long pauses. Unlike slips of the tongue, such dysfluencies have been perceived as unfortunate and uninformative elements of the speech stream. In contrast to this view, Fox Tree argues that these phenomena form an integral part of the communicative enterprise. She argues that certain paralinguistic phenomena are used to get around some of the problems inherent to the communication medium of spontaneous speech. They are necessary to coordinate conversation, achieve the grounding criterion, indicate turn units, create a coherent discourse, and warn listeners of upcoming dysfluencies and production trouble.

In summary, the chapters in this volume address a wide range of issues relevant to the cognitive modelling of language production processes. Current theories of phonological encoding, word retrieval, sentence formulation, conceptual structure, and spontaneous speech are evaluated. In addition, the relevance of these models for other areas of language research is addressed, including language comprehension, neuropsychology, language development, and linguistic theory. Finally and most importantly, what emerges from the chapters in this volume is a new set of research questions that should keep us busy well into the third millennium.

REFERENCES

CHAPTER TWO
Producing words: How mind meets mouth

Kathryn Bock
Department of Psychology and the Beckman Institute for Advanced Science and Technology, University of Illinois, USA
Zenzi M. Griffin
Department of Psychology, Stanford University, USA

INTRODUCTION
One of the premises of contemporary cognitive science is that an organism’s response to a stimulus requires sensory, perceptual, and memory processes on an order of complexity unexpected from observations of simple reactions to physical stimulation. It was a touchstone of the cognitive revolution in the late 1950s and early 1960s that these interpretative processes are a proper subject for psychology, setting the goal of explaining how a stimulus is processed through successive mental transformations. In contrast, within behaviourist conceptions that regarded the relationships between stimuli and responses as little more than learned reflexes (Skinner, 1953; Watson, 1913), the role that mental transformations played was on a par with the hypoten in the term stimulus-response (S-R) psychology.

Put into the terms of the theoretical framework that cognitive psychology supplanted, the thrust of the cognitive revolution was therefore directed toward elaborating the S of S-R theories. Cognitive theories have given considerably less attention to elaborating the mental processes by which interpretations or other representations of meaning guide and gain expression in action. There are many reasons for this inattention. First and foremost, action suffered from its associations with behaviour and the eponymous -ism. What behaviourists were seen as neglecting most egregiously were the processes of stimulus interpretation, and these processes judiciously took priority. Second, to begin to explain the cognitive processes that guide action, one must have in hand some framework.
for representing interpretations, because interpretations are the starting point for the formulation of goals. To be firmly grounded, the elements of such a framework should come from theories of stimulus processing. Third, the practical problems of conducting controlled experiments stymied efforts to explore the cognitive processes of action in many domains, including the psychology of language. And finally, with many experiments requiring only button-press responses, there was little demand for a cognitive theory to explain how the index fingers are controlled. For all of their merits, these things had an unfortunate by-product. To a considerable degree, cognitive psychology lost sight of the need to explain how actions are guided and how they are learned. Reflecting this, many views of cognition tacitly assume that the essential elements of action comprise its motor components alone. Figure 2.1 caricatures this state of affairs with a cognitive model that includes multiple stages of stimulus processing and no stages of response preparation.

Psycholinguistics is one of the areas that disdained the action side of the cognitive relationship between stimuli and responses, and it is the one that we will highlight in this chapter. Benignly reflecting psycholinguistics' neglect of action, in a 1974 review of experimental psycholinguistics, Johnson-Laird defined the field as the study of language comprehension, with some justification. Little work had been done, explicitly, on language production. This single-minded emphasis on the cognitive processes of interpretation (language comprehension) coupled with the disappearance of experimental investigation of the cognitive processes of expression (language production) led to a marked imbalance in our understanding of adult language use. This distanced the scientific enterprise from the prototype of skill in language, an ability to speak fluently.

There were scattered efforts to change the balance (Carroll, 1958; Deeke, 1984; Garrett, 1975; Osgood, 1971), and these efforts have gathered strength and coherence as the problems of language production become more widely recognized (Bock, 1995; Clark, 1996; Dell, 1995; Garrett, 1988; Leveti, 1989).

Our goal in this chapter is to broaden this recognition by considering how certain findings from research on language comprehension may bear on our understanding of language production, and vice versa. Our primary focus will be the mechanisms of lexical retrieval. To begin, we sketch how the cognitive processes of expression seem to be organized, in general, and how word production works, in particular. We then contrast the lexical processes of production with the lexical processes of comprehension, revisiting some key claims about word recognition in light of the implications from research on word production.

THE ANATOMY OF LANGUAGE PRODUCTION

The components of the language production system that are directly implicated in the formation of an utterance include a representation of the message, at least two types of lexical retrieval operations, several structural procedures responsible for combining and ordering words and sounds, and all the intricacies of articulatory planning and execution. In Figure 2.2 we offer a rough sketch of this system along with examples to illustrate selected operations that must be implemented to retrieve a single word (rabbit) in the course of creating a spoken utterance (e.g., "There's a rabbit").

In keeping with our emphasis on lexical processes, we will have little to say about the mostage apart from assuming that there is one. The key feature of messages is that they stand at the interface between thought and language. They are not verbal or word-based: Messages may be much the same in nature whether the eventual utterance is in English, Giugga Yimithirih, Japanese, or Zoltan' (Potter, So, Von Eckhardt, & Feldman, 1984). In the simple laboratory tasks that are used in most research on word production, a message is rarely more than a single concept. Picture-naming tasks evoke concepts with pictures of concrete objects (like the rabbit in Figure 2.2). In classic Stroop studies, in which speakers are asked to name the color of the ink in which a word is printed, the concept is evoked by the colored stimulus (the ink itself or a color patch; Stroop, 1935). More generally, a message can be viewed as the penultimate link in the traditional information-processing chain from sensory processing to output. As such, it is a perceptual or conceptual categorization of the input.
Rhymes with “Babbit”. It comprises the word’s phonological segments and perhaps also its stress pattern and syllabification. Phonological encoding uses this information to smoothly integrate the sound segments for a word (r, a, h, b, r, etc.) into the stream of speech. In connected speech, this means that a phrase such as “another rabbit” may be uttered by speakers of North American English as “a NOB bit”, combining the r of rabbit and the r of another. This process of resyllabification has familiar reflections in children’s word play (“I scream, you scream, we all scream for ice cream”) and song lyrics (“Mickey mouse and davy doots and little lissy divvy”), and contributes to one of the central problems in auditory word recognition, the problem of segmenting words (e.g., hearing “The mean of them all” as “The mean of the mall”). For accounts of speaking, the implication is that retrieving the sound form of a word is not sufficient to produce the word, without additional work.

By setting aside the processes of conceptual identification and phonological encoding, we have exhausted almost everything that there is to word production as it is sometimes conceived. But since there is an entire chapter to come, one may suspect that there is more to be said. What more there is comprises the lexical processes that Figure 2.2 depicts as components of grammatical and morphological processing, lexical selection, and lexical retrieval.

The motivation for linking lexical selection and retrieval to grammatical processing stems from the kinds of information that words carry about their structural and positional privileges and requirements. In everyday language use, words are rarely produced in isolation. Instead, they occupy places within strings of words, with their positions determined in part by their grammatical categories (e.g., in most English declarative sentences, at least one noun will precede a verb) and with their pronunciations determined in part by their positions (e.g., in many dialects of American English, the definite article the is pronounced “thuh” when it precedes a word that begins with a consonant and “thee” when it precedes a word that begins with a vowel). Thus, in order to create normal utterances, speakers must recover information about the grammatical classes of words and about the words’ sounds before phonological encoding can be completed.

Implicit in the architecture of Figure 2.2 is the claim that information about grammatical class is recovered during lexical selection and that information about sound is recovered during lexical retrieval. Selection makes it possible to assign syntactic roles and retrieval makes it possible to inflect word forms.
appropriately for their positions in an utterance (see Bock, 1995 and Bock & Levelt, 1994, for some of the evidence for these assumptions in the context of sentence production). In the next sections we further detail the processes of selection and retrieval, and then survey some of the reasons for believing that these processes occur not only in connected speech, but even when words are produced in isolation.

Lexical selection and retrieval

Lexical selection involves finding a lexical entry (technically, a lemma) that adequately conveys some portion of a message, ensuring that there exists a word in one’s mental lexicon that will do the job. A rough analogy is looking for a word in a reverse dictionary, which is organised semantically rather than alphabetically. If the desired meaning is listed in the dictionary with a single word that expresses the sense, there is an appropriate word to be had in the language; if not, the search fails. The mental lexicon is presumably accessible in comparable fashion, permitting speakers to determine whether they know a word that will convey the meaning they intend. Most English speakers, for example, will find at least one lemma for their concept of a member of the family Oryctolagus cuniculus.

Locating a lemma yields basic information about how the impending word combines with other words. This corresponds to information about grammatical class (noun, verb, adjective, etc.) and other grammatical features that control a word’s combinatorial privileges and requirements (e.g., nouns must be specified as mass or count, and if count, as singular or plural; verbs must be specified as intransitive or transitive, and if transitive, as simple or ditransitive, etc.). The lemma for an instance of Oryctolagus cuniculus, for example, is a noun, count, and singular.

Once a lemma is found, the next step is retrieving the word’s morphological form (technically, its lexeme). In connected speech, this may encompass inflectional processes that adjust a word’s morphological structure to its syntactic environment (e.g., making a verb singular or plural). These lexical retrieval processes yield an abstract specification for the morphology of the selected word, a representation suitable for guiding the process of phonological encoding. So, retrieving the lexeme for the singular count noun that denotes a member of the family Oryctolagus cuniculus should yield a specification of the morphological structure of /r/ë/ob/.

Two steps to pronunciation?

Contemporary theories of how we produce language are largely agreed on a distinction along the lines of the one we have drawn between lexical selection and lexical retrieval (though see Caramazza, 1997). The distinction can be found in theories that are as different in other respects as those of Butterworth (1989), Dell (1986), Garrett (1975), Levelt (1989), and MacKay (1982). The key feature is the existence of two lexically specific steps separating meaning from sound: in order to convey a meaning with a word, the speaker must locate both a lemma and a lexeme in the mental lexicon. However, apart from production theories, the need for lexical selection (lemma retrieval) is sometimes disputed (Starreveld & LaHeij, 1996). This warrants consideration of the arguments that motivate the lemma hypothesis.

The first argument is logical, and reflects a basic property of language: The mapping from meaning to sound is arbitrary. This fact has consequences for the representation of words in the mental lexicon (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997). Suppose a speaker wants to produce a word that denotes the concept FEMALE PARENT. This concept is part of a semantic field that includes other concepts like FATHER, MAN, and WOMAN, whose meanings are partially indicated in Table 2.1. Consider the relationship between the matrix in the table and the conditions for an /m/ at the beginning of a word, as an English speaker would utter in expressing the FEMALE PARENT concept with the word mother. For /m/ to be used appropriately in this case, both the FEMALE and PARENT features are relevant. However, if a speaker wishes to express the concept ADULT MALE with the word man, neither the FEMALE nor the PARENT feature is relevant to the utterance of /m/. The upshot is that if a single one of the features (FEMALE, PARENT, ADULT, or MALE) is active, initial /m/ should be blocked, but if both FEMALE and PARENT or ADULT and MALE are active, initial /m/ is enabled. This is an instance of exclusive-OR, a well-known computational problem for single-layer associative networks (Minsky & Papert, 1969). Its solution entails that the mapping from meanings to sounds, even for very simple sets of meaning and sounds, includes an additional set of representations. In connectionist frameworks this is a layer of hidden units, and in an interpreted model of language production, it is a layer of lexical entries.

From an empirical standpoint, the most readily appreciated evidence for the lemma hypothesis comes from the universal tip-of-the-tongue experience. This is the familiar frustration of being unable to retrieve the sounds of a word that one knows and—the source of the frustration—knows one knows. This

<table>
<thead>
<tr>
<th>Target Word</th>
<th>Conceptual Features</th>
<th>Phonological Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>mother</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>father</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>woman</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>man</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
metacognitive awareness is sometimes called a “feeling of knowing”. It implicates the identification of an appropriate lemma in the mental lexicon without lexical retrieval, since the phonological form is unavailable.

A concrete link between the tip-of-the-tongue (TOT) state and lemma identification can be found in the kinds of information that are available to speakers in such states. It is common knowledge among psychologists that partial phonological information can sometimes be retrieved: In an early study, Brown and McNeill (1966) discovered that speakers in tip-of-the-tongue states could accurately report the beginnings of otherwise-inaccessible words between 45% and 20% of the time. However, the lemma hypothesis makes the prediction that other kinds of information should be accessible in tip-of-the-tongue states. Specifically, if a tip-of-the-tongue state signals the successful identification of a lexical entry, grammatical information about the word should be available even if phonological information is not.

This proposition has been examined in several studies involving Italian subjects. Like many other languages, Italian has grammatical genders that control (among other things) the determiners that nouns take and the forms of the adjectives that modify nouns—the kinds of information that are needed for a word to combine appropriately with other words. Excepting those few nouns whose grammatical genders reflect biological gender, gender is semantically arbitrary and unpredictable from the word’s meaning or reference. These properties were exploited by Vigliocco, Antonini, and Garrett (1997) and Miozzo and Caramazza (1997) in research on the information available to Italian speakers in tip-of-the-tongue states. The results showed that speakers reliably identified the genders of words whose phonological properties were inaccessible, and did so even for words whose phonological forms offer no gender clues (though there are clear phonological correlates of gender in Italian, many words have no overt gender marking or have contrary marking). The rate at which gender was correctly reported was significantly greater than chance—over 80% in Vigliocco et al. (1997) and about 70% in Miozzo and Caramazza (1997).

Interestingly, in Miozzo and Caramazza’s research the rates for correct forced-choice identifications of initial sounds were virtually identical to the rates for gender identification. This gives no support to a subsidiary prediction of the lemma hypothesis that grammatical information should be more reliably available than phonological information. Unfortunately, Miozzo and Caramazza’s identification test always queried gender before sound, and almost 20% of the initially reported tip-of-the-tongue states were resolved by retrieval of the word during the search for gender information, compared to only 1% during the search for initial phoneme information. It is impossible to tell whether these differential resolution rates affected the results, or how.

In English, the distinction between mass and count nouns has grammatical consequences that require the lexical entries for individual words to indicate whether the word is mass or count. Mass nouns cannot be pluralised (at least not with the same sense as the singular form: compare wheat and wheats), cannot take indefinite articles (compare a bread and a bagel), cannot be used with quantifiers like every (compare every bean and every corn), and so on. Although the grammatical distinction has an obvious semantic basis, it is nonetheless unpredictable: Beans are indistinguishable from corn in most of the ways that matter semantically, but only beans are grammatically count. Noodles and pasta are almost synonymous, yet the former is a count noun and the latter a mass noun. This suggests that the mass/count distinction is a grammatical feature that might be accessible to English speakers in tip-of-the-tongue states, and it is (Vigliocco, Vishton, Martin, & Garrett, 1999).

Perhaps the most striking support for the lemma hypothesis comes from a case report about an Italian agrammatic aphasic (Buddecker, Miozzo, & Zanuttini, 1995). This patient had enormous difficulties naming even common objects. Despite this deficit, his ability to correctly identify the grammatical genders of words that he was unable to retrieve was nearly perfect, even when his performance was at chance on tests of his ability to identify the first sounds in the same words from two alternatives provided. Similar results from a French-speaking aphasic have been reported by Hensaff Goon, Bruckert, and Michel (1989), providing evidence for a lexically specific step of word retrieval that gives access to abstract grammatical information.

Apart from the disorders of aphasia and the disruptions to normal speech caused by tip-of-the-tongue states, there is evidence for the lemma hypothesis from recent studies using an event-related-potential measure called the lateralised readiness-potential. The lateralised readiness-potential (LRP) indexes electrophysiological activity in the brain prior to an overt response. Because the LRP begins to develop as soon as task-relevant perceptual and cognitive information become available to the motor system, it can be used to assess the preparation of the information that guides responding.

Using the lateralised readiness-potential to examine the time course of lexical access, van Turennout, Hagoort, and Brown (1997) asked Dutch speakers to name pictures of common objects. In three experiments, the LRP showed clear evidence of semantic information being activated prior to phonological information. A second series of experiments (van Turennout, Hagoort, & Brown, 1998) tested the lemma hypothesis directly. In these studies, speakers were occasionally cued to judge the grammatical gender or the sound of a Dutch word they were in the process of producing, with the judgements yoked in a way that made it possible to assess the relative time courses for the development of the relevant information. The results showed that grammatical gender was accessible earlier than phonological information, consistent with the lemma prediction.

