



HAL
open science

Chapter 1. Models of language production and the temporal organization of lexical access

Niels Schiller, F.-Xavier Alario

► **To cite this version:**

Niels Schiller, F.-Xavier Alario. Chapter 1. Models of language production and the temporal organization of lexical access. Mikel Santesteban; Jon Andoni Duñabeitia; Cristina Baus. Bilingualism through the Prism of Psycholinguistics, 17, John Benjamins Publishing Company, pp.28-53, 2023, Bilingual Processing and Acquisition, 9789027213969. 10.1075/bpa.17.01sch . hal-04241417

HAL Id: hal-04241417

<https://hal.science/hal-04241417>

Submitted on 13 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Schiller, N. O., & Alario, F.-X. (2023). Models of Language Production and the Temporal Organization of Lexical Access. In M. Santesteban, J. A. Duñabeitia, & C. Baus (Eds.), *Bilingualism Through the Prism of Psycholinguistics: In Honour of Albert Costa*. John Benjamins Publishing Company.

<https://doi.org/10.1075/bpa.17.01sch>

CHAPTER 1

Models of language production and the temporal organization of lexical access

Niels O. Schiller^{1,2} & F.-Xavier Alario³

¹ Leiden University Centre for Linguistics (LUCL), Leiden University, The Netherlands

² Leiden Institute for Brain and Cognition (LIBC), Leiden University, The Netherlands

³ Aix-Marseille Université, CNRS, LPC, Marseille, France

This chapter focusses on models of language production and the temporal organization of the processes involved in lexical access. Albert Costa contributed significantly to advancing our understanding of lexical access. Through the investigation of multilingual speakers, he added an important dimension to this discussion. Here, we provide a discussion of discrete and cascaded/interactive accounts of lexical access and summarize some of the crucial experimental evidence that has been offered for and against these different theories, roughly over the years when Albert contributed to these debates. We will also discuss whether or not lexical selection is competitive and what research on multilingual speech production contributed to our knowledge on lexical access. The discussion is mainly restricted to behavioral data, with some references to electrophysiological studies. We conclude that considering the experimental evidence a strict version of discrete processing models is difficult to maintain. Research on bilingual language production has made important contributions to the discussion, however, the picture is not yet complete.

1. Speaking – a complex behavioral skill

Speech production often seems to be a simple and fast process. Our average speaking rate is in the order of several syllables per second. Depending on the language and its average word length, this means that even within a couple of minutes, speakers can produce hundreds of words, maybe get close to a thousand. At the same time, they make relatively few mistakes or speech errors. In a non-formal conversation, presumably we make fewer than one error per 1,000 words (Deese, 1984). This suggests that the (neuro-)cognitive architecture underlying speech production is quite robust.

However, speaking is not a simple process. Rather, it is one of the most complex forms of skilled behavior that human beings are equipped with and execute on a regular basis. Producing speech involves the planning of several processes some of which have been suggested to be serial (Lashley, 1951). Our task as speakers is to translate the message, we intend to convey into motor actions of our speech apparatus, i.e., our tongue, our jaws and our lips, such that air molecules are set into motion and our interlocutors can extract semantic content from the sensations caused by these vibrations.

The complexity of speaking can be illustrated with the example of articulation, that is, the final execution processes of planned speech. One may compare the act of speaking with directing a symphony, where the speaker is the conductor, the articulatory-motor plan is the score, and our articulatory organs are the instruments. Producing speech is a complex interaction between the movement of the articulators on a gestural score and their timing is extremely sensitive (Browman & Goldstein, 1986; Goldstein & Fowler, 2003). MacNeilage (2008) noted that there are about 40 muscles involved in the movements of our articulatory organs and estimated that every five milliseconds speakers execute one muscle event, that is, about 225 different muscle activations for every second of speech. Speakers need to come up with articulatory-motor plans that orchestrate these muscles in subtle ways. Fine-grained changes in the positioning of the articulators can have huge effects on the auditory perception of the produced speech signal. Getting control over the movements of our articulators is a motor skill that requires a lot of practice. In fact, newborn infants start acquiring it after a couple of months in their so-called babbling phase, long before they produce their first words (for a review, see Guenther, 2003, 2016).

The example of articulatory execution illustrates the complexity of speech production, but it is only one of the stages involved in word production. Speech production can be broken down into various other levels and processes. The individual processes are related to each other according to a certain hierarchy and sequencing, and it is important to describe the temporal organization of these processes. It is generally accepted that conceptual processing (what is to be said) precedes articulatory processing (how exactly are words pronounced). However, the nature of many processes occurring in between, including their ordering, is still debated. Following a classic distinction that Albert Costa contributed exploring, we will discuss two classes of speech production models (discrete vs. cascaded/interactive) and provide empirical arguments in favor and against those models thereby showing how the discussion developed over the years. Furthermore, we will debate whether or not lexical selection is competitive, and finally extend our discussions to bilingual word production.

2. Language Production Models

Over the course of the last decades, researchers developed comprehensive models of language production to account for and describe the process of speech production in great detail (e.g., Bock & Levelt, 1994; Caramazza, 1997; Dell, 1986, 1988; Fromkin, 1971; Garrett, 1975, 1980; Levelt, 1989; Levelt, Roelofs, & Meyer, 1999; Mahon et al., 2007; Oppenheim, Dell, & Schwartz, 2010; Roelofs, 2000, 2018; Roelofs & Ferreira, 2019). Speaking entails the activation of meaning from memory, the selection of words from our mental lexicon, the retrieval of morpho-syntactic features of words, the activation, selection and encoding of phonological form (segments, syllables, metrical structures, etc.), and the execution of articulatory-phonetic motor programs (for an overview see Griffin & Ferreira, 2006). It should be mentioned that, so far, speech production models are mostly limited to the discussion of isolated words or short utterances whereas in real-life verbal interactions, speakers create narratives, that is, (combinations of) sentences, and often engage in dialogues as well (see Pickering & Garrod, 2004; 2013). However, the issue of dialogue production lies beyond the scope of this chapter.

The retrieval of lexical entries from long-term memory, that is, the so-called mental lexicon (Schiller, 2021), is an essential component of the speech production process. The mental lexicon is often taken to refer to the part of our language processing system that hosts the word forms we know and their corresponding meanings, that is, a neurocognitive version of a dictionary. The retrieval of lexical items from the mental lexicon is called lexical access (for an overview see Schiller & Verdonschot, 2015). The meaning component of a lexical item, including conceptual-semantic processing, precedes the encoding of its form, that is, morphological processing (see overviews in Schiller & Verdonschot, 2019; Schiller, 2020) and phonological-phonetic processing (or phonological encoding; see Schiller, