MIT Press • Open Encyclopedia of Cognitive Science

Language Production

Zenzi M. Griffin[1]

[1]The University of Texas at Austin

MIT Press

Published on: Dec 04, 2024

DOI: https://doi.org/10.21428/e2759450.b076f599

Speakers are typically aware of the ideas they want to express but not the steps involved in getting from those ideas to a series of motor movements. Consider describing an image with the sentence, “Bees are stinging a man.” The idea that starts the production processes contains no order between the semantic representations of BEES, MAN, and STING, just their roles in the event. The BEES are the agents who act, and the MAN is the patient being acted on. The speaker must decide the order in which to express them. The semantic representation of MAN may be chosen to start the sentence because speakers prefer to mention humans before nonhumans despite BEES performing the action. Word order or grammatical marking must indicate that it is the BEES who sting the MAN and not vice versa. The expression of these concepts follows the syntax of the language. In English, the action STING follows the subject of the sentence BEES, whereas it would come at the end of the sentence in Japanese. The verbs expressing STING need to agree in number with the word expressing BEES (“are stinging” vs. “is”). In addition, the speaker must decide on the words to use. For example, the representation of MAN could be expressed as “guy,” “dude,” “victim,” or another context-appropriate word. The speaker must also decide whether to use “a” vs. “the” before the selected nouns. The sounds of the words are retrieved and organized before they can be converted into motor movements. Language production research is concerned with studying these steps.

History

Much can be learned by examining what happens to a system when it falls apart. The roots of language production research are often attributed to Merringer and Meyer’s analysis of spontaneous speech errors by adult speakers (as cited in Fromkin, 1971). The late 1960s and early 1970s saw a boom in speech error research (e.g., Fromkin, 1971; Garrett, 1975; Nooteboom, 1973), which produced the first influential theories outlining the stages of language production. Scholars also created methods for inducing errors in the lab (Baars et al., 1975). This control over content allowed researchers to manipulate variables that were correlated with errors in spontaneous speech to uncover causation. This work was fundamental in establishing the processing steps and building blocks involved in planning to speak.

Starting in the 1980s, much experimental work concerned how speakers decide the syntactic structure of their sentences [see Psycholinguistics]. Researchers often ask participants to describe line drawings of scenes (e.g., bees stinging a man). Presenting pictures allows researchers to specify a message that constrains what is said across participants without giving them specific words or syntactic structures to use. For example, studies showed that people prefer to mention humans at the start of sentences even if they are not agents (e.g., “A man is stung by bees”; Sridhar, 1988).

As computers became widely available, researchers increasingly measured the time from the onset of a picture of an object (e.g., a bed) on a computer screen to when a participant begins to articulate its name (a.k.a., *naming latency*). In picture-naming experiments, participants are asked to name objects or actions as quickly and accurately as possible. They find, for example, that common words (e.g., “bed”) are

produced faster and more accurately than uncommon ones (“bell”; the *word frequency effect*—e.g., Oldfield & Wingfield, 1964). Producing object and action names often takes over 900 ms when participants are not repeating or pretrained on their names (e.g., Szekely et al., 2005). Under 150 ms of this time is typically devoted to recognizing the familiar objects or actions depicted (Potter, 1975). Instead, most of this time is devoted to the many steps of word production.

In 1989, Wilhelm Levelt published the seminal book *Speaking* (Levelt, 1989). In it, he exhaustively reviewed work on language production and established much of the terminology and theoretical framework for the future work described below.

Studies measuring response latencies to produce simple sentences sometimes found that manipulating the difficulty of words at the end of an utterance (as opposed to the beginning) did not delay speech onset (e.g., Schriefers et al., 1998). This supported the argument that fluent sentence production can be *incremental*, meaning the speaker prepares upcoming parts of an utterance while articulating earlier parts. Around the start of the new millennium, researchers started monitoring speakers’ eye movements as they described pictures (Griffin & Bock, 2000; Meyer et al., 1998). This method goes beyond response latencies by providing direct insight into what happens *before* and *after* the articulation of an utterance begins. Although speakers can recognize what is depicted in simple line drawings within tenths of a second without moving their eyes (e.g., Bock et al., 2003), they nonetheless typically gaze at the objects they mention just before saying their names (Griffin & Bock, 2000). For example, when describing bees stinging a man, a speaker looks at the man for about a second before saying “A man.” While articulating “A man is stung by,” the speaker looks at the bees for about a second before saying “bees.” The time spent looking at the objects corresponds to the time needed to name the objects on their own and reflects how difficult it is to retrieve their names (Meyer et al., 1998).

