involved linguistic

elements that are similar to one another. He suggested that errors might come about because of transient changes in the memory strength of words or sounds that cause them to appear in places where they do not belong. This is reminiscent of claims in some contemporary production theories (e.g. Dell, 1986), close relatives of the model we have sketched in this chapter.

Freud, in contrast, speculated that the true sources of errors lie beyond the mundane linguistic similarities that Meringer emphasised. As almost everyone knows, Freud proposed that errors were the result of unconscious intentions that coloured or played havoc with the consciously intended message. He offered the instance of a professor who said "In the case of the female genital, in spite of the tempting ... I mean, the attempted..." (Freud, 1924/1935, p. 38). Another of Freud's examples was analysed as originating in the underlying ill-feeling of a worker toward his boss. The worker called on his colleagues before a meal to "burp [*aufzustossen*] to the health of our chief" instead of "drink a toast [*anzustossen*] to the health of our chief" (Freud, 1924/1935, p. 38).

Freud's idea that speech errors reveal unconscious motives was clearly more provocative than Meringer's. It was so provocative, in fact, that there was very little work on speech errors from a cognitive or linguistic perspective for many years after Freud produced his psychodynamic analysis.

Yet there are fundamental weaknesses in the Freudian account. One is that hardly any speech errors have a clear Freudian interpretation. Most errors are at best innocuous and at worst downright boring, as the examples in Table 12.1 testify. Freud may have been misled by the properties of the sample of speech errors he observed or, more likely, he overemphasised the most interesting of the mistakes that he encountered. Ellis (1980) showed that Freud's collection of errors is indeed unrepresentative of the distribution that arises in everyday speaking.

A second problem with the Freudian account is also apparent in Table 12.1. Almost all speech errors show a strong influence of purely linguistic factors. Consider the "toasting/burping" example again, in the original German form: "Ich fördere sie auf, auf das wohl unseres chefs aufzustossen." The substituted word *aufzustossen* is very similar to the intended word *anzustossen*, and the word *auf* appeared twice prior to its erroneous appearance, suggesting that it was very strongly primed. Freud considered this and dismissed it out of hand (Freud, 1924/1935, pp. 53–54), embracing as the only possibility that the production of a word or phrase entails the representation of a relevant meaning within what we have termed the speaker's message. The linguistic constraints on errors, their most prominent characteristics, have no satisfactory explanation in this framework.

None of this means that Freud was demonstrably wrong in his hypotheses about the causes of errors. None the less, his hypotheses explain very little about most mistakes in speech, overlooking or omitting most of the data. The study of speech errors therefore returned to Meringer's original interest in what errors can tell us about how people talk, instead of what they might reveal about the darker recesses of people's thought. Building from this work on errors, experimental research has begun to uncover a variety of facts about the kinds of processes that are involved in speaking.

EXPERIMENTAL RESEARCH ON LANGUAGE PRODUCTION

Although analyses of errors in spontaneous speech served as the starting point for the contemporary study of language production, there is much more for a theory to explain. Errors are rare events, particularly when we consider them in light of ordinary speech achievements. Measurements of normal speech rates give average values of about 150 words per minute (Maclay & Osgood, 1959) or 5.65 syllables per second (Deese, 1984). Although this speech is liberally sprinkled with pauses and false starts, outright error is very uncommon. For example, in a tape-recorded corpus of nearly 15,000 utterances, Deese (1984) counted only 77 syntactic anomalies—roughly one in every 195 utterances. Heeschen (1993) reported a similarly low incidence of syntactic errors in spoken German. Errors of lexical selection and retrieval (such as semantic and phonological word-substitutions) are even less common, with attested rates averaging under one per 1000 words (Bock & Levelt, 1994). Sound errors are rarest of all, occurring less than once in every 2000 words. The implication is that the most challenging facts about speaking stem from its general accuracy and fluency.

These normal levels of speech performance reflect the workings of an information-handling system of great complexity and considerable efficiency. To discover the system's properties, we must rely on subtle but powerful experimental techniques that can reliably detect the fleeting cognitive operations that give rise to speech. These techniques make it possible to examine the details of production processes in rigorous and systematic ways. To illustrate the experimental study of language production, we will present experiments designed to illuminate the workings of the components of the model in Fig. 12.1.