The implication of these results is that production processes may yield lexically specific information prior to the point at which they achieve access to word forms. In terms of the model in Figure 2.2, lexical selection can be initiated before lexical retrieval, making way for the combinatorial-grammatical properties of words. Among the questions raised by this evidence is whether lemmas (perhaps including grammatical properties) are automatically accessed in the course of
producing single words outside of grammatical contexts. These questions have both theoretical and methodological ramifications for accounts of how we recognize and understand words, and in the next section we consider what those ramifications may be.

UNDERSTANDING AND PRODUCING WORDS

The cognitive mechanisms of word production have various counterparts in processes hypothesised to occur during auditory and visual word recognition. This seems only natural. Recognition is logically the reverse of the production mapping from meaning to sound, and speakers and listeners share the same linguistic knowledge, if they have the same native language. One consequence of this mirroring is that very similar issues occupy the literatures on word recognition and production. Among these issues there are various matters of experience (how often a word is produced or encountered), lexical organisation (phonological and morphological representation), access paths (what kinds of information are normally retrieved and with what priorities), and the flexibility of access (whether the access to information involves discrete or cascaded stages, whether stages interact, whether all stages are mandatory, whether the system is modular or open to influence from contextual information, and so on).

Despite these broad similarities, there are basic disparities in the processing problems that have to be solved during recognition and production. Production demands the creation or retrieval of different kinds of structures (including the phonological and metrical structures of words) whose properties may not be reconstructed to similar levels of detail during recognition. The production of a word begins with the meaning that the speaker has in mind (a meaning that is unambiguous, at least to the speaker) and proceeds from there through a retrieval process that may allow the simultaneous activation of all of a word’s phonological properties, prior to the sequencing of the word’s segments. In contrast, the auditory recognition of a word requires the accumulation over time of acoustic cues, segmentation of the word from a continuous stream of speech, and in due course, the recovery of meaning. All of this happens reliably despite some notorious problems with the acoustic signal: The input may be degraded (Samuel, 1981), a partial segment of speech may be momentarily ambiguous (Marlson-Wilson, 1989; Tanenhaus, Spiro-Knowlton, Eberhard, & Sedivy, 1995; Zwitserlood, 1989), or the segmentation cues may be misleading (Cutler & Butterfield, 1992). Even with these hurdles in the speech stream successfully crossed, a correctly segmented and recognised word may remain ambiguous in meaning (Simpson, 1994).

Some consequences of the disparities between word recognition and word production are that the processes almost certainly differ in how they access lexical information (Stallvo, McLeod, & Lewis, 1985), in the kinds of lexical information they access (Cutting, 1997), and in how they are affected by repeated use (Cutler, 1988). Other kinds of differences emerge from comparison of three roughly contemporaneous models of word production (Dell, 1986, 1988), auditory word recognition (Elman & McClelland, 1988; McClelland & Elman, 1986), and visual word recognition (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). Although the models share many basic architectural and processing assumptions (all are connectionist and localist and use spreading activation as a retrieval mechanism), they differ in important respects. The word production model (Dell, 1986) includes syntactic category information as part of each word’s lexical entry along with phonological frames (“wordshapes” in Dell, 1988) to order the word’s sound segments. Neither of these things can be found in the models of word recognition, which concentrate instead on the identification of individual sounds or letters as the first step toward word identification, and rely on time or space to convey information about order. Although these differences reflect contrasts in emphasis as well as contrasts in principle, they serve to highlight the dissimilarities in theoretical focus that characterise the literatures of word production and word recognition.

The dissimilarities in presumed cognitive mechanisms as well as in explanatory goals make it risky to generalise from how recognition works to how production works, and vice versa. In spite of this, research on word recognition has come to rely on many techniques that have production as their end product, and of necessity incorporate the cognitive processes of production in some form. Since the cognitive processes of production may be fairly complex, as we have argued, the variability that these processes introduce into the data in studies of word recognition may be considerable. Moreover, the variability may be misattributed to processes of recognition.

In the next section we will survey some of the research that has used production tasks to explore recognition processes. Our aim is to disentangle the components of subjects’ performance that reflect recognition from those that reflect the cognitive mechanisms of word production.

From recognition to production: one easy step?

Word production is often used in experimental psychology as a task for indexing successful recognition in many kinds of visual and auditory tasks. A few of its applications include assessing the consequences of lexical priming and masking on perceptual identification tasks, indexing the automatization of visual word recognition with Stroop tasks, as well as straightforward testing of picture, object, and auditory or visual word recognition. In a perceptual identification task, a briefly presented word must be named aloud. In traditional Stroop experiments, the task is to name aloud the ink colour in which a word is printed. In picture- and object-recognition tasks, the name of the depicted object must sometimes be produced aloud. In word-recognition tasks, the word itself may be repeated or pronounced aloud. In each case, the theoretical focus is commonly on the interpretative components of stimulus processing that yield recognition, although the automaticization of verbal or cognitive processing necessarily includes
pronounced a visually presented word aloud, and later produced the same word when naming a pictured object. In the picture-to-picture printing condition, subjects saw and named the same picture twice. The picture-to-picture condition was designed to promote implicit retention in the form of perceptual priming, which Park and Gabrieli defined as the facilitation from naming the same picture twice. The facilitation was assessed relative to the time needed for naming the pictured object in the word-to-picture condition. By this measure, across experiments there were roughly 94 ms of facilitation attributable to perceptual retention.

The logic of this test assumes that the word-production requirement of pronouncing a visually presented word is equivalent to the word-production requirement of naming a visually presented picture. There are reasons to doubt this: The pronunciation of a visually presented word may not require the selection or retrieval steps shown in Figure 2.2, since skilled readers seem to have the ability to retrieve the sounds of words from their orthography alone (enabling them to fluently pronounce nonwords like zat, for example). Naming a picture, however, demands a more complicated mapping from meaning to sound. In support of this, Wheeldon and Monsell (1992) showed that the production of a picture name is primed more by a prior production of the same name elicited by a definition than by pronunciation of the visually presented word form. Still more striking, the priming is much less likely to be observed when the prime is a homophone of the target word (e.g., the playing-card spade is less likely to prime the digging-tool spade). So, even when the phonological form and articulation of a target word is identical to that of the prime word, the amount of repetition priming can be negligible. The critical component of this lexical repetition priming, Wheeldon and Monsell argue (1992; also Monsell, Matthews, & Miller, 1992), has a focus after the activation of meaning and before the execution of pronunciation.

This kind of priming is uncontrolled in the experiments by Park and Gabrieli (1995) and in related work on normal (Cave, 1997; Cave, Box, & Cobb, 1996) and amnesic speakers (Cave & Squire, 1992). It is large in magnitude (ranging from 50 to 120 ms in the experiments by Wheeldon & Monsell, 1992) and long-lasting (persisting over 60 to 120 trials in the conditions of Wheeldon and Monsell’s studies). Its contribution to Park and Gabrieli’s results can be seen most clearly in their unexplained finding that pictures with low name-agreement produced significantly more repetition priming than pictures with high name-agreement. Name agreement (or codability; Lachman, 1973; Mitchell, 1989) is a factor that affects lexical selection during word production (Griffin & Bock, 1998; Johnson, 1992; Roelofs, 1992; Schacter, Christenfeld, Ravina, & Bilous, 1991). Similar variables have potent effects on name selection in studies of face recognition, and have been identified as components of production (Bredart & Valentille, 1992; Ellis, Flude, Young, & Burton, 1996). In short, some of the effect that has been attributed to implicit perceptual memory in research...
Perceptual memory for words. A popular task in the implicit-memory literature is word-fragment completion. As commonly implemented, subjects examine a list of words in order to make some judgement about each one (for example, how legible the word is, whether it rhymes with "Babbit," whether it denotes an animal). Later, subjects are given an ostensibly unrelated test in which words appear as fragments, with several letters replaced by blank spaces (e.g., _a_b_t). A variant of the task is word-stem completion, in which only the final letters are removed (e.g., tab__). By examining how many fragments are correctly completed (assuming that the fragments have unique solutions), or how many are completed with words from the original list rather than with alternative completions (e.g., with #hia that is #hgin), and comparing this performance to that of subjects exposed to the original list, a measure of the effect of the initial exposure to the word is obtained. As with other implicit memory tasks, the initial exposure has been found to benefit the fragment-completion performance of both normal and amnestic subjects regardless of whether they can explicitly recollect their previous encounter with the stimulus. So, like picture naming, fragment completion yields dissociations between procedural and declarative memory.

