Models of word production

Willem J.M. Levelt

Research on spoken word production has been approached from two angles. In one research tradition, the analysis of spontaneous or induced speech errors led to models that can account for speech error distributions. In another tradition, the measurement of picture naming latencies led to chronometric models accounting for distributions of reaction times in word production. Both kinds of models are, however, dealing with the same underlying processes: (1) the speaker’s selection of a word that is semantically and syntactically appropriate; (2) the retrieval of the word’s phonological properties; (3) the rapid syllabification of the word in context; and (4) the preparation of the corresponding articulatory gestures. Models of both traditions explain these processes in terms of activation spreading through a localist, symbolic network. By and large, they share the main levels of representation: conceptual/semantic, syntactic, phonological and phonetic. They differ in various details, such as the amount of cascading and feedback in the network. These research traditions have begun to merge in recent years, leading to highly constructive experimentation. Currently, they are like two similar knives honing each other. A single pair of scissors is in the making.

How do we generate spoken words? This issue is a fascinating one. In normal fluent conversation we produce two to three words per second, which amounts to about four syllables and ten or twelve phonemes per second. These words are continuously selected from a huge repository, the mental lexicon, which contains at least 50–100 thousand words in a normal, literate adult person. Even so, the high speed and complexity of word production does not seem to make it particularly error-prone. We err, on average, no more than once or twice in 1000 words. This robustness no doubt has a biological basis; we are born talkers. But in addition, there is virtually no other skill we exercise as much as word production. In no more than 40 minutes of talking a day, we will have produced some 50 million word tokens by the time we reach adulthood.

The systematic study of word production began in the late 1960s, when psycholinguists started collecting and analyzing corpora of spontaneous speech errors (see Box 1). The first theoretical models were designed to account for the patterns of verbal slips observed in these corpora. In a parallel but initially independent development, psycholinguists adopted an already existing chronometric approach to word production (Box 1). Their first models were designed to account for the distribution of picture naming latencies obtained under various experimental conditions. Both approaches are happily merging in current theorizing, all existing models have a dominant kiss-like: their ancestry is in either speech error analysis or it is in chronometry. In spite of this dual perspective, there is a general agreement on the processes to be modeled. Producing words is a core part of producing utterances; explaining word production is part of explaining utterance production.

All current models of word production are network models of some kind. In producing an utterance, we go from some communicative intention to a decision about what information to express – the ‘message’. The message contains one or more concepts for which we have words in our lexicon, and these words have to be retrieved. They have syntactic properties, such as being a noun or a transitive verb, which we use in planning the sentence, that is in ‘grammatical encoding’. These syntactic properties taken together, we call the word’s ‘lemma’. Words also have morphological and phonological properties that we use in preparing their syllabification and prosody, that is in ‘phonological encoding’. Ultimately, we must prepare the articulatory gestures for each of these syllables, words and phrases in the utterance. The execution of these gestures is the only overt part of the entire process.

Two kinds of model

All current models of word production are network models of some kind. In addition, they are, with one exception, all ‘localist’, non-distributed models. That means that their
Box 1. Historical roots of word production research

The study of word production has two historical roots, one in speech error analysis and one in chronometric studies of naming.

The speech error tradition

In 1885, Cattell published a substantial corpus of German speech error that they had diligently collected (Ref. a). The corpus, along with the theoretical analysis they provided, inspired subsequent studies and research traditions. One important distinction they made was between meaning-based substitutions (such as *bit* ‘(piece)’ for *note* ‘(note)’) and form-based submissions (such as *bun* ‘(bun)’ for *bun* ‘(bun)’), acknowledging that there is often a phonological connection in meaning-based errors (i.e. the over-representation of mixed errors was practically all meaning-driven). Why does a patient say of her parents that they have Geiz (‘good’) instead of Geiz (‘cleverness’)? Because she had suppressed her real opinion about her parents – oh, all the errors we would make! A second, now classical distinction that Meringer and Mayer introduced was between exchanges (e.g. *wolf* instead of *wolf*), anticipations (false starts for false starts), perseverations (here away for here away) and blends or contaminations (sewed, blending avoid and evade).