Message creation

Processing in the message component is responsible for determining the communicative content of the intended utterance. One thing a message

should do, if it is to succeed in communication, is ensure that the expressions that are used to refer to things will be understandable to the listener. For instance, if one wants a listener to pick out a particular photograph from among a set of photographs of buildings in New York City, one must choose a referring expression that makes contact with what the listener knows. A particular building could be referred to in any of several ways, such as "the Citicorp building" or "the building with the slanty roof". The first expression will work well if the person one is talking to is familiar with the names of buildings in New York City; otherwise, the second expression will work better. The decision about which expression to use is a decision about what to put in one's message.

Isaacs and Clark (1987) studied exactly this situation. In their experiment, one person (the director) viewed a display of postcards depicting landmarks in New York City. The second person in the experiment (the matcher) viewed a display that had the same postcards in a different order. Neither participant could see the other's display. The only task was to get the matcher's postcards into the same order as the director's. However, none of the pictures were visibly labelled, so the partners had to come up with their own expressions to refer to them. So the director could say something along the lines of "The first picture is the Citicorp Building" or "The first picture is a building with a slanty roof".

The twist in the experiment was that some of the directors and matchers were knowledgeable about New York City, and some were not. Isaacs and Clark found that the participants quickly determined whether their partners could successfully identify pictures based on proper names (*Citicorp Building*) or needed more descriptive expressions (*building with a slanty roof*). The directors adapted their subsequent utterances appropriately, using more descriptive expressions when their listeners were unfamiliar with New York City landmarks and more proper names when directors and matchers both were knowledgeable.

This is an example of just one factor that enters into determining the substance of an utterance. It may seem obvious, but its implications are far-ranging. The content of a message typically includes more than "just the facts" that the speaker intends to convey, going beyond them to incorporate information specifically tailored to the communicative context (Clark & Bly, 1995). Message formulation can demand a great deal of problem solving, much more so than the rather mechanical operations of the encoding processes. One upshot is that speakers frequently fail to plan their messages adequately, and communication may suffer as a result (Horton & Keysar, 1996).

In addition to specifying the content of an utterance, a message must also signal the relative prominence of its components. Most important, it must indicate which element is the topic of the utterance, what the

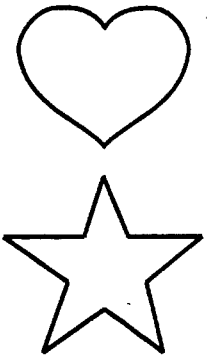

FIG. 12.4 Which component of this display is more prominent?

utterance is about. Things can become topics by attracting attention, a process illustrated in a study by Forrest (1993). In Forrest's experiments, speakers watching a computer screen had to describe the physical relation between two pictured objects like those shown in Fig. 12.4.

Immediately before the pair of objects was displayed, a cue was presented at the screen location where one of the objects would appear, drawing the speaker's attention to that location. Forrest found that this attentional manipulation strongly influenced how speakers described the scenes. Speakers were more likely to say "The heart is above the star" when their attention was directed to the heart than when it was directed to the star, and more likely to say "The star is below the heart" when their attention was directed to the star than when it was directed to the heart. Thus, it appears that an element that a speaker is attending to, an element that constitutes the intended topic, is likely to be given a prominent position in the utterance.

Grammatical encoding: Structural processes

Grammatical encoding is responsible for translating the message into a series of words, and comprises assigning syntactic functions (like subject) and arraying words in a grammatical order. Consider how one might describe the event pictured in Fig. 12.5. One could say, "The boy is being awakened by the alarm clock"—a passive sentence with "the boy" as subject—but one could also say, "The alarm clock is awakening the boy"—an active sentence with "the alarm clock" as subject. The idea is virtually the same in either case. What determines which noun is assigned the syntactic function of subject? Forrest's (1993) results suggest that function assignment takes the most prominent message element, the topic, and

FIG. 12.5 A target picture adapted from Figure 5 in Bock, Loebell, and Morey (1992).

assigns it the role of subject. Since animate objects (humans and animals especially) tend to attract attention in events, a likely subject is *the boy*.