Word-fragment completion incorporates a word-production task, taking it out of the realm of simple perception (cf., Jacoby & Dallas, 1981). Similar arguments can be made about the gating procedure used in studies of spoken-word perception (Grosjean, 1980). The cognitive mechanisms that these tests call upon therefore include production components as well as perceptual processes. Some evidence for this can be found in the kinds of exposure conditions that benefit fragment-completion performance in the absence of any need to create a visual representation of the word. Kornatyn and Nico (1992) found equivalent priming of fragment completion from reading a word and from noting the word given its definition. Likewise, mentally translating words into the language to be used at test produced a large fragment-completion effect (Budden, Romilla-Marks, & Baden, 1994). These effects cannot be attributed to conceptual processing: Budden et al. (1994) showed that imaging the referent of a word at study produced no priming of fragment completion than mentally translating the word did, and Weldon (1991) showed that viewing a picture of the object denoted by a test word is less effective for later fragment completion than mentally generating the word. Moreover, the harder it is for participants to mentally name a picture at study (due to short presentation time or simultaneous shadowing of other words), the smaller the fragment-completion effect becomes (Weldon & Jackson-Barrett, 1993).

The upshot is that implicit memory, certainly as it is measured by picture naming and word-fragment completion, appears to have action components as well as perceptual components. These action components indicate that explanations for implicit memory (and related phenomena such as cryptomnesia; Brown & Murphy, 1989) must go beyond stimulus processing to encompass response preparation.

A related argument against perceptual-priming accounts of implicit memory has been made by Ratcliff and McKoon (1997; Ratcliff, Albrighton, & McKoon, 1997). They attribute implicit-memory effects to a decision bias in memory retrieval. Our contention is that when words constitute the domain of study, the explanation may be more parsimoniously grounded within a specific cognitive system that handles the information processing that people normally engage in when using words in everyday language use. For this kind of account, the implication of research on implicit memory is that the procedures of lexical selection and retrieval can be primed and are likely to be primed by overt and covert word production.

The cognitive neurobiology of perception and word recognition

Another major development in contemporary cognitive science is the increasing use of neurophysiological measures of cognitive processing. Among these measures, neuroimaging techniques such as positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and magneto-encephalography (MEG) have shown great promise for elucidating the brain mechanisms that support human cognition. The interpretation of the results from experiments using these kinds of measures depends critically on cognitive models of the tasks that subjects perform. But here too, the task models often appear to underestimate the cognitive complexities of formulating words for production.

In one of the first groundbreaking efforts to use neuroimaging techniques for localising the brain processes engaged in word recognition, Petersen, Fox, Posner, Mintun, and Raichle (1989) employed positron emission tomography to examine changes in blood flow during the processing of visually presented words. One of their manipulations addressed the semantic processing of words such as hammer by asking subjects to generate and produce words for actions typically carried out with the object denoted by the word (so, the subject would be expected to say "hammer"). To isolate semantic processing from the visual and articulatory components of the task, Petersen et al. examined the cortical activation patterns obtained from a control condition. In the control task, subjects saw the stimulus word and pronounced it aloud. These activation patterns were then subtracted from those obtained during the performance of the verb generation task. The reasoning behind the subtraction was that the difference between the two sets of patterns represented semantic processing, with visual processing and articulation stripped away.

Examined in this comparison was the nature of the processing that is required to support the voluntary retrieval and phonological assembly of a spontaneously
generated word (e.g., pound). These processes include locating the lexical entry for an appropriate verb, retrieving its word form, and carrying out phonological encoding. Because the only one of these processes likely to have been tapped in the word naming task was phonological encoding, the activities associated with the other processes remained in the measure for semantic processing of the stimulus word hamster. It may even be the case that phonological encoding was controlled for, given that the control word differed from the word articulated during the experimental task (see also Price, Moore, Humphreys, Frackowiak, & Frith, 1996).

The appropriateness of control conditions is essential to the functional localisations that PET research aims for, and Petersen et al. understood the risks of their subtraction approach. In subtracting out the regional activations that stem from incidental components of the task, the aim is to isolate the brain regions that subserve the specific cognitive ability that is the target of study. This demands that the controls be fitting companions for the experimental tasks. Furthermore, for the subtraction logic to work well, it is important at the outset to use a task that produces robust patterns of activation. Word-generation tasks do this, yielding strong signals in broad regions of activation, and for these reasons they continue to be used in PET research. However, the benefits have a number of drawbacks. Different studies find different patterns of activation, in part because of differences in the control conditions, and consequently there is considerable disagreement about the functional interpretations of the observed activity (Frith et al., 1996).

Some of these disagreements arose as an immediate result of efforts to resolve the inadequacies of the word-presentation control used by Petersen et al. (1989), and more recent research has begun to move questions about word retrieval toward centre stage (e.g., Warburton et al., 1996). Ironically, however, the complexity of word generation (the very thing responsible for all the blood that the task stirs up) appears capable of threatening simple efforts to control it. There are problem-solving components common with what we dubbed message generation ("what's the most likely thing to be done with this object?"), unusual metalinguistic judgements about whether or not a retrieved word is of the kind that is supposed to be produced ("is this word a verb?"), as well as uncertainties about the type and the grain of phonological information that is retrieved when generation is silent (as it sometimes is in imaging studies). Although some of these difficulties have been recognised, many have not because word production continues to be viewed as a single simple step from meaning to sound. When its complexities are fully acknowledged, it becomes clear that the picture that emerges may be different (Indefrey & Levelt, 1999; Levin, Pramstrale, Meyer, Helmines, & Salmelin, 1998).

Word recognition

The cognitive processes of word production often come into play in more traditional research paradigms in cognitive psychology. Many of these traditional paradigms were developed to study visual word recognition and its interactions with attention, and the results occupy a vast literature (Balota, 1994; Seidenberg, 1983). The central issues in this literature have to do with the perceptual and cognitive pathways that give access to word lexical information. These pathways have been hypothesised to make use of visual information (in the form of letters), phonological information (as a consequence of recoding letters into sounds), and semantic information (in the form of top-down contributions to perceptual processes). One of the perennial debates is about how to separate these processing events along these access pathways from events that are incidental to, or consequent upon, contact with lexical information. The debate is reflected in a division of processing effects into prelexical and postlexical, with postlexical effects deemed irrelevant to most accounts of the recognition process.

Since most of word production is postlexical by necessity, our assessment of relevance is the reverse: Some postlexical effects in word recognition may be straightforward reflections of production processes. The best known of these involve articulation. In delayed pronunciation tasks, visually presented high-frequency words are initiated and completed slightly more rapidly than otherwise comparable low-frequency words. This is not only because they are recognised faster, but also because they are articulated more fluently (Balota, Boland, & Shields, 1989; Balota & Chumbley, 1985). However, some postlexical effects on word recognition may be both more subtle and much larger. This is especially likely when word recognition is assessed in circumstances that tap cognitive components of word production like lexical selection and retrieval.

Identifying these circumstances requires a brief detour into the dual-route theory of word recognition (Coltheart, Curtis, Atkins, & Haller, 1993) and its connectionist competitors (Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989). According to dual-route theories, visual-word recognition can take place in either of two ways: (1) Through the use of grapheme-phoneme correspondence rules that convert print into sound (the assembled route); or (2) through an immediate visual route from the word form to its stored lexical entry (the direct route). When a word is read orally, the pronunciation may be achieved by either route. In both cases the pronunciation itself comes about through processes such as those of phonological encoding shown in Figure 2.2. However, the assembled route goes from print to phonological encoding, whereas the direct route goes from the visual word form to a whole-word representation. In some theories (see Seidenberg, 1993 for review) the whole-word representation is analogous to the lexeme (the word form tapped during retrieval in production) whereas in others it bears more similarities to an abstract lexical entry (compatible in some ways to the wotna that is tapped during selection in production). In connectionist frameworks (Plaut et al., 1996; Seidenberg & McClelland, 1989), lexical information is retrieved in the same manner for all words, with different degrees of fluency depending on the frequency with which particular spelling-to-sound correspondences are encountered in reading. However, complete
versions of the Plaut et al. model make use of two different kinds of information (semantic and phonological) in arriving at pronunciations, with the division of labor between them trading off in various ways. On this scenario, the circumstances in which semantics weighs more heavily in getting to a pronunciation should be the circumstances that are most likely to make use of existing processes for production, analogous to the use of the direct route in dual-route theory.