Language production research has historically focused on spoken language. A growing body of research demonstrates many similarities between spoken and signed language production despite the difference in modality (Emmorey, 2023) [see Sign Language]. Acquiring a native spoken or signed language differs from learning a written language in that they do not need to be explicitly taught. In addition, speaking and signing typically occur with the time pressure of saying something in the moment to an interlocutor without the opportunity for extensive consideration or editing as found in writing. Despite differing from spoken and signed production in many ways, several basic findings such as frequency effects replicate with writing (e.g., Bonin et al., 2002). This article focuses on spoken languages.

Core concepts

Message planning

Speaking starts with an intention to express something (Fromkin, 1971; Levelt, 1989). The content is first established with *macroplanning*, which consists of creating a goal (e.g., convincing a friend to watch a particular movie) and the conceptual content associated with it (e.g., pointing out the strength of its

reviews, the skill of its lead actor, and the fact that it is only playing this week). Because all of this information cannot be squeezed into the production system at once, the content for an utterance is packaged into a *preverbal message* during a *microplanning* stage. A message can correspond to a proposition or two (e.g., that the movie is only playing this week) that is used to generate a sentence or a single concept like naming an object ("chair").

Messages are shaped by what the speaker thinks the listener knows or can infer (*audience design* based on *common ground;* see Yoon & Brown, 2023). For example, messages include information about whether an element has been discussed recently in a conversation, allowing a pronoun like "she" or "it" to be used instead of "the actor" or "the movie."

Grammatical encoding

Messages need to be converted into sequences of words that express the relationships between the parts of each message. Across languages, there is a tendency to express previously mentioned and human referents early in sentences (e.g., start with "man" rather than "bees"; Sridhar, 1988).

Aside from message content, another influence on word order is the structure of sentences that speakers have recently said or heard. Even when no words overlap between sentences, people are more likely to say a sentence with the same structure as one they have recently heard (Bock, 1986). For example, after hearing a *passive* sentence like, "A passerby was jostled by a drunk," in which the patient precedes the agent, a speaker is more likely to describe a picture as "A man is being stung by bees," which has the same sequence than if they heard the active form, "A drunk jostled a passerby." This phenomenon has several labels but is usually called *syntactic priming* or *structural persistence*. It occurs in natural conversations as well as the lab. This phenomenon suggests that there are procedures for forming sentences that are independent of the words used and that are changed by experience.

Two-step word retrieval: Lemma selection and phonological encoding

Speech errors throw light on the structure of word representations. The most common type of word error involves substituting an intended word with one that has a similar meaning in context ("couch" for "chair"; see Dell et al., 1997). The selection of a word based on meaning can go awry while correctly retrieving the sounds of the intruding word (e.g., substituting "couch" for "chair" but pronouncing "couch" correctly). Phonological errors involve the correct word being selected based on meaning but sound retrieval goes awry (saying "bear" for "chair"). This suggests that there is a step of production primarily concerned with selecting words based on meaning rather than sound and a subsequent one in which sounds are retrieved without much concern for meaning (Fromkin, 1971; Nooteboom, 1973). If retrieval happened in one step, one would not expect errors to be divided as cleanly into semantic and phonological types as they are. Errors can occur at either stage, and in unimpaired speakers, they involve near misses in that the result is semantically or phonologically related to the intended word.

Further support for this separation of word and sound stages comes from the *tip-of-the-tongue* phenomenon, in which a speaker has a strong sense of knowing which word they want to say but cannot come up with all the sounds (signers experience the analogous *tip-of-the-fingers* state; see Emmorey, 2023). This occurs most often for proper names (e.g., "Who directed Pink Flamingos?") and uncommon words ("What do you call a word that is spelled the same forward and backward?"; Burke et al., 1991). People in a tip-of-the-tongue state usually know syntactic information about the missing word such as its grammatical gender in languages that have it (e.g., Vigliocco et al., 1997). So, an Italian speaker failing to retrieve the sounds for the word "pen" would nonetheless know that it has feminine grammatical gender and thus would be preceded by "la" instead of "il." This word-specific knowledge indicates that speakers in a tip-of-the-tongue state have not just identified the intended concept. People in tip-of-the-tongue states sometimes have partial phonological information about the intended word, such as the first sound and number of syllables (Brown & McNeill, 1966). Providing additional semantic information about the intended word provides little or no benefit, whereas providing sounds is very helpful (Meyer & Bock, 1992). As a result, theories posit a word representation called a *lemma*, which marks that there is a word in a speaker's vocabulary that expresses a particular meaning and has specific grammatical properties but does not contain information about the sounds of the word (Kempen & Hoenkamp, 1987). Speakers first select a lemma and then retrieve and organize its sounds (*phonological encoding*).

Many phenomena indicate that the lemmas for semantically related words compete with each other for selection. One way to envision this is that putting a concept into words involves activating a set of semantic features (Fromkin, 1971). For example, the concept CHAIR could be represented with features such as IS FURNITURE, HAS LEGS, IS MANUFACTURED, HOLDS ONE PERSON, and USED FOR SITTING. While some features are relatively specific to chairs, others are shared with other objects such as couches, stools, and tables. The activated features spread activation to corresponding lemmas (e.g., Dell et al., 1997). So, the "chair" lemma will be highly activated but so will lemmas for "couch," "stool," and "table." The most highly activated lemma is selected. Due to noisy activation or residual activation from recent use, another highly active lemma may be selected by mistake. Lemmas that do not share semantic features with "chair" receive no activation from its features and are therefore unlikely to be accidentally selected. Activation of semantically related lemmas increases the time needed to produce a word (Levelt, 1989). For example, producing a semantically related word increases the time to name a picture (e.g., "chair" is slower after "couch" compared to after "lime"; e.g., Damian & Als, 2005).

One of the biggest influences on picture-naming latencies is a variable called *picture name agreement* (see also *uncertainty* and *codability*; e.g., Szekely et al., 2005). The more words that are considered appropriate names for an object or action, the longer it takes to name it. For example, a picture of a baby might be named "baby" by speakers 99% of the time, even though it could also be called an "infant" or "child." The picture would be considered to have high name agreement. In contrast, another picture might elicit "couch" 55% of the time and "sofa" 45%. This would constitute medium name agreement, and all else being equal, a picture of a baby would be named faster than a picture of a couch. Although name

agreement is calculated by looking at the variety of names produced across speakers, evidence suggests that it reflects the activation of multiple lemmas within a speaker (Balatsou et al., 2022). Assuming that semantic features activate lemmas for lower name-agreement pictures less than high agreement ones accounts for name-agreement effects.

A *morpheme* is the smallest unit of language that carries meaning. For example, "bees" contains the morpheme for "bee" and a plural "s," whereas "babysitting" contains "baby," "sit," and "ing" [see Morphology]. Relatively little research has addressed the production of morphologically complex words, and some theories omit them (see Wheeldon & Konopka, 2018). Some morphological processes may intervene between lemma selection and phonological encoding, but we will not further explore them here.

Having selected a lemma, the speaker must retrieve phonological information about the intended word, sometimes termed a *lexeme* (Kempen & Hoenkamp, 1987). This information includes not only the individual sounds (sometimes described as *phonemes* or *segments*) but also how they combine with syllable frames to form syllables. A *syllable* is a group of sounds that include a vowel. For example, the words "tack" and "cat" are each one syllable long, and they contain the same segments but in different syllable positions. Theories of production often posit a *syllable frame* that combines and orders initial consonants, a vowel, and then final consonants (e.g., Dell et al., 1997).

Syllable frames and the speech sounds that make up those syllable frames may be retrieved independently. The evidence for independence comes from the ability to prime a syllable frame when the speech sounds within the syllables differ (Costa & Sebastian-Gallés, 1998). So, all else being equal, "ta-co" would prime "ki-wi" because both have two syllables composed of a consonant and vowel, but "lim-bo" wouldn't prime "ki-wi" because it has an extra consonant ("m") in its first syllable frame. Other work suggests that syllable frames are filled incrementally, that is, speech sounds are retrieved in the order they will be spoken (Meyer, 1991). Thus, the "l" in "limbo" would be inserted in the syllable frame for "lim" before the "m" is.

The sounds that words share with other words in a person's vocabulary also influence the speed and success of phonological encoding in several ways (see Wheeldon & Konopka, 2018). For example, words that overlap in sounds with many other words tend to be less vulnerable to speech errors and tip-of-the-tongue states. However, the details of how the retrieval of a word is influenced by similar-sounding words is complicated.

Phonetic processing and articulation

Traditionally, most production researchers have tacitly or explicitly assumed that once abstract speech sounds have been retrieved and ordered as part of phonological encoding, they are converted directly into motor movements such that further investigation is better left to researchers focused on motor movements rather than language. So, ironically, the physical act of speaking is often ignored by those

who study the processes that result in speech (see, e.g., Buchwald, 2014; Bürki, 2023). Nevertheless, fine-grained differences in how speech sounds are produced must become explicit. For example, the "p" sound in "pan," "span," and "nap" are articulated slightly differently, although we think of them all as "p."

Questions, controversies, and new developments

How far in advance do speakers plan aspects of their utterances?

Researchers often assume that speakers strive to produce fluent speech (Clark & Clark, 1977). Pausing or saying "um" or "uh" in the middle of an utterance suggests that some aspect of the following utterance has not been planned yet. Such disfluencies are extremely common. For example, disfluencies occur once every 20 words in spontaneous speech (Shriberg, 1994), whereas speech errors occur roughly once every 1,000 words (Deese, 1984). Disfluencies suggest that speakers often plan less than an entire utterance down to the level of speech sounds before articulating the first word. While perfectly fluent speech could result from planning an utterance entirely before speaking, research shows that it may often be due to last-second preparation when the timing works out to have each unit ready when needed (Griffin, 2003). The consensus is that people rarely, if ever, plan entire utterances before beginning to speak even when fluent (e.g., Kempen & Hoenkamp, 1987; Levelt, 1989). Instead, production is incremental, with processing occurring at multiple levels simultaneously. For example, the lemma for an upcoming word is selected while an earlier word is articulated (e.g., selecting "bees" while saying "man").

Much work has addressed the minimum amount of planning prior to speech onset at different steps in the production needed for fluent speech. However, variations in methodology make it challenging to establish a minimum or modal degree of planning prior to speech onset because speakers operate under different constraints and with different goals across studies. Research on the scope of planning before speech has addressed issues such as the degree to which speakers prefer to know the gist of an event and select a sentence structure before speaking (e.g., Norcliffe et al., 2015); when verbs are selected (e.g., Schriefers et al., 1998); the time course for encoding nouns in complex noun phrases like, "The acorn that is over the bear..." (Allum & Wheeldon, 2007); whether final morphemes in a word can be encoded before initial ones (e.g., "print" before "re" in "reprint"; Roelofs, 1996); whether articulation of a word can begin before all of its sounds have been activated (e.g., do you need to know that "limbo" ends in "o" before starting to articulate the "l" Roelofs, 2002); and so on. It is uncontroversial to say that when speaking fluently, speakers at least select the lemma and initial sound of the first content word in an utterance before speaking; otherwise, they would have nothing to say. The rest is subject to debate.

Priming as learning across the lifespan

Language production research focuses on processing in adult speakers who have acquired their language and might be thought to have stable representations of their language with no need for change. However, producing a word or sentence has a long-lasting impact on speaking. For example, repetition priming for producing words can persist for weeks (Cave, 1997). Semantic interference from producing "couch" affects

the time and accuracy for producing "chair" over multiple seconds and several intervening, unrelated words (e.g., "flower" and "truck"; e.g., Schnur et al., 2006). Likewise, syntactic priming can be detected over the production of many structurally unrelated sentences (Bock & Griffin, 2000). Whereas earlier models focused on accounting for immediate effects of produced and perceived words (e.g., Levelt et al., 1999), more recent ones have incorporated learning mechanisms to accommodate long-lasting effects (e.g., Oppenheim et al., 2010). Each time someone says a word or uses a syntactic structure, they become faster and more accurate when doing so again. Learning is often conceptualized as strengthening the connections between representations that lead to successful speech, allowing for greater activation to spread between them (e.g., Damian & Als, 2005). For example, each time a lemma is selected, its semantic features become better able to activate it in the future. In addition, some models also incorporate a weakening of connections between activated features and lemmas that are not selected (see Oppenheim et al., 2010). So, after producing "chair," the overlapping features such as USED FOR SITTING activate "couch" less, making it slightly slower to produce. To account for long-lasting syntactic priming, the mappings between roles in a message (e.g., agent and patient) and the ordering of the phrases expressing them in a sentence are strengthened with use (Chang et al., 2006).

Conceptualizing priming effects as the result of slight but persistent strengthening mappings between representations allows these phenomena to connect to language acquisition, which is traditionally thought of as learning [see Language Acquisition]. For example, the syntactic priming effects seen in neurotypical adults are argued to rest on fundamentally the same learning mechanisms that underlie other learning (Bock & Griffin, 2000), such as children learning to produce passive sentences (e.g., Kumarage et al., 2022), second language learners acquiring new syntactic structures (e.g., Shin & Christianson, 2012), and the relearning of the syntactic structures by agrammatic aphasics (e.g., Lee & Man, 2017).