Bock, Loebell, and Morey (1992) relied on this tendency to examine the relationship between function assignment and constituent assembly. In their experiment, subjects were presented with a series of spoken sentences and pictures, one by one. The subjects repeated each sentence and described each picture. The critical manipulation in the experiment lay in the properties of the sentences that preceded the target pictures (the "prime" sentences). The sentences varied on two dimensions, as illustrated in Table 12.2. One dimension was the animacy of the subject and object noun phrases: The subject of the prime sentence was either an animate noun phrase ("five people") or an inanimate noun phrase ("the boat"). The goal of the animacy variation was to prime the function assignment process: The type of assignment pattern used in one sentence (that is, in

TABLE 12.2
Sample Priming Sentences

Priming Condition	*Example Sentence*
Active, animate subject	Five people carried the boat.
Active, inanimate subject	The boat carried five people.
Passive, animate subject	Five people were carried by the boat.
Passive, inanimate subject	The boat was carried by five people.

the priming sentence) should tend to be repeated in a subsequent sentence (the sentence used to describe the target picture). For example, if the subject of a prime sentence was inanimate (either "The boat carried five people" or "The boat was carried by five people"), there may be a tendency to try to assign the inanimate entity in the target picture to the subject function (saying "The alarm clock wakened the boy"). The events depicted in the target pictures always included one animate and one inanimate entity, either of which could serve as the sentence subject (as in "The alarm clock wakened the boy" or "The boy was wakened by the alarm clock"). However, with function priming, there should be more sentences with inanimate subjects after primes with inanimate subjects than after primes with animate subjects.

The second variation in the priming sentences was in their syntactic form: They could be in the active voice ("Five people carried the boat"; "The boat carried five people") or in the passive ("Five people were carried by the boat"; "The boat was carried by five people"). The goal of this manipulation was to prime the constituent assembly process. Having used a particular syntactic structure for the prime sentence, the constituent assembly process may tend to repeat that structure in describing the target picture. So, active descriptions of the target picture should be more likely following active prime sentences than following passive prime sentences.

Both of these manipulations affected speakers' descriptions: First, inanimate subjects were more likely following primes with inanimate subjects, showing that function assignment tends to repeat the previous mapping of animacy to syntactic function. Second, active sentences were more likely following active primes than following passive primes, showing a tendency for constituent assembly to encore a recently produced structure. This argues that function assignment and constituent assembly are separable processes.

Grammatical encoding: Lexical processes

Recall that lexical selection is the process of identifying an entry from the mental lexicon for conveying the intended meaning. Earlier, we described the case of a brain-damaged speaker, which showed that a word can be selected and its grammatical properties accessed without its sounds becoming available. Is this true for normal speakers? This seems to be what happens when one is in a tip-of-the-tongue state: One has a word in mind but the sounds of the word are inaccessible. To find out whether speakers in a TOT state have in fact selected a particular word (rather than a concept), Vigliocco, Garrett, and Antonini (1997) asked speakers of Italian who were having a TOT experience to identify the gender of the word that they were unable to retrieve. Even when they knew nothing

about how the word sounded, the speakers were able to identify the gender over 80% of the time. Because the *concepts* expressed by these nouns were neither masculine nor feminine, the ability of the Italian speakers to identify the gender of a wayward noun must be due to their having selected a particular word. This argues that word meanings and word forms are separately represented, and that lexical retrieval is a necessary component of the production process.

Phonological encoding

When the form of a word is retrieved, its sounds must be individually encoded in preparation for ordered production. We pointed out earlier that sound exchanges seem to involve phonemes of the same phonological category (consonant or vowel), in similar phonetic environments, and in similar syllable position. For example, in the vowel exchange "accipital octivity" (see Table 12.1), the exchanging vowels were both followed by the consonant /k/ and occurred in the same syllable position.