In drawing parallels to the cognitive processes of word production, the most relevant recognition results are those that reflect the workings of the direct route or the weights of semantic connections. Consider the standard finding that consistency matters much more for low- than for high-frequency words, with consistent words pronounced faster than inconsistent primarily when the words are low in frequency (see Jared, 1997, for review and qualification). The dual-route account of this interaction involves races between the two routes, and different outcomes depending on word frequency. For high-frequency words, the direct route tends to outpace the assembled route regardless of consistency, but for low-frequency words, the direct route beats the assembled route only when a word's orthography is consistent with its pronunciation in connectionist terms, low-frequency inconsistent words are slow because of the uncertainty of their mapping from spelling to sound, which is a joint consequence of sharing spelling-to-sound correspondences with few other words, of conflicting with many other words in spelling-to-sound correspondences, and occurring infrequently themselves (Seidenberg & McClelland, 1989). In Plaut et al.'s (1996) formulation, inconsistent low-frequency words are also the first to suffer from the withdrawal of a semantic contribution to the spelling-to-sound mapping, as a consequence of the weakness of the phonological mapping route.

With this issue as background, we examined some familiar phenomena for more hints about the workings of word production. Since both of the leading theories of word recognition use interactions with frequency as a litmus test for the involvement of production-relevant representations in pronunciation tasks, our cursory survey of the word recognition literature initially focused on frequency effects. Proceeding in the same vein, we then turned to word-recognition experiments that examined effects of semantic and syntactic priming and used tools likely to tap selection and retrieval processes in word production. Although most of our discussion deals with visual word recognition (because that is what most of the research deals with), we cite related findings from studies of spoken-word recognition where they are available.

**Frequency effects in visual-word recognition**

Reading skill is normally a product of explicit instruction. Because of this, the ability to read and the corresponding ability to recognise a visually presented word are paramount examples of skills that are learned to variable levels of proficiency depending on specific environmental support and practice. Repeated experience with printed words may be a precondition of acquiring and maintaining the skills. The consequence is that words that we read often are recognised readily, whereas words that we read less often are recognised less readily. Using various estimates of the frequencies with which words are encountered in print, untold numbers of experiments (beginning with Cattell, 1885b) have shown a recognition advantage for words that are likely to be familiar, and the advantage increases with increasing familiarity. Glaser (1992) estimated that increases in the printed frequency of a word speeds its naming by 30 ms for each log10 increment in frequency.

It is harder to make the connection from repeated production of specific words to production fluency for those words. In part, this is simply because of the difficulty of objectively estimating how often speakers produce particular words. But since almost all estimates of spoken-word frequency and familiarity correlate highly with estimates of printed-word frequency and familiarity, investigators have used these measures somewhat interchangeably to assess how frequent use affects word production.

These assessments reveal broad-based similarities between frequency effects in production and recognition. First, just as frequent words are more readily recognised, they are more readily produced, both in terms of speed and accuracy. Speed advantages for frequent words have been observed in experiments on picture naming (usually with controls for the recognisability of the depicted objects; Griffin & Bock, 1998; Hutton & Kohnik, 1984; Jenkins & Levitt, 1994; Lachman, Shafter, & Hart, 1974; Oldfield & Wingfield, 1965; Wingfield, 1968), word translation (Cattell, 1887; de Groe, 1992; Jenkins & Levitt, 1994), and close completion (Griffin, in press). Accuracy differences are seen in the incidence of speech errors on high- versus low-frequency words (Doll, 1990; Stoneberger & MacWhinney, 1986). Second, the impact on production fluency of frequency differences among words is greater in the lower than in the higher frequency ranges, yielding an inverse logarithmic relationship between frequency and production latencies reminiscent of the inverse logarithmic relationship seen in word recognition, though the slope is much greater. For picture naming, Glaser (1992) estimated the reduction in production latencies with increases in the printed frequency of a word at over 250 ms for each log10 increment in frequency. Finally, frequency effects are likely to be neutralised in strongly supportive contexts, both in word recognition (Becker, 1979; Grosjean & Itzler, 1984; Inhoff, 1984) and in word production (Daneman & Green, 1986; Griffin & Bock, 1998).

Two debates about frequency that are shared between the literatures on word recognition and production have to do with the nature of the experiences that yield frequency effects and the locus of the effects in the perceptual, cognitive, or action systems. These questions about causation and processing locus are tightly linked, inasmuch as any experience with a word has multiple components,
and any or all of the components could have consequences for subsequent encounters with the same word.

The traces of what kinds of experience? The nature of the requisite experience is not well understood, either for recognition or production. Despite the intuitive plausibility of a relationship between the recognizability of a word and the number of times the word is actually encountered in print, the validity of frequency counts as predictors of word recognizability has long been questioned (Gernsbacher, 1984; Howe, 1966). The same is true for word retrievability. One of the variables that seems to do a better job accounting for visual-word naming latencies (e.g., Brown & Watson, 1997; Rubin, 1980) and picture-naming latencies (e.g., Morrison, Ellis, & Quinlan, 1992; Snodgrass & Yudofsky, 1996) is age of acquisition. Age-of-acquisition norms reflect judgments of how early in life a particular word was learned, and they correspond well with empirical observations of word learning in children (Carroll & White, 1973a,b; Gilhooly & Gilhooly, 1980). But because they also correlate highly with measures of word frequency-in-print, separating the contributions is difficult, and it is tempting to view them as interchangeable measures of the same thing: Repeated encounters with words over the lifespan.

Consequently, investigators have been loath to sacrifice widely available frequency norms for a less convenient and seemingly less objective measure, so word frequency has remained the variable of choice for indexing the effects of experience. However, there are good grounds for taking seriously the evidence that age of acquisition particularly affects word production (Gilhooly & Logie, 1980a,b; Morrison et al., 1992; Snodgrass & Yudofsky, 1996). Chief among them are the relationships between age of exposure and language learning itself: No amount of later exposure seems to fully compensate for early deprivation of certain kinds of linguistic experience (Mayberry, 1993, 1994), and the phonology of language may be the most sensitive component (Bahrick, Hall, Goggin, Bahrick, & Bergtr, 1994; Felge, 1987; Fiege, Yeni-Komshian, & Liu, 1999; Scovel, 1989; Snow, 1979). The strong association between accent-free speech and age of exposure to a second language or dialect offers a familiar example of age-dependent effects in language acquisition. Together with evidence that frequency effects in production (probably subsuming age-of-acquisition effects) may be tied to phonological form (Dell, 1990, Exp. 1; Huttenlocher & Kibbie, 1985; Jescheniak & Levelt, 1994), the possibility that phonology has strong age dependencies offers an argument for viewing age of acquisition as an important (and perhaps the most important) index of the accessibility of word forms in adults (Brown & Watson, 1987).

In word recognition, it is less clear how age of acquisition and word frequency-in-print compare as metrics of familiarizability for visual word forms, per se. Gilhooly and Logie (1980a,b) found frequency-in-print to be a better predictor than age of acquisition for visual recognition thresholds (and for auditory recognition of faint speech), and a worse predictor for naming times. From this they argued that word frequency-in-print may be a better index of visual recognizability, whereas age of acquisition may be a better index of retrievability for production.

Where are the traces of experience to be found in processing? Gilhooly and Logie’s argument adds obvious the connection between the experiential sources of exposure effects and the processing locus for the effects. In line with their hypothesis that word-specific experience can influence the uptake of visual information, there is variability associated with word frequency in the word-identification times that are measured when monitoring eye movements in reading (Rayner, 1977; Rayner & Duffy, 1986). The story gets more complicated, however, with findings of word-frequency effects for fixation times during reading but not during visual search (Rayner & Raney, 1996), and frequency effects associated with the post-lexical decision component of word-nonword judgements (Balota & Chumbley, 1984; see Figure 2.1). These and similar results have prompted the conclusion that in word-recognition tasks, frequency variations have a greater impact later in processing than they do in the early perception of visual or auditory word forms (Balota, 1994; Cutler, 1995).

In word production, there is mounting evidence that the experience that is indexed by frequency counts (whatever that experience may be) has consequences for the retrieval of phonological word-forms (Dell, 1990, Exp. 1; 3; Jescheniak & Levelt, 1994; also see Huttenlocher & Kibbie, 1983). At the same time, Dell (1990, Exp. 4) found one effect that is not easily explained by this view: When sound errors create words, there is no reliable tendency for the words so created to be higher in frequency than the intended words. Dell was able to account for this result as well as all of the other phonological-frequency effects with a computational model in which lexical selection (i.e., the lemma) was the production locus for the impact of frequent use. Because the message conditions for selection vary greatly with the context of speech, one implication of this view is that an effect of frequent is to be found in the mapping from context-specific referential concepts to lemmas (cf., Viskovich & Humphreys, 1991). Such effects are well known under the heading of codability, name agreement, and typicality (Griffin & Bock, 1998).