The nature of competition

Early models of word production assumed that semantically related words such as "couch" and "sofa" battled for selection in so that the more available "sofa" was, the longer it would take to select "couch" (Levelt et al., 1999). Effectively, everything was delayed until the single best word was selected. More recent approaches and data argue for a "good enough" approach to production, in which speakers will produce the first available word that adequately expresses their intention (Goldberg & Ferreira, 2022; Oppenheim et al., 2010). So, the selection of a lemma or syntactic structure is more of a horse race in which the strength of the second horse does not slow the first horse instead of a battle to the death in which stronger competitors delay the victory of the winner.

Multilingualism

The default for psycholinguistic studies is that participants self-identify as native speakers of the tested language (e.g., Dell et al., 1997; Meyer, 1991; Roelofs, 1996). The majority of people, particularly those outside the United States, can carry out conversations in at least two languages (e.g., Byers-Heinlein,

2019). Many of the studies discussed here were carried out with bilingual participants, ignoring the potential influence of another language. However, there is a vibrant subarea of psycholinguistics concerning how the knowledge of two or more languages influences production. The primary finding is that a speaker's first-learned or dominant language influences processing in their other language, and a second or nondominant language also influences processing of the first language (see Bailey et al., 2024).

Cognates are words that share sounds and meaning across languages such as the English "guitar" and Spanish "guitarra." Bilinguals are faster to produce cognates than noncognates (translation equivalents that are not similar in form, e.g., English "dog" and Spanish "perro"; Costa et al., 2000). Cognates are also less likely to trigger tip-of-the-tongue states than matched noncognates (Gollan & Acenas, 2004). This may be attributed to lemmas in both languages becoming active and providing converging activation to overlapping sound representations (Costa et al., 2000).

Syntactic priming occurs across languages (see van Gompel & Arai, 2018). For example, the German sentence, "Der Musiker verkaufte dem Agenten etwas Kokain" ("The musician sold the agent some cocaine"), would prime speakers to say, "A woman is giving a girl a lunchbox," rather than "A woman is giving a lunchbox to a girl." Cross-language syntactic priming has been replicated across many pairs of languages, including typologically unrelated ones. It is often considered evidence for shared syntactic representations.

Cross-linguistic research

English and Dutch were the predominant languages used in language production research in the 20th century (the Max Planck Institute for Psycholinguistics officially opened in the Netherlands in 1980). As of 2009, only approximately 30 out of the world's 5,000 languages had been explored in laboratory experiments on sentence production (Jaeger & Norliffe, 2009). Although some linguists argue for many universal tendencies across languages, there is great variability among them along many dimensions that a priori should affect processing (see Evans & Levinson, 2009). For example, languages differ in their basic word order. In English, a typical word order like "A girl ate pasta'" has the order of subject, verb, object. In contrast, about 13% of languages place the main verb ("ate") at the start of the sentence, and over 56% put the verb after the subject and object (Hammarström, 2016). Work on such languages has been important in testing the extent of incrementality in sentence production and influences on the choice of sentence structure (e.g., Momma et al., 2015; Norcliffe et al., 2015).

The search for the basic building blocks used in sentence production presents a strong example of the hazard of making universal claims based on a limited set of languages. Logically, we form sentences by combining words in novel ways, so words or lemmas are building blocks in speaking. In addition, morphemes (such as "kick" and "-ing" in "kicking") are the smallest units of meaning in a language and combine to form words. Speech error patterns reinforce the intuition that these are the units combined in speaking. For example, words accidentally exchange with other words in sentences (e.g., "She ate a park in the sandwich"). Morphemes can be recombined by accident, as in saying, "I'm not in the read for

mooding," instead of "mood for reading," in which the morphemes "mood" and "read" exchange while the progressive "-ing" remains in place. Such errors are attested in languages that are unrelated to Indo-European ones (Wells-Jensen, 2007). These observations suggest that universal steps in production involve selecting and placing words and morphemes in sequences.