Many linguists and psychologists have continued this tradition (Ref. b), but an eldritch resonance (probably triggered by the work of Cohen; Ref. c) began in the late 1960s. In 1973, Fromkin edited an influential volume of speech error studies, with part of her own collection of errors as an appendix (Ref. d). Another substantial corpus was built up during the 1970s, the MIT-CU corpus: It led to two of the most influential models of speech production: (i) Garrett discovered that word exchanges (such as *left* and *right*) can span some distance and mostly preserve grammatical category as well as grammatical function within their clauses (Ref. e). Sound-form exchanges (such as *pack* for *padd*); on the other hand, ignore grammatical category and preferably happen between close-by words. This indicates the existence of two modular levels of processing in sentence production, a level where syntactic functions are assigned and a level where the ordering of forms (morpheme, phoneme) is organized. (2) Shattuck-Hufnagel’s scan-copier model captured phonological encoding (Ref. f). A core notion here is the existence of phonological frames, in particular syllable frames. Sound errors tend to preserve syllable position (as in the case in pack, or in a papa smoker for pipe smoker). The model claims that a word’s phonemes are retrieved from the lexicon with their syllable position specified. They can only land in the corresponding slot of a syllable frame.

In 1976, Baars, Merkey and MacKay (Ref. g) developed a method for eliciting speech errors under experimentally controlled conditions, two years after Brown and McNeill had created the paradigm of eliciting speech errors by claiming that innocent form errors are practically all naming errors. There is, for instance, difficult to name the word word when it is written in red. But naming the word was not affected by the word’s color. Rosinski et al. interested in the automatic word reading skills of children, transformed the Stroop task into a picture/word interference task (Ref. h). The children named a list of object drawings. The drawings contained a printed word that was to be ignored. Alternatively, the children had to name the printed words, ignoring the objects. Object naming suffered much more from a semantically related interfering word than word naming suffered from a meaning-related interfering object, confirming the pattern typically obtained in the Stroop task. Lupker set out to study the differences between line drawings and words. Finseid showed that when a small circle was named as ‘circle’ took, on average, 619 ms, but when named as a ‘rat’ took, 1450 ms. He also discovered that different codas to be accessed. They are not graphemic codes, because Potter et al. obtained the same picture-word difference in Chinese (Ref. i). The dominant current view is that there is a direct access route from the word to its phonological code, whereas the line drawing first activates the object concept, which in turn causes the activation of the phonological code – an extra step. Another classical discovery in the picture-naming tradition (by Oldfield and Wingfield; Ref. j) is the word frequency effect (see main article).

In 1955, Stroop introduced a new roteic paradigm, now called the ‘Stroop task’ (Ref. k). The stimuli are differently colored words. The subject’s task is either to name the color or to say the word. Stroop studied what happened if the word was a color name itself. The main finding was that color naming is substantially slowed down when the colored word is a different color name. Now, for instance, difficult to name the word green when it is written in red. But naming the word was not affected by the word’s color. Rossin the automatic word reading skills of children, transformed the Stroop task into a picture/word interference task (Ref. h). The children named a list of object drawings. The drawings contained a printed word that was to be ignored. Alternatively, the children had to name the printed words, ignoring the objects. Object naming suffered much more from a semantically related interfering word than word naming suffered from a meaning-related interfering object, confirming the pattern typically obtained in the Stroop task. Lupker set out to study the differences between line drawings and words. Finseid showed that when a small circle was named as ‘circle’ took, on average, 619 ms, but when named as a ‘rat’ took, 1450 ms. He also discovered that different codas to be accessed. They are not graphemic codes, because Potter et al. obtained the same picture-word difference in Chinese (Ref. i). The dominant current view is that there is a direct access route from the word to its phonological code, whereas the line drawing first activates the object concept, which in turn causes the activation of the phonological code – an extra step. Another classical discovery in the picture-naming tradition (by Oldfield and Wingfield; Ref. j) is the word frequency effect (see main article).
to the subject at different SOAs with respect to picture onset. The distractor words were either semantically or phonologically related to the target word, or unrelated. This paradigm and its many later variants made it possible to study the relative time courses of the target name's semantic and phonological encoding in much detail.