It is tempting to look for a resolution to some of the uncertainties about the workings of frequency with a theory that bridges the mechanisms of comprehension and production. Word-recognition tasks that increase reliance on direct or semantic routes to word pronunciation may show the same kinds of frequency effects that picture-naming tasks do (which is to say that they may reflect age of acquisition more than frequency in print; Brown & Watson, 1987), whereas tasks that lean more on spelling-to-sound mappings may show larger effects of visual-word frequency. Some of the frequency biases in lexical decision may ride on the consequences of some operation like lexical selection (Monell, Duyts, & Haggard, 1990, Paap, McDonald, Schvaneveldt, & Noel, 1987). Of course, if
our experiences with words have many different facets, each with a proprietary frequency effect, there will be no single cause or processing locus. But for the same reason, distinguishing between the production and comprehension consequences of using words may make it easier to untangle the effects of experience at the same time that it yields an integrated view of language use.

To evaluate the prospects for this kind of integration, we used the existing literature to assess differences between various recognition and production tasks in their susceptibility to age-of-acquisition and frequency effects. From our own normative data and from various manuscripts and published papers (see Table 2.2), we culled average response times for individual words in visual-word naming, lexical decision, and picture naming. For as many of the words as possible, we also obtained age-of-acquisition norms (from Carroll & White, 1973a and Gilhooly & Logie, 1980a) and spoken- and written-word frequencies (from the CELEX database; Baayen, Feng, & van Rijn, 1995). Table 2.3 gives descriptive statistics for the words in the sample on each measure. For each of the three tasks, Figure 2.3 shows the linear regressions on spoken frequency (top panel) and age of acquisition (bottom panel). The results for written frequency were comparable to those for spoken frequency but generally weaker, so we omitted them for present purposes.

For each of the three task data sets, we performed two separate stepwise multiple-regressions. Age of acquisition and spoken frequency were entered first into alternative analyses, to determine how much residual variance could be accounted for by the unforced variable. For picture-naming times, age of acquisition and spoken frequency each accounted for significant variance regardless of the order in which they were entered, with age of acquisition being a slightly better predictor. For visual-word naming and lexical decision, only age of acquisition reliably predicted response times. So, regardless of whether frequency was entered into the regression before or after age of acquisition, it failed to account for a significant share of the response-time variability across items.
The results for picture naming are generally consistent with previous reports in showing age of acquisition to be the better predictor of naming times (Morrisson et al., 1992; Snodgrass & Vanderky, 1996), although the superiority of age of acquisition over spoken frequency was less pronounced in our item-based analyses. For both of the remaining tasks, age of acquisition was a markedly better predictor than spoken frequency; in fact, it was the only significant predictor of the two. Its predictive power across all three tasks is consistent with there being commonalities that are unlikely to be found in initial processing, since picture naming involves no word display.

Regardless of the impact of early acquisition or frequent exposure on word recognition and production, there is limited to the role that repeated exposure can play in word use. In recognition and production, the magnitude of frequency effects diminishes or disappears entirely when words appear in constraining contexts (Becker, 1977; Griffin & Bock, 1998). For this reason, a great deal of attention has been focused on the question of how supportive contexts facilitate word recognition. The simplest and most-studied contexts are those in which single words are presented in succession or as pairs, in so-called lexical priming.

**Priming**

The several phenomena of priming arise from successive exposures to related words. When the relationship between a prime and a target word happens to be one of identity, the main issues that arise have to do with repetition priming and the procedural/declarative memory dissociation, as described earlier. Here we will restrict our attention to the consequences of exposure to successive words that are related in meaning (semantic priming) or linked by a canonical sequence of syntactic categories that create a grammatically acceptable pairing of words (syntactic priming). A prototypical experimental paradigm involves seeing one word, a prime (e.g., cat) and then another word, the target (e.g., dog) on which some task is performed. The task that we will emphasise is vocal pronunciation (“word naming”), which produces high levels of accuracy but interesting variations in the amount of time required to perform the tasks relative to controls. Ideally, the control conditions include the same target words, but preceded by unrelated primes.

The most common effect of priming is to facilitate recognition of target words, with recognition operationalised as response time in naming and lexical decision tasks, or as accuracy in perceptual identification tasks. However, the explanations for the observed facilitation are points of considerable conjecture and controversy (Halota, 1990; Hodgson, 1991; McNamara, 1992; Neely, 1991; Norris, 1986, Ratcliff & McKoon, 1988; Seidenberg, 1995). To add to the frays, we will consider whether some or even most of the typically observed effects of semantic and syntactic priming reflect the cognitive retrieval processes of production.

**Semantic priming.** Semantic priming is an effect of semantic relationships between primes and target words. The semantic relationships that figure most prominently in analyses of semantic priming are of two kinds, taxonomic and associative. Taxonomic relationships involve concepts from the same semantic fields or categories, such as types of artifacts (desk-table as items of furniture, shirt-scarf as items of clothing) or natural kinds (dog-wolf as canine, chewy-pie as fruits). Associative relationships have their origins in co-occurrences of words or the objects they denote, and exemplify a more or less consistent meaning similarity. Some associated words are also taxonomic relatives (shirt-suit), whereas others are united by event scripts (milk-mouse), by normal predication (milk-white), and so on. The question is whether the recognition of the prime can facilitate the recognition of a semantically related target word by some mechanism specific to the input pathway.

If semantic relationships automatically prime the input pathway, their effects should be observable in recognition regardless of context. So, semantic priming in word recognition should be observable for target words seen in isolation or in sentences. Contesting this, Hess, Fox, and Carroll (1995) and Williams (1988) have convincingly demonstrated that the priming effects, which can be readily observed for words in isolation, vanish when the same word pairs are placed into sentence or discourse contexts. Hess et al. (1995) argue that priming comes not from the meanings of words per se, but from the contextual predictability or coherence of a word. As contexts change from single-word primes, to sentences, to entire discourse, what is predictable or coherent also changes. After reviewing the evidence, Seidenberg (1995) likewise suggests that whenever semantic priming is observed, it is a consequence of the selective activation of semantic features rather than activation of word meanings as wholes.

Although this kind of contextually modulated meaning activation may have little impact on the recognition of word forms, as Hess et al. suggest, it is central to the selection of words for production. Speakers begin with a nonverbal message that is influenced by the communicative context, and the features of the message reflect the idiosyncrasies of particular speakers in particular places and times. As a consequence, messages are unlikely to fully instantiate all of the features of meaning for a single word, but will approximate the conditions for using the word to a greater or lesser degree. The selection of a word indicates not only its conditions (its meaning features) but the contextually specific features of the message than do the features of other equally accessible words. Given this, the mechanism responsible for the bulk of semantic priming effects observed in word recognition may be the same kind of lexical selection that precedes word production.

A potential defect in this argument comes from the comparative strength of semantic priming for taxonomic relatives versus associates in word-naming tasks. In naming, taxonomic priming is less likely to be observed than associative priming. But in production, taxonomic relationships are necessarily more central
than associations, reflecting the kinds of meaning features that normally drive lexical selection. For example, semantic word substitutions (e.g., saying “your daughter” when “your sister” was intended) typically involve taxonomically related words rather than associatively related words (Hotopf, 1980). If semantic priming in visual-word naming arises from mechanisms that are likewise involved in normal word production, one might expect naming to benefit from selective activation between taxonomically related primes and targets. Yet there is little or no priming between them.

The resolution to this apparent contradiction can be found in research by Wheelock and Monsell (1994). Wheelock and Monsell showed that naming a pictured object took more time when the naming episode followed the naming of a pictured taxonomic relative. Interestingly, this effect was slightly (but not entirely) offset by a shortened facilitation. Thus, the ability of semantic priming between taxonomic relatives in single word naming may be due to a trade-off between a small and short-lived facilitation that is quickly overwhelmed by the large and long-lasting competition between closely related words. In semantic priming tasks, the presentation of primes and targets activates two conceptually similar words. Since naming requires the selection of just one of the words, this dual activation creates competition. Competition in turn creates errors in speech, and it slows production, as Wheelock and Monsell’s findings attest.