Representations of single sounds are also implicated as building blocks by common speech errors such as saying "what the bell" instead of "what the hell" in Indo-European languages such as English, Dutch, and Spanish and have therefore been considered universal building blocks of sentences (e.g., Bock, 1991). However, as research has expanded beyond Indo-European languages, the question of sublexical units has become more complicated. For example, errors are more likely to involve whole syllables than individual sounds in Mandarin Chinese (see Alderete & O'Séaghdha, 2022). Likewise, Japanese speech errors tend to involve phonological units that are bigger than a single sound (*morae*). English has about 9,000 different syllables, and Dutch has at least 12,000, whereas Mandarin has 400 syllables (1,200 including tones), and Japanese only has 100 morae. These cross-linguistic differences in error patterns and vocabulary suggest that unit-like behavior at the level of phonological encoding is influenced by how frequently sound sequences are combined in the vocabularies of different languages. In other words, words fall apart when the following sound is not very predictable (*low transitional probability*), and this varies across languages [see Statistical Learning]. Fortunately, researchers are increasingly studying production in typologically different languages, allowing true universal processing principles to be discovered (e.g., Wells-Jensen, 2007; Norcliffe et al., 2015; Sauppe, 2017).

Broader Connections

Action production

To speak is to act. Although language can be seen as an unparalleled system of combining symbols of many levels of abstraction, it shares many characteristics with producing other motor actions (Lashley, 1951). This is most easily seen in comparisons with playing music, which also involves sequencing movements with a complex hierarchical structure. Ordering errors across the two domains share many similarities. In addition, the timing of executing one action (articulating a word or flexing an arm) while preparing the next is similar in many ways (Griffin, 2003).

Working memory

Baddeley's (1992) model of *working memory* (which is responsible for the short-term storage and manipulation of information) includes a *phonological loop* for rehearsing verbal material such as word lists [see Working Memory]. Psycholinguists have often considered how working memory capacity interacts with language production processes (see Schwering & MacDonald, 2020).

Cognitive neuroscience

Historically, neuroimaging and electrophysiological measures have not played a large role in language production research because the movements required for speaking disrupted measurement. Also, the research questions often differed, with imaging typically focusing on localization rather than understanding processes [see Neuroscience of Language]. However, researchers are increasingly finding ways to circumvent the limitations of neuroimaging techniques and relate brain activity to production processes. Such studies, for example, have found very early phonological effects in word production, supporting the hypothesis that phonological encoding can begin before lemma selection is complete or that the stages are less distinct than typically assumed (Strijkers et al., 2017).

Acknowledgments

I am indebted to Maria Bedny, Michael Frank, Bill McNeill, Lauretta Reeves, and Bryan Register for comments on drafts of this text and Eri Pilon for copyediting.

Further reading

- Buchwald, A. (2014). Phonetic processing. In M. Goldrick, F. S. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 245–258). Oxford Press. https://doi.org/10.1093/oxfordhb/9780199735471.001.0001
- Goldrick, M. (2014). Phonological processing: The retrieval and encoding of word form information in speech production. In M. Goldrick, F. S. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 228–244). Oxford Press. https://doi.org/10.1093/oxfordhb/9780199735471.001.0001
- Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007
- Wheeldon, L. R., & Konopka, A. (2023). Grammatical encoding for speech production. Cambridge University Press. https://doi.org/10.1017/9781009264518

References

- Alderete, J., & O'Séaghdha, P. G. (2022). Language generality in phonological encoding: Moving beyond Indo-European languages. *Language and Linguistics Compass*, *16*(7). https://doi.org/10.1111/lnc3.12469 ↩
- Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 791–810. https://doi.org/10.1037/0278-7393.33.4.791 ↩
- Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *14*(4), 382–391.

https://doi.org/10.1016/s0022-5371(75)80017-x

↩

- Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

↩

- Bailey, L. M., Lockary, K., & Higby, E. (2024). Cross-linguistic influence in the bilingual lexicon: Evidence for ubiquitous facilitation and context-dependent interference effects on lexical processing. *Bilingualism: Language and Cognition*, *27*(3), 495-514. https://doi.org/10.1017/S1366728923000597

↩

- Balatsou, E., Fischer-Baum, S., & Oppenheim, G. M. (2022). The psychological reality of picture name agreement. *Cognition*, 218, 104947. https://doi.org/10.1016/j.cognition.2021.104947

↩

- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*(3), 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

↩

- Bock, J. K. (1991). A sketchbook of production problems. *Journal of Psycholinguistic Research*, *20*, 141–160. https://doi.org/10.1007/bf01067212

↩

- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, *129*(2), 177–192. https://doi.org/10.1037//0096-3445.129.2.177

↩

- Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, *48*, 653–685. https://doi.org/10.1016/s0749-596x(03)00007-x

↩

- Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, *93*, 89–114. https://doi.org/10.1348/000712602162463

↩

- Brown, R., & McNeill, D. (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*(4), 325–337. https://doi.org/10.1016/S0022-5371(66)80040-3

↩

- Buchwald, A. (2014). Phonetic processing. In M. Goldrick, F. S. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 245–258). Oxford Press.

https://doi.org/10.1093/oxfordhb/9780199735471.001.0001

↩

- Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language*, *30*(5), 542–579. https://doi.org/10.1016/0749-596X(91)90026-G

↩

- Bürki, A. (2023). Phonological processing: Planning the sound structure of words from a psycholinguistic perspective. In R. J. Hartsuiker & K. Strijkers (Eds.), *Language production* (pp. 66-96). Routledge.

↩

- Byers-Heinlein, K., Esposito, A. G., Winsler, A., Marian, V., Castro, D. C., Luk, G., Brown, B., & DeJesus, J. (2019). The case for measuring and reporting bilingualism in developmental research. *Collabra: Psychology*, *5*(1), 37. https://doi.org/10.1525/collabra.233

↩

- Cave, C. B. (1997). Very long-lasting priming in picture naming. *Psychological Science*, *8*(4), 322–325. https://doi.org/10.1111/j.1467-9280.1997.tb00446.x

↩

- Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272. https://doi.org/10.1037/0033-295x.113.2.234

↩

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

↩

- Costa, A., & Sebastian-Gallés, N. (1998). Abstract phonological structure in language production: Evidence from Spanish. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(4), 886–903. https://doi.org/10.1037/0278-7393.24.4.886

↩

- Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1283–1296. https://doi.org/10.1037/0278-7393.26.5.1283

↩

- Damian, M. F., & Als, L. C. (2005). Long-lasting semantic context effects in the spoken production of object names. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1372.

https://doi.org/10.1037/0278-7393.31.6.1372

↩

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

↩

- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in normal and aphasic speakers. *Psychological Review*, *104*(4), 801–838. https://doi.org/10.1037/0033-295x.104.4.801

↩

- Emmorey, K. (2023). Sign production: Signing vs. speaking: How does the biology of linguistic expression affect production? In R. J. Hartsuiker & K. Strijkers (Eds.), *Language production* (pp. 233–256). Routledge.

↩

- Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*(5), 429–448. https://doi.org/10.1017/S0140525X0999094X

↩

- Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, *47*, 27–52. https://doi.org/10.2307/412187

↩

- Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9, pp. 133–177). Academic Press. https://doi.org/10.1016/s0079-7421(08)60270-4

↩

- Goldberg, A. E., & Ferreira, F. (2022). Good-enough language production. *Trends in Cognitive Sciences*, *26*(4), 300-311. https://doi.org/10.1016/j.tics.2022.01.005

↩

- Gollan, T. H., & Acenas, L. A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 246-269. https://doi.org/10.1037/0278-7393.30.1.246

↩

- Griffin, Z. M. (2003). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin & Review*, *10*(3), 603–609. https://doi.org/10.3758/bf03196521

↩

- Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*(4), 274–279. https://doi.org/10.1111/1467-9280.00255

↩

- Hammarström, H. (2016). Linguistic diversity and language evolution. *Journal of Language Evolution*, 1(1), 19–29. https://doi.org/10.1093/jole/lzw002

↩

- Jaeger, T. F., & Norcliffe, E. J. (2009). The cross-linguistic study of sentence production. *Language and Linguistics Compass*, *3*(4), 866–887. https://doi.org/10.1111/j.1749-818x.2009.00147.x

↩

- Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, *11*(2), 201–258. https://doi.org/10.1207/s15516709cog1102_5

↩

- Kumarage, S., Donnelly, S., & Kidd, E. (2022). Implicit learning of structure across time: A longitudinal investigation of syntactic priming in young English-acquiring children. *Journal of Memory and Language*, 127, 104374. https://doi.org/10.1016/j.jml.2022.104374

↩

- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). Wiley.

↩

- Lee, J., & Man, G. (2017). Language recovery in aphasia following implicit structural priming training: A case study. *Aphasiology*, *31*(12), 1441–1458. https://doi.org/10.1080/02687038.2017.1306638

↩

- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.