References


Glasers, F.J. (1984) Logical encoding. Only Roschels WEAVER model15,16 has a fully developed phonological component. A fragment of the WEAVER lexical network is shown in Fig. 2.
The main strata in this network are the same as those in the interactive model. There is a conceptual/semantic level of nodes, a lemma stratum and a phonological or form stratum. But the model is only partially interactive. There are good reasons for assuming that conceptual and lemma strata are shared between production and perception, hence their interconnections are modelled as bi-directional. But the form stratum is unique to word production; it does not feed back to the lemma stratum. Therefore it is often called the discrete (as opposed to ‘interactive’) two-step model. Although the model was designed to account for response latencies, not for speech errors, the issue of ‘mixed’ speech errors cannot be ignored and it has not been. The explanation is largely post-lexical. We can strategically monitor our internal phonological output and intercept potential errors. A phonological error that happens to create a word of the right semantic domain (such as rat for cat) will have a better chance of ‘slipping through’ the monitor than one that is semantically totally out of place (such as mat for rat). Similarly, an error that produces a real word will get through easier than one that produces a non-word. There is experimental evidence that the monitor is indeed under strategic control. Still, the causation of mixed errors continues to be a controversial issue among models of word production.

Conceptual preparation

The first step in accessing content words such as cat or select is the activation of a lexical concept, a concept for which you have a word or morpheme in your lexicon. Usually, such a concept is part of a larger message, but even in the simple case of naming a single object it is not trivial which lexical concept you should activate to refer to that object. It will depend on the discourse context whether it will be more effective for you to refer to a cat as cat, animal, siamese or anything else. Rosch has shown that we prefer ‘basic level’ terms to refer to objects (car rather than automobile, dog rather than corgi, etc.), but the choice is ultimately dependent on the perspective you decide to take on the referent for your interlocutor. Will it be more effective for me to refer to my sister as my sister or as that lady or as the physicist? It will all depend on shared knowledge and discourse context. This freedom of perspective-taking appears quite early in life and is ubiquitous in conversation.
Working models of word production begin where perspective-taking ends: at the activation of a target concept to be expressed. The representation of a target concept, however, varies among models. The two preferred variants are just the ones exemplified in Figs 1 and 2. Concepts are either represented as decomposed, or as non-decomposed or ‘whole’. The issue is controversial, but arguments have been accumulating for using whole-concept representations in models of word production. One argument is the so-called ‘hyperonym problem’. If you activate some set of semantic features as a representation of the notion ‘cat’, the notion ‘animal’ will involve a proper subset of these features. Hence, it is indeterminate which of the two will ultimately be expressed. This is not an advantage: hyperonym speech errors are rare in any case and you need extra machinery to prevent the hyperonym problem from arising.
Lexical selection

In the chronometric tradition lexical selection has been studied with interference paradigms, in particular picture-word interference (see Box 1). The recurring finding has been that naming an object is slowed down when a distracter word is presented with the picture; the effect is stronger when the distracter word is semantically related to the target than when it is semantically unrelated and it is at maximum when picture and distracter word are presented simultaneously. The WEAKER model provides an accurate quantitative account of a wide range of picture-word interference data, with only a few free parameters. How does it work? When you are naming a picture of a sheep and you decide to go for the basic level term, you will activate the lexical concept sheep as your target and activation spreads to the corresponding lemmas. In the semantic network activation spreads to related concepts, such as goat and llama. They, in turn, spread activation to their lemmas. During any unit time interval the probability of selecting the target lemma sheep from the mental lexicon is the ratio of that lemma’s degree of activation and the total activation of all lemmas (including goat, llama and sheep). This is called Luce’s ratio, and it allows for the computation of an expected selection latency. In other words, there is competition between semantically related lemmas. Active alternatives slow down the selection process (even though a special checking mechanism in WEAKER normally prevents them from replacing the target). If you present the semantically related word ‘goat’ as a distracter, the already co-activated lemma goat will receive an additional boost, thereby becoming a strong competitor to sheep. By contrast, if you present a semantically unrelated word, such as ‘chair’, as distracter, there will be no convergence of activation and, correspondingly, competition will be relatively weak. That explains the semantic-inhibition effect.