If the appearance of semantic priming in visual-word naming is attributable to events within the production system, one kind of information that should become available as a consequence is information relevant to the target word’s syntactic privileges. This would occur if the manipulations that gave rise to semantic priming in naming set into motion the selection process for word production. To see if this happens, we turn next to the literature on syntagmatic priming.

**Syntagmatic priming.** By syntagmatic priming, we mean a tendency for a word from one grammatical form-class to prime (e.g., those words that serve as grammatically acceptable successors to the prime. For example, the noun people might serve as a syntagmatic prime for the verb eat, analogous to a kind of response that is sometimes observed in tests of word association (and contrasting with the kind of response called paradigmatic, involving words from the same form-class, such as people-eat). (Ervin-Tripp, 1961).1

There is good evidence for syntagmatic priming in word naming (Beland, 1997; O’Scaighdha, 1997; West & Stanovich, 1985). The syntactic origins of the effects are clear in work by O’Scaighdha (1997) and by Peterson, Burgess, Def, and Eberhard (in press). O’Scaighdha obtained distinct semantic and syntactic effects on naming times in contexts that varied systematically in semantic and syntagmatic congruence. Peterson et al. presented naming targets to subjects within highly constraining discourse contexts, along the lines of “The man was old and feeble and it was believed that he would soon kick the...”. At the offset of the context, a word was presented visually that formed a syntactically congruent continuation (e.g., thing) or a syntactically incongruent continuation (e.g., gun). Despite the semantic anomaly of both the congruent and incongruent continuations, there was consistent facilitation for the congruent word. Peterson et al. even obtained syntagmatic priming for nonwords that were morphologically marked as nouns (glitter) or verbs (glatted).

On the assumption (common in research on word recognition) that syntactic category information is a by-product of lexical access or morphological analysis, such effects are unlikely to arise from the activation of input pathways during visual or auditory word-perception. In fact, if prime processing is restricted to input pathways, syntagmatic priming in naming may disappear: Semon (1901) failed to observe it when the primes were presented very briefly and masked.1

The constellation of conditions that yield syntagmatic priming in word recognition strongly suggest that when it is seen, it is correlated with the kinds of syntactic information required for production. Specifically, the grammatical category of the prime word becomes available and effective in influencing the response to a syntactically congruent target word. As the work by O’Scaighdha (1997) and Peterson et al. (in press) implies, the source of the effect is likely to be the congruence between the syntactic class of the target word and the available slots in a structural frame. One prediction from this explanation is that the effect of syntactic incongruence should be inhibitory, and it seems to be: West and Stanovich (1986, Exp. 4; see also Fedemeier & Bates, 1997) found that words in syntactically congruent contexts were named no faster than the same words in contexts that were neither semantically nor syntactically informative (e.g., The next word is...). In syntactically incongruent contexts, the words were named more slowly.

**Interference**

The complement to priming in word recognition is interference: Sometimes exposure to words retards processing instead of speeding it. The mechanisms of interference are at least as controversial as those of priming, if not more so, and have an even longer history in cognitive psychology by virtue of being intertwined with the operations of selective attention. We will sidestep the attentional issues to focus on the best-known example of an indirect word-recognition test that produces interference, the Stroop naming task.

**Stroop interference.** In its most familiar variants, the Stroop task (Stroop, 1935) involves the presentation of words printed in assorted ink colors, and requires naming the ink colors rather than the words themselves. The words, however, denote colors that are congruent or incongruent with the ink colors. For example, the word "blue" printed in blue ink would be a congruent stimulus (since the correct naming response fits both the ink colour and the word itself),
whereas the word red printed in blue ink is incongruent. Ink-colour naming for incongruent stimuli can be dramatically slowed, relative to the naming of congruent stimuli (a comprehensive review of research on the naming of Stroop stimuli can be found in MacLeod, 1991).

The participant's main goal in a standard Stroop experiment is to accomplish the fast and fluent utterance of a word, beginning with the minimal "messages" formed from ink-colour perception. The task of word recognition is to identify the name of the word, and the conflict is then presented for word production. This task is that the processing required for word production. However, relatively little attention has been given to how the production system that occurs. Most of the task's applications have been to questions about attention, visual word-recognition, and the development of word-recognition skills in reading. Since it is normally impossible to avoid recognising the word that constitutes the Stroop stimulus, the interference that arises during ink-colour naming has been used as an index of word recognition.

Yet the locus of much of the interference is to lie within the production system. One indication is that eliminating the verbal response reduces interference, although competition among lexical entries will remain if lexical selection is needed to mediate a button-press response (Keele, 1972) or typing of the word (Logan & Zbrodoff, 1998). Consequently, Stroop interference can be exploited to gauge the time course of events during word production. One experimental task that does so is picture-word interference.

**Picture-word interference.** In the picture-word interference paradigm, auditory or visual distractor words are presented at some point during the naming of a picture, analogous to the combination of word and colour name in the traditional Stroop task. To isolate the source of the interference, the distractors may be semantically related to the picture name, phonologically related, or unrelated. In addition, to determine when the interference occurs during the retrieval and production of the word, the timing of the distractor can be manipulated.

The results from this work have been to illuminate when during single-word production the information about meaning and sound is used. There are two findings of note. First, the effect of a semantic distractor is normally inhibitory, in line with the standard Stroop effect. Inhibition is maximal when an auditory distractor slightly precedes the to-be-named picture, peaking when the distractor precedes picture onset by about 150 ms. At this interval, the distractor appears to disrupt the selection of the lemma for the picture name roughly 300 ms later, or 150 ms after picture onset. (Levelt et al., 1998). At its peak, the magnitude of the inhibitory effect (relative to an unrelated control condition) is about 30 ms. Second, in contrast to the semantic inhibition effect, phonologically related distractors tend to facilitate picture naming. Maximal facilitation is seen when an auditory distractor and the to-be-named picture have simultaneous onset, so that the distractor's phonology can facilitate the phonological encoding of the picture name beginning about 300 ms after picture onset. The magnitude of the facilitation is in the neighbourhood of 50 ms (Schriefers, Meyer, & Levelt, 1990).

These parameters are comparable to those for Stroop tasks (Glasner, 1992). In Stroop naming, maximum inhibition occurs when the distractor is presented within a narrow temporal window from 100 ms before to 100 ms after the target. Since Stroop distractors are usually visual (but see Cowan, 1989), the necessary correction for the duration of an auditory distractor in picture-word interference tasks puts the inhibitory effects into the same temporal frame as Stroop tasks. Likewise, the magnitude of Stroop inhibition (relative to an unrelated distractor) is in the region of 30 to 100 ms depending on the nature of the target, putative picture-word interference in the same range but at the low end (unsuitable to cross-modal presentation of the distractor). To our knowledge there have been no systematic investigations of modulations of Stroop naming by phonological distractors over time, although picture-word studies consistently show facilitation from graphemically and phonologically related visual distractors (e.g., Lupker, 1962; Rayner & Springer, 1986).

Pursued systematically, these kinds of co-occurrences could serve to clarify the complex empirical picture that surrounds research on Stroop naming in word recognition and attention. The evidence from picture-word interference argues for a contribution to Stroop inhibition from within the word-production system.

In that system, there are at least two factors that can influence the magnitude of Stroop effects, stemming from competition for selection and facilitation of phonological coding for speech. There may also be a third factor, to do with the grammatical properties of distractors and targets (Elki, 1977; Micco & Caramazza, 1997; Schriefers, 1993). Although the classic Stroop effect reflects more than disruptions to word production alone, the identification and isolation of the production components of the effect may be prerequisites to an explanation of the perceptual and attentional components that have engaged so many cognitive psychologists for so many years.

**Summary**

Research on word recognition has produced a wide range of results that are relevant to explaining the cognitive mechanisms of word production, as well as to explaining word recognition proper. Although the foregoing survey barely scratched the surface of the literature, it illustrates how the processes of production can influence performance in many ostensibly perceptual paradigms. These include implicit memory tasks, tasks used for neuroimaging studies, and the traditional tasks of word recognition research, such as visual-word naming, lexical priming, and Stroop naming. In these tasks, the less salutory implications of production processing arise in the absence of adequate controls for the influence of production mechanisms, with the consequence that important experimental results may be compromised. On the positive side, the emergence of a converging body
of work on word production raises the prospects for an integrated account of the mechanisms of word recognition and production. We will examine some of these prospects in the next section.

HOW ARE WORD PRODUCTION AND RECOGNITION LINKED?

At least logically, word production and word recognition are mirror-image events. Were they mirror images psychologically, as well, the foregoing exercise in contrast and comparison might have yielded a richer picture. For all the reasons noted earlier, we did not expect to find such neatness, and so its absence is hardly dismaying. What we did find, however, supports some hypotheses about points of convergence and divergence in lexical processing that merit brief mention.

Separate pathways for word recognition and production

Writing in 1987, Monsell argued for separate input and output lexicons, with separate but linked processors for phonological features in recognition and production. This bridge served as an intermediate link between the processing of orthography and the processing of phonology. However, the point of ultimate convergence between lexical input and output was identified as a shared representation of the semantic and syntactic features of words, roughly comparable to the kinds of information that are combined during lexical selection in Figure 2.1. From this view, one would expect certain key effects to emerge from recognition and production tasks that depend on access to a word’s semantic and syntactic properties, and different effects to emerge when tasks can be accomplished without tapping information about the meaning or structural privileges of a specific word.

Some direct evidence for Monsell’s conception comes from a series of experiments by Cutting (1997). To explore the kinds of processes that are shared by word recognition and word production, Cutting used picture naming coupled with a prime-processing manipulation. To force different ways of processing priming words, subjects were instructed to repeat one word from a dichotically presented two-word prime and to ignore the second word. Words spoken in a female voice were always to be repeated, and words spoken in a male voice were always to be ignored. For example, some participants would hear the words tiger (spoken in a female voice) and sandal (spoken in a male voice) simultaneously, under instructions to repeat only the word spoken in the female voice. Other participants would hear the same two words recorded in different voices (so that they would repeat sandal). Then, 1500 ms later, both groups of subjects would see and name aloud a picture of a lion. The question was whether and how the semantic properties of the primes would influence the time taken to name the depicted object. This influence was assessed relative to the picture naming times after unrelated primes.

Because masked and otherwise unattended words are known to be processed to a level that facilitates subsequent recognition (e.g., Tipper & Driver, 1988), Cutting assumed that the ignored word would traverse perceptual/recognition processes only, whereas the produced word would be processed through the entire comprehension and production systems. The hypothesis was that when the semantic relative tigre was to be ignored, it would influence the subsequent production of lion only to the degree that comprehension processes overlap those of production. In other experiments, the phonological properties of attended and ignored primes were manipulated (e.g., liar served as a phonological prime for lion).

Cutting’s results showed a clear pattern of priming from semantically related words regardless of whether they were to be repeated or ignored. For phonological primes, however, only produced words had an impact on subsequent picture naming. The same patterns of priming were obtained when the prime words were visually rather than auditorily presented, with the cues to ignore or produce the words conveyed by ink colour. The implication is that the production and comprehension systems share semantic processes but divide phonological processes.

These findings support Monsell’s suggestions about the relationships between word production and word comprehension. They argue that recognition tasks with spoken response measures are very likely to call on information that is in the province of language production proper, and that variations in findings across traditional recognition tasks may in part be explainable in terms of which production-relevant processes come into play. In particular, if recognition and production call on different phonological processing systems, whenever recognition is operationalised in terms of spoken-word responses it is likely that some portion of what is measured reflects the processes involved in preparing to talk.

CONCLUSIONS

Talking is a goal-directed action. It is a kind of action that is central to human communication and culture, as well as a kind of action that psychologists exploit for the narrower purposes of research on cognition. Word production is in widespread use as a measure of successful perception, successful word recognition, successful memory, and much more. Although there has been little acknowledgement of the complexity of the high-level preparatory processes that eventuate in the articulation of a word, our survey suggests that these preparatory processes have a great deal in common with the components of normal word production in meaningful, connected speech. If so, much of the research that has been done to uncover the processes of word recognition may ultimately tell us at least as much about how words are retrieved for speaking as about how they are recognised.
Despite the complications that this introduces for current models of word recognition, in the long run it promises a better balanced explanation of language performance. Listeners (and readers) are also speakers, and their ability to do both of these things with their language is the foundation of successful communication. From Cattell (1886a,b) onwards, psychologists have understood that there are significant cognitive events subsequent to perception and recognition. Cattell himself was struck by how much more variability was to be found in the additional time taken for naming something ("will time" in Cattell's terminology) than in the time taken for recognizing the same object ("perception time"). A picture of a complex object took little more time to recognize than a simple colour, but considerably longer to name, even with numbers of alternatives equated. Cattell recognized that "will time" includes articulation but also considerably more; what seems to include can begin to be appreciated in light of the cognitive architecture of word production. By putting the pieces of the recognition and production puzzles together, we come closer to achieving the science of cognition that Cattell foresaw as the culmination of studies of naming in psychology.

In retrospect, Cattell's optimism about such matters raises questions about why goal-directed action in general, and word production in particular, have received scant attention in contemporary cognitive psychology. The legacy of behaviorism is surely one culprit, as we noted in the introduction. Watson's original conception of cognition embraced the vaguely mundane speculation that "thought processes are really motor habits in the larynx" (1913, p. 174), the prototype for what Bock (1996) dubbed the mind-in-the-mouth assumption. The counter-reaction to behaviorist ideas only compounded the problem: Action is behavior, so cognitive psychology shunned it and thereby overlooked the role of cognition in directing and achieving goals.

One manifestation of this general disdain for action is inattention to how people control their responses within experimental tasks. Task analysis is surely shallowing, and seems especially barren when applied to the kinds of arbitrary responses that psychologists favour. Lever-pulling and button-pressing are used as tasks precisely because of their assumed neutrality with respect to the decisions that motivate them. But recognizing the risks of misconstruing how readers carry out the laboratory tasks they are assigned, psychologists have increasingly called for more and better task analysis. This is especially true in research on word recognition (Balota, 1994; Seidenberg, 1995). Balota's work illustrates the importance of task analysis for understanding the "neglected decision stage" (Balota & Chumbley, 1984, 1985) where, one imagines, Cattell's "will time" resides along with the high-level processes of word production.

Word production is different in some fundamental ways from the simpler kinds of decisions that experimental subjects make. It is deeply embedded within an action system that is central to normal human behavior. The action system is one that we begin to use daily around the age of one and deploy thereafter during almost every waking hour of every day. Ignoring the ecological niche that words normally occupy invites unhappier surprises similar to those that ensued when psychology first confronted the influence of instinctive and prepared action systems on the behavior of laboratory rats and other animals (Hayne & Stevenson-Hinde, 1973; Marler & Terrance, 1984; Seligman & Hages, 1972). Rats easily learn to run to avoid shock; they resist learning to turn a wheel for the same purpose. Pigeons easily learn to peck a key for food; they resist learning the same action to turn off shock. In short, when natural behaviors are co-opted for the convenience of experimentation without adequate consideration of their normal function for an animal, it can be difficult to understand or fully explain the effects of an experimental manipulation.

Words have typical functions, too. They are selected for their fitness not only to the context of communication and the intended meaning, but also for the properties that allow them to be woven into connected speech. The combinational privileges of words are an integral part of what we know about them, and are crucial to their normal use. Many laboratory tasks nonetheless rely exclusively on the elicitation of single words that are rarely uttered alone. Much can be learned from this work. Much more might be gained by acknowledging the properties of words that are intrinsic to their functions in natural language, and by making allowances for how these properties complicate the journey from mind to mouth. Though logically irrelevant to naming a picture or pronouncing a word, these properties and their associated processes may be no more easily sidestepped for words in the laboratory than for words in the wild.

ACKNOWLEDGEMENTS

The preparation of this paper was supported in part by research and training grants from the National Institutes of Health (ROI HD2101) and T32 MH18990, and the National Science Foundation (PHY 94-11627, 97-14540, and a National Science Foundation Minority Graduate Fellowship). We thank Urban Chader, Todd Reising, and Hilary Silber for their assistance with the frequency and age-of-acquisition analyses, and Linda Woodland and two anonymous reviewers for their comments on a previous draft of the manuscript.

NOTES


2. Lexical (word/nonword) decision is a commonly used word-recognition task, yet its validity as a test of recognition is suspect. In particular, it is believed to be swayed by post-recognition processes brought into play to make the word/nonword discrimination, such as using apprehended relationships between primes and targets to bias "word" decisions (Balota & Chumbley, 1984). Since unrelated primes and targets by definition lack a relationship, frequency times to them may be slow for reasons that have nothing to do with the recognition process itself, and related primes fast, for
REFERENCES


46 BOCK AND CRIPPIN


2. HOW MIND MEETS MOUTH 47