↩

- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1–38. https://doi.org/10.1017/s0140525x99001776

↩

- Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language*, *30*(5), 69–89. https://doi.org/10.1016/0749-596X(90)90050-A

↩

- Meyer, A. S., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition*, *20*(6), 715–726. https://doi.org/10.3758/bf03202721

↩

- Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, *66*(2), B25-B33. https://doi.org/10.1016/s0010-0277(98)00009-2

↩

- Momma, S., Slevc, L. R., & Phillips, C. (2015). The timing of verb selection in japanese sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(5), 813–824. https://doi.org/10.1037/xlm0000195

↩

- Nooteboom, S. G. (1973). The tongue slips into patterns. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 144–156). De Gruyter. (Original published in 1969)

↩

- Nooteboom, S. G. (1973). The tongue slips into patterns. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 144–156). De Gruyter. (Original work published 1969)

↩

- Norcliffe, E., Konopka, A. E., Brown, P., & Levinson, S. C. (2015). Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience*, *30*(9), 1187–1208. https://doi.org/10.1080/23273798.2015.1006238

↩

- Oldfield, R. C., & Wingfield, A. (1964). The time it takes to name an object. *Nature*, *202*(4936), 1031–1032. https://doi.org/10.1038/2021031a0

↩

- Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

↩

- Potter, M. C. (1975). Meaning in visual search. *Science*, *187*(4180), 965–966. https://doi.org/10.1126/science.1145183

↩

- Roelofs, A. (1996). Serial order in planning the production of successive morphemes of a word. *Journal of Memory and Language*, *35*(6), 854–876. https://doi.org/10.1006/jmla.1996.0044

↩

- Roelofs, A. (2002). Spoken language planning and the initiation of articulation. *Quarterly Journal of Experimental Psychology*, *55*(2), 465–483. https://doi.org/10.1080/02724980143000488 ↩
- Sauppe, S. (2017). Symmetrical and asymmetrical voice systems and processing load: Pupillometric evidence from sentence production in Tagalog and German. *Language*, *93*(2), 288–313. https://doi.org/10.1353/lan.2017.0015 ↩
- Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language*, *54*(2), 199-227. https://doi.org/10.1016/j.jml.2005.10.002 ↩
- Schriefers, H., Teruel, E., & Meinhausen, R. M. (1998). Producing simple sentences: Results from picture-word interference experiments. *Journal of Memory and Language*, *39*(4), 609–632. https://doi.org/10.1006/jmla.1998.2578 ↩
- Schwering, S. C., & MacDonald, M. C. (2020). Verbal working memory as emergent from language comprehension and production. *Frontiers in Human Neuroscience*, *14*, 68. https://doi.org/10.3389/fnhum.2020.00068 ↩
- Shin, J. A., & Christianson, K. (2012). Structural priming and second language learning. *Language Learning*, *62*(3), 931-964. https://doi.org/10.1111/j.1467-9922.2011.00657.x ↩
- Shriberg, E. E. (1994). *Preliminaries to a theory of speech disfluencies* (Doctoral dissertation, University of California, Berkeley). ↩
- Sridhar, S. N. (1988). *Cognition and sentence production: A cross-linguistic study.* Springer-Verlag. ↩
- Strijkers, K., Costa, A., & Pulvermüller, F. (2017). The cortical dynamics of speaking: Lexical and phonological knowledge simultaneously recruit the frontal and temporal cortex within 200 ms. *NeuroImage*, *163*, 206–219. https://doi.org/10.1016/j.neuroimage.2017.09.041 ↩
- Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., Arévalo, A. L., Vargha, A., & Bates, E. (2005). Timed action and object naming. *Cortex*, *41*(1), 7–25.

https://doi.org/10.1016/s0010-9452(08)70174-6

↩

- van Gompel, R. P. G., & Arai, M. (2018). Structural priming in bilinguals. *Bilingualism: Language and Cognition*, *21*(3), 448–455. https://doi.org/10.1017/s1366728917000542

↩

- Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, *8*(4), 314–317. https://doi.org/10.1111/j.1467-9280.1997.tb00444.x

↩

- Wells-Jensen, S. (2007). A cross-linguistic speech error investigation of functional complexity. *Journal of Psycholinguistic Research*, *36*(2), 107–157. https://doi.org/10.1007/s10936-006-9036-5

↩

- Wheeldon, L. R., & Konopka, A. E. (2018). Spoken word production: Representation, retrieval, and integration. In S. A. Rueschemeyer & M. G. Gaskell (Eds.), *The Cambridge handbook of psycholinguistics* (2nd ed., pp. 335-371). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198786825.001.0001

↩

- Yoon, S. O., & Brown-Schmidt, S. (2023). Understanding language use in social contexts. In R. J. Hartsuiker & K. Strijkers (Eds.), *Language production* (pp. 284–303). Routledge. https://doi.org/10.4324/9781003145790-12

↩