Activation spreading through a semantic network (of whatever type) is also the obvious explanation for semantic naming errors, the dominant speech error type (about two-thirds of errors in a normal picture naming task are semantic in character). But what is a semantic error? A particular choice of words may have its cause in perspective-taking. If a speaker decides to name a depicted dog as an ‘animal’ or a ‘collie’, that may well be an intentional act rather than an error. There is a substantial literature on the types of semantic (and other) errors produced by aphasic patients, which will not be covered in the present review. It is a major challenge to predict these error distributions by ‘damaging’ the normal network. Dell et al. have set an impressive example. They successfully modeled the naming errors (semantic and other) of a diverse set of aphasic patterns by manipulating no more than two parameters in their interactive two-step model: the weight on the network connections and the decay rate of the nodes’ activation.

The timing of lexical selection is not explicitly modeled in the speech-error based models. In the interactive two-step model the selection moment is determined from outside. When you produce a sentence, the moment of selecting the most activated lemma is dictated by when it is to be inserted in the grammatical frame. The selection moment is usually given a constant default value in modeling error distributions.

Both whole-concept and featural representations allow for precise semantic referencing (of the type ‘a dog is an animal’), but this inferential potential plays no role in the factual word production process.
Box 2. Implicit priming

The method of implicit priming was introduced by Meyer to study the time course of phonological encoding, that is the speaker’s construction of a spoken word’s form (Ref. 3). The initial and major discovery, which has been repeatedly confirmed, was that a word’s form is built up incrementally, starting with the first segment. Apparently, phonological word shapes do not come as whole templates; rather they are generated afresh, time and again, from beginning to end.

The method is exemplified in Table 1. Subjects learn a set of three semantic–word-associations (A–B); for instance set 1 in the leftmost column. Then, an A-word from the set appears on the screen and the subject produces the corresponding B-word as fast as possible. The word onset latency is measured by voice key. The A-words from the set are repeatedly presented in random order and at each trial the naming latency of the B-word is registered. Then the subject is presented with set 2, the triple in the second column of the table below, and the same procedure is run for that set. Finally, set 3 is run in the same way.

The response-words in a set share a phonological property. The B-words in set 1 are lower, hot, and time; they share the initial syllable /l/. Similarly, the B-words in set 2 share the initial syllable /m/; and those in set 3 share the initial syllable /l/. Such sets sharing a phonological property are called ‘homogeneous’ and the shared property is called the ‘implicit prime’.

Can the subject use this implicit prime when running through the set? Whether the subject can prepare for the first syllable of the response word can be tested by comparing the homogeneous condition with a heterogeneous condition; that is, one in which there is no implicit prime. The heterogeneous condition is created by confounding the A-B-pairs in such a way that they no longer share the first syllable. For instance, the first set of the heterogeneous condition (fourth column in the table) has loser, brown, and major as response words. Each word pair is its own control in the experiment: it appears in both the homogeneous and the heterogeneous condition.

In the homogeneous condition there is no implicit prime, hence the subject cannot prepare anything. When Meyer did the experiment exemplified in Table 1 (in Dutch), he found that response latencies were significantly shorter in the homogeneous condition than in the heterogeneous condition. Apparently, subjects can prepare for the response word’s first syllable. If prosodification is really incremental (i.e. starting at the beginning of the word), subjects should not be able to prepare for the second syllable of a bisyllabic word. And indeed, a not using response words ending in the same syllable, such as murder, pound, build, showed no implicit priming whatever. Generally, there was always implicit priming for words that shared any beginning part of the word, but never for words sharing any final part, nor even for monosyllabic rhymes, such as dog, feed, seed. In addition, the longer the shared word-beginningstretch, the stronger the priming.

These robust findings have led to the suggestion that, normally, the speaker does not instantiate articulation before the whole word has been encoded. If encoding is incremental, which is now well-established, this should predict a robust word-length effect. But this is rarely obtained (Ref. 3). It is unknown under which conditions a speaker does complete a word’s phonological encoding before initiating speech.

Implicit priming can also be used to test whether a subject must know which syllable to stress in prosodification. For instance, in Table 1 all response words have the same metrical shape: they are all first-syllable-stressed. In that case, there is strong implicit priming. But will there be implicit priming if the response words do not share stress position? Roelofs and Meyer tested this in Dutch, using sets of response words such as: ma-tre-rie – ‘make’/ ma-de-lief – ‘makeup’/ ma-ne-ri-ra – ‘manuscript’/ ma-wl-de-eb – ‘dairy’ (Ref. 4). Here two words in the set have second syllable stress and two have third syllable stress. There was not the slightest hint of implicit priming in this condition of variable metrical stress. This shows that you cannot prepare for the first syllable if you don’t know where the word’s stress is to go – in order to prepare you must know the word’s ‘metrical frame’. References


Table 1. The implicit priming method: priming the first syllable of bisyllabic words

<table>
<thead>
<tr>
<th>Homogeneous condition</th>
<th></th>
<th>Heterogeneous condition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Set 1</strong></td>
<td><strong>Set 2</strong></td>
<td><strong>Set 3</strong></td>
</tr>
<tr>
<td>single-toner</td>
<td>signal-beacon</td>
<td>captain-major</td>
</tr>
<tr>
<td>place-local</td>
<td>priest-beadle</td>
<td>cardis-maker</td>
</tr>
<tr>
<td>fruit-lotus</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Morpho-phonological encoding

When you are planning the sentence ‘they are selecting me’, you must retrieve from your lexicon the two morpheme-size codes select and ing (see Fig. 5), and compute their syllabification and accent structure in context (see-ing-ing). This naturally divides the process into ‘code retrieval’ and ‘prosodification’.

Code retrieval

An item’s morpho-phonological code consists of its morphological make-up, its metrical shape and its segmental make-up (see Fig. 3, Step 1 and Step 2). Retrieving that information follows activation/seletion of the lemma. Much ink and many subjects have been spilled over this issue. In the WEAVER model, the activation and retrieval of a phonological code is strictly conditional on selecting the corresponding lemma. For instance, when your target word is eat, you first select its lemma and only then spread activation to its phonological code (cat). This predicts that alternative active, but non-selected lemmas (such as the lemma for age) do not spread any activation to their phonological codes. Initial experimental evidence showed that, in picture naming, there is semantic but not morpho-phonological activation of same-category alternatives (if cat is the target, dog is semantically but not phonologically active).

If prosodification is really incremental (i.e. starting at the beginning of the word), subjects should not be able to prepare for the second syllable of a bisyllabic word. And indeed, a not using response words ending in the same syllable, such as murder, pound, build, showed no implicit priming whatever. Generally, there was always implicit priming for words that shared any beginning part of the word, but never for words sharing any final part, nor even for monosyllabic rhymes, such as dog, feed, seed. In addition, the longer the shared word-beginning stretch, the stronger the priming.

These robust findings have led to the suggestion that, normally, the speaker does not instantiate articulation before the whole word has been encoded. If encoding is incremental, which is now well-established, this should predict a robust word-length effect. But this is rarely obtained (Ref. 3). It is unknown under which conditions a speaker does complete a word’s phonological encoding before initiating speech.

Implicit priming can also be used to test whether a subject must know which syllable to stress in prosodification. For instance, in Table 1 all response words have the same metrical shape: they are all first-syllable-stressed. In that case, there is strong implicit priming. But will there be implicit priming if the response words do not share stress positions? Roelofs and Meyer tested this in Dutch, using sets of response words such as: ma-tre-rie – ‘make’; ma-de-lief – ‘makeup’; ma-ne-ri-ra – ‘manuscript’; ma-wl-de-eb – ‘dairy’ (Ref. 4). Here two words in the set have second syllable stress and two have third syllable stress. There was not the slightest hint of implicit priming in this condition of variable metrical stress. This shows that you cannot prepare for the first syllable if you don’t know where the word’s stress is to go – in order to prepare you must know the word’s ‘metrical frame’. References


Table 1. The implicit priming method: priming the first syllable of bisyllabic words

<table>
<thead>
<tr>
<th>Homogeneous condition</th>
<th></th>
<th>Heterogeneous condition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Set 1</strong></td>
<td><strong>Set 2</strong></td>
<td><strong>Set 3</strong></td>
</tr>
<tr>
<td>single-toner</td>
<td>signal-beacon</td>
<td>captain-major</td>
</tr>
<tr>
<td>place-local</td>
<td>priest-beadle</td>
<td>cardis-maker</td>
</tr>
<tr>
<td>fruit-lotus</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Morpho-phonological encoding

When you are planning the sentence ‘they are selecting me’, you must retrieve from your lexicon the two morpheme-size codes select and ing (see Fig. 5), and compute their syllabification and accent structure in context (see-ing-ing). This naturally divides the process into ‘code retrieval’ and ‘prosodification’.

Code retrieval

An item’s morpho-phonological code consists of its morphological make-up, its metrical shape and its segmental make-up (see Fig. 3, Step 1 and Step 2). Retrieving that information follows activation/seletion of the lemma. Much ink and many subjects have been spilled over this issue. In the WEAVER model, the activation and retrieval of a phonological code is strictly conditional on selecting the corresponding lemma. For instance, when your target word is eat, you first select its lemma and only then spread activation to its phonological code (cat). This predicts that alternative active, but non-selected lemmas (such as the lemma for age) do not spread any activation to their phonological codes. Initial experimental evidence showed that, in picture naming, there is semantic but not morpho-phonological activation of same-category alternatives (if cat is the target, dog is semantically but not phonologically active).
Outstanding questions

- How should error-based and chronometric models be further reconciled computationally and empirically?
- What causes a speech error? Is it caused by occasional cascading or occasional feedback in a normally non-cascading, feed-forward system? Is it the product of noise in a normally cascading interactive system? Or is the origin of speech error something else entirely?
- How does the word-production network relate to the word-perception network? How is self-monitoring realized in this combined system?
- How are syllabic and larger gestures computed from a syllabified phonological code? Is there anything like a repository of syllabic gestural scores?
- If phonological word encoding is an incremental process, why is it that naming a short word is harder faster than naming a long word?
- Which brain regions subserve the core components of conceptual/semantic preparation, lexical selection, phonological code retrieval, prosodicification, phonetic encoding, articulation and self-monitoring?
word stress goes to the first full-voweled syllable (morning, yellow, forget – the ‘-s’ in the latter word is not full-voweled, but rather a neutral ‘-sshw’-sound). This can be automatically produced in incremental syllabification. But when a word has a deviant stress pattern, the automaticity breaks down17,18 (see Box 2 for an example). A word’s deviant metrical frame is probably stored as part of its phonological code; it guides the deviant prosodification. Languages differ, however, in their default metrics.

The distinction between accessing a word’s phonological code and in subsequent rapid syllabification is crucial for understanding the neural architecture of word production. A meta-analysis of imaging studies in word production19 suggests that accessing the code involves Wernicke’s area, whereas prosodification involves the posterior inferior frontal cortex.

Phonetic encoding and articulation

As incremental prosodification proceeds, the resulting syllabic and larger prosodic structures should acquire phonetic shape. As a speaker you will incrementally prepare articulatory gestures for the syllables in its prosodic content. A core feature of the WEAVER model is the notion of a syllabary20. Statistics show that native speakers of English or Dutch do 80 percent of their talking with no more than about 500 different syllables21 (although these languages have many more than 10 000 different syllables). The syllabary is postulated as a repository of such overused, high-frequency syllabic gestures, one ‘syllabic score’ for each. Each time a new phonological syllable, such as ‘hit’ (/hɪt/), or ‘hit’ in the latter word is not full-voweled, is composed, the corresponding gestural score is triggered. The score specifies which motor tasks (such as closing the glottis or releasing lip closure) are to be performed21 in order to generate the syllable. In WEAVER there is always competition among gestural scores. The activation spreads from individual segments to all syllabic scores in which they participate (see Fig. 2). Hence, similar syllabic scores tend to be co-activated. The occasional non-selection will resemble the target gesture. Selection latency is determined by Luce’s rule (as in the case for lemma selection).

There are further restrictions in selecting a syllabic score for execution. Repeat use of a particular type of syllable, for instance in producing the nonsense phrase dog-of/der (where dog and the following of/der are both CVCC syllables), may facilitate articulation22. General scores of similar types (such as CV–CV or CYC) can apparently co-activate one another. Finally, WEAVER and the two-step interactive model have a structural representation of each segment. In both models the units of phonological encoding are whole phonemes (for which there is good experimental evidence23), but their features, such as ‘voiced’, ‘nasal’, ‘sonorous’, are already ‘visible’ to the process of syllabification (see legend to Fig. 2). During the next stage, phonetic encoding, these features function in the construction of articulatory gestures. The study of speech movement planning has become a discipline of its own24,25 and is not covered in the present review.

Conclusion

There is still a long way to go before the two research traditions emerging from speech error analysis and from naming chronometry are fully reconciled. But there has been lively and highly constructive interaction, leading to a much improved understanding of the processes involved in lexical selection and phonological encoding. One unifying force has been computational modeling. Current implemented models share their major strands, they are localist and symbolic; they compete with similar linguistic representations. Another unifying force will hopefully proceed from brain imaging (see Ref. 57 for a recent review of imaging studies of word processing). It is in the processing models that should guide the design of brain imaging experiments in word production, not native intuition as it is too often the case25. The return will be convergence of evidence for or against particular processing components and their interactions.

References

Gary Dell. I gratefully acknowledge helpful commentary by Antje Meyer and by Gary Dell. Therefore, its mechanisms may be comprehensively addressed. However, the organization provided by this approach is likely to be insufficient for understanding the complex and dynamic processes involved in word production. In this article, I summarize current theoretical approaches to word production and outline the following

Acknowledgement

There is still a long way to go before the two research traditions emerging from speech error analysis and from naming chronometry are fully reconciled. But there has been lively and highly constructive interaction, leading to a much improved understanding of the processes involved in lexical selection and phonological encoding. One unifying force has been computational modeling. Current implemented models share their major strands, they are localist and symbolic; they compete with similar linguistic representations. Another unifying force will hopefully proceed from brain imaging (see Ref. 57 for a recent review of imaging studies of word processing). It is in the processing models that should guide the design of brain imaging experiments in word production, not native intuition as it is too often the case25. The return will be convergence of evidence for or against particular processing components and their interactions.

References

I gratefully acknowledge helpful commentary by Antje Meyer and by Gary Dell. Therefore, its mechanisms may be comprehensively addressed. However, the organization provided by this approach is likely to be insufficient for understanding the complex and dynamic processes involved in word production. In this article, I summarize current theoretical approaches to word production and outline the following

Acknowledgement

There is still a long way to go before the two research traditions emerging from speech error analysis and from naming chronometry are fully reconciled. But there has been lively and highly constructive interaction, leading to a much improved understanding of the processes involved in lexical selection and phonological encoding. One unifying force has been computational modeling. Current implemented models share their major strands, they are localist and symbolic; they compete with similar linguistic representations. Another unifying force will hopefully proceed from brain imaging (see Ref. 57 for a recent review of imaging studies of word processing). It is in the processing models that should guide the design of brain imaging experiments in word production, not native intuition as it is too often the case25. The return will be convergence of evidence for or against particular processing components and their interactions.

References

I gratefully acknowledge helpful commentary by Antje Meyer and by Gary Dell. Therefore, its mechanisms may be comprehensively addressed. However, the organization provided by this approach is likely to be insufficient for understanding the complex and dynamic processes involved in word production. In this article, I summarize current theoretical approaches to word production and outline the following

Acknowledgement

There is still a long way to go before the two research traditions emerging from speech error analysis and from naming chronometry are fully reconciled. But there has been lively and highly constructive interaction, leading to a much improved understanding of the processes involved in lexical selection and phonological encoding. One unifying force has been computational modeling. Current implemented models share their major strands, they are localist and symbolic; they compete with similar linguistic representations. Another unifying force will hopefully proceed from brain imaging (see Ref. 57 for a recent review of imaging studies of word processing). It is in the processing models that should guide the design of brain imaging experiments in word production, not native intuition as it is too often the case25. The return will be convergence of evidence for or against particular processing components and their interactions.