Dell (1984) tested this observation experimentally, using a procedure developed by Baars, Motley, and MacKay (1975). In this procedure, the speaker is presented visually with pairs of words, like the following:

bid meek
bud meek
big men
mad back

After certain pairs, the speaker is cued to say the words out loud. Sometimes speakers make errors when saying the words, and these are recorded. In order to increase the likelihood of an error, the target trials are preceded by several trials that are designed to bias the subject to make a slip. For example, if the target trial is the pair *mad back*, the three preceding trials would all have pairs in which the first word started with a /b/ sound and the second with an /m/ sound, biasing the subject to slip on the target pair and mistakenly say *bad mack*. Dell compared the likelihood of a slip occurring for pairs like *mad back*, in which both words have the same vowel, and pairs like *mad bake*, in which the words have different vowels. He found that exchanges of word-initial phonemes (*bad mack* or *bad make*) were more likely when the words contained the same vowel, confirming the pattern hinted at in naturalistic errors.

An important aspect of phonological encoding is placing sounds in syllable frames, as illustrated at the bottom of Fig. 12.3. This helps to ensure the correct ordering of phonemes, because syllable frames specify the order of consonants and vowels within the syllable. Syllables with

different orderings are said to have different consonant-vowel (CV) structures. So, for example, the syllables "kem" and "til" have the same structure—both are consonant-vowel-consonant (CVC)—even though they are made up of different phonemes. The syllable "tilf" has a different structure (CVCC).

To test experimentally whether the phonological component represents the consonant-vowel structure of syllables during production, Sevald, Dell, and Cole (1995) asked speakers to repeat a pair of nonsense words as often as they could in a four-second interval. The two words had either the same or different CV structures. In the shared-structure conditions, the first two syllables of the pair had the same structure, e.g. "kem tilfer" (note that the second word is made up of the syllables "til" and "fer"). In the different-structure conditions, the two syllables differed in structure, e.g. "kem tilfner" (the second word is made up of "tilf" and "ner"). If the phonological component represents the CV structure of syllables, and this structure can be re-used from one syllable to the next, then speakers should be able to say "kem tilfer" more often in four seconds than "kem tilfner". Of course, the pair that has the most phonemes would be expected to take longer to say, regardless of structure, so Sevald et al. also compared utterances like "kemp tilfner" and "kemp tilfer", where the shared-structure pair had more phonemes than the non-shared pair.

The results showed that speakers could say the shared-structure pairs more quickly than the non-shared pairs. Even more strikingly, repeating the exact same phonemes as well as the structure ("til tilfer" or "tilf tilfner") did not speed production any more than just repeating the structure. So the study suggests that part of phonological processing is representing syllables in terms of their CV structure, independent of the particular sounds that instantiate that structure.

The experimental approach to the study of language production has largely confirmed the picture developed through analysis of spontaneous speech errors, but there are a number of advantages of experimental methods over purely observational investigations. First, experiments avoid some of the biases inherent in the collecting of errors. Some speech errors are simply more noticeable or more easily remembered than others. Experiments can investigate the full range of errors and so give a more accurate picture of production failures. Just as important, experiments extend the investigation of language production to include aspects of normal error-free production that cannot be investigated adequately through naturalistic observation. For instance, because speakers often talk about animate entities (such as other people), in naturally occurring speech the subjects of sentences are often animate. It is only through experimental manipulation that this normal correlation of animacy and subjecthood can be teased apart in order to investigate (for example) the

process of function assignment. Experimental approaches to the study of language production therefore promise to shed new light on the workings of the language production system, going well beyond the understanding provided by the analysis of speech errors.

THE IMPORTANCE OF LANGUAGE STRUCTURE

We have already alluded to one of the most striking facts about language production, but it deserves explicit mention. At every level of processing, there are powerful structural constraints that govern the arrangements of elements. In speech errors, this shows up as restrictions on the elements that interact with one another. When words interact with one another, they come from the same grammatical class. When sounds interact with one another, they come from the same phonological (vowel or consonant) class. The consequence is that the basic structural patterns of the utterances are preserved even in errors.

This principle is vividly clear in stranding errors. In "you ordered up ending some fish dish" (said instead of "you ended up ordering some fish dish"; Garrett, 1993), the past tense and progressive affixes (*-ed* and *-ing*) occurred in the correct locations. Had they moved along with the word stems, the sentence would have become ungrammatical: "You ordering up ended some fish dish." But the inflections generally do not move, and the syntactic structure remains intact.

Notice that the structure is stable in the face of radical distortions of the speaker's intended meaning. Errors such as "dinner will be served at wine" (Fromkin, 1973) and "a room in your phone" are abysmal failures as vehicles for the speaker's communicative intention. (Take a moment to work out what the speakers of these errors actually intended. Both errors involve word exchanges.) Yet the utterances observe normal grammatical constraints. The rarity with which errors make sense, hidden or otherwise, is a challenge not only to claims like Freud's but also to any theory that overlooks the complex information processing system that mediates the translation of thoughts into language.

Considerations such as these have led contemporary theories of language production to emphasise the structural and information-processing constraints on speech, rather than the processes determining what the speaker means to convey. As one of the contemporary pioneers of language production research pointed out (Garrett, 1980a, p. 216), "The production system must get the details of form 'right' in every instance, whether those details are germane to sentence meaning or not". For example, verbs in English agree in number with their subjects ("She sneezes" vs. "They sneeze"), and this agreement operates in virtually every utterance that a speaker produces. Yet number agreement affects meaning

hardly at all. If number inflection is omitted by mistake, the sentence is still understandable. In the past tense, number is completely unmarked on most verbs ("She sneezed" and "They sneezed"). Even so, when number marking is required, speakers almost always get it right. Thus, the creation of linguistic structure is central to any account of what people do when they talk.

LANGUAGE PRODUCTION IN PSYCHOLINGUISTICS

Speaking is only one part of what we do with language, because speakers are also listeners. Neither speaking nor listening can be fully explained without the other, so the relationship between language production and language comprehension is a natural target of curiosity. Both put our knowledge of language to work. But they put it to work in different ways, inasmuch as they differ in their goals. Production starts from a meaning to be conveyed and then works to convert that message ultimately into a series of speech sounds. Comprehension, by contrast, starts with speech sounds (or written letters) and works toward determining the meaning conveyed by those sounds.

As a consequence of this fundamental difference in the problems that the two systems must solve, we might expect them to differ considerably in their operation. For instance, the comprehension system must have ways to deal with ambiguous input, because natural language is rife with ambiguities. When one hears a sentence such as *The spy observed the man with binoculars*, one must decide whether the spy is using the binoculars to observe the man, or whether the spy is observing a man who is carrying binoculars. There is presumably no ambiguity in the mind of the speaker of the sentence, though, about which message is intended. Conversely, it is possible that comprehension of some sentences does not require a complete syntactic analysis of the sentence; the meanings of the words might be enough to give the listener a good idea of the meaning of the sentence as a whole. If one hears a sentence containing "villagers," "soldiers", and "massacred", one can understand that the soldiers massacred the villagers without having to decide which noun was the subject of the sentence and which the direct or indirect object. Of course, one might be wrong—the villagers may actually have massacred the soldiers—but there may be times when just knowing the words in their context is enough to understand the speaker, without a complete syntactic analysis of the utterance. But in producing a sentence, a speaker necessarily assigns syntactic functions to every element of the sentence; it is only by deciding which phrase will be the subject, which the direct object, and so on that a grammatical utterance can be formed—there is no way around syntactic processing for the speaker.

Although there is reason, then, to believe that production and comprehension may operate in different ways, there is also reason to believe that at some level they draw on the same linguistic knowledge. After all, every one of us both speaks and understands our native language. Communication occurs because speakers and listeners know the same code, a code that governs how arrangements of sounds and words convey meaning, allowing us (Pinker, 1994, p. 15):

> to shape events in each other's brains with exquisite precision. I am not referring to telepathy or mind control or the other obsessions of fringe science; even in the depictions of believers these are blunt instruments compared to an ability that is uncontroversially present in every one of us. That ability is language. Simply by making noises with our mouths, we can reliably cause precise new combinations of ideas to arise in each other's minds.

CONCLUSION

Titchener's (1909) introspection failed him when he claimed that talking is as easy as reading from a memory manuscript. The errors that people make reveal that speech is the product of a complex information-processing system that must piece together words and sounds to convey messages. We have sketched an outline of that system and showed how the workings of its components are being explored through experimental research. As this research proceeds, our sketch will come closer to a blueprint of how normal speakers draw on their linguistic knowledge to formulate utterances. The confluence of this explanation with models of normal comprehension, with accounts of the development of fluent speaking ability in children, and with descriptions of the disintegration of speech due to brain injury and the diseases of ageing, should offer a better understanding of the cognitive architecture of human language and its contribution to human communication.

ACKNOWLEDGEMENTS

This chapter is abridged and adapted from Bock (1995). Its preparation was supported in part by grants from the National Institutes of Health (R01 HD21011, F32 DC00141) and the National Science Foundation (SBR 94-11627). We thank David Huitema for comments on an early draft of the manuscript.

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 from artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior, 14*, 382–391.

Badecker, W., Miozzo, M., & Zanuttini, R. (1995). The two-stage model of lexical retrieval: Evidence from a case of anomia with selective preservation of grammatical gender. *Cognition, 57*, 193–216.

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.

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, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99*, 150–171.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Clark, H.H., & Bly, B. (1995). Pragmatics and discourse. In J.L. Miller & P.D. Eimas (Eds.), *Handbook of perception and cognition. Vol. 11: Speech, language, and communication* (pp. 371–410). Orlando, FL: Academic Press.

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. (1995). Speaking and misspeaking. In L.R. Gleitman & M. Liberman (Eds.), *An invitation to cognitive science: Language* (Vol. 1, pp. 183–208). Cambridge, MA: MIT Press.

Ellis, A. (1980). On the Freudian theory of speech errors. In V.A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 123–131). New York: Academic Press.

Fink, B.R. (1986). Complexity. *Science, 231*, 319.

Forrest, L.B. (1993). *Syntactic subject and focus of attention*. Unpublished masters thesis, University of Oregon, Eugene, OR.

Freud, S. (1924/1935). *A general introduction to psychoanalysis*. New York: Washington Square Press.

Fromkin, V.A. (Ed.). (1973). *Speech errors as linguistic evidence*. The Hague, The Netherlands: Mouton.

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. (1980a). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (pp. 177–220). London: Academic Press.

Garrett, M.F. (1980b). The limits of accommodation: Arguments for independent processing levels in sentence production. In V.A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 263–271). New York: Academic Press.

Garrett, M.F. (1993). Errors and their relevance for models of language production. In G. Blanken, J. Dittmann, H. Grimm, J. Marshall, & C.-W. Wallesch (Eds.), *Linguistic disorders and pathologies* (pp. 72–92). Berlin: Walter de Gruyter.

Heeschen, C. (1993). Morphosyntactic characteristics of spoken language. In G. Blanken, J. Dittman, H. Grim, J. Marshall, & C. Wallesch (Ed.), *Linguistic disorders and pathologies* (pp. 16–34). Berlin: Walter de Gruyter.

Horton, W.S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition, 59*, 91–117.

Isaacs, E.A., & Clark, H.H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology: General, 116*, 26–37.

James, W. (1890). *The principles of psychology, Vol. 1.* New York: Dover.

MacKay, D.G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia, 8*, 323–350.

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

Meringer, R., & Mayer, K. (1895/1978). *Versprechen und Verlesen.* Amsterdam: John Benjamins. (Original work published 1895)

Pinker, S. (1994). *The language instinct.* New York: Morrow.

Roget, P.M. (1852). *Thesaurus of English words and phrases, classified and arranged so as to facilitate the expression of ideas and assist in literary composition.* London: Longman, Brown, Green, & Longmans.

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.

Sevald, C.A., Dell, G.S., & Cole, J.S. (1995). Syllable structure in speech production: Are syllables chunks or schemas? *Journal of Memory and Language, 34*, 807–820.

Titchener, E.B. (1909). *Lectures on the experimental psychology of the thought-processes.* New York: Macmillan.

Vigliocco, G., Antonini, T., & Garrett, M.F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science, 8*, 315–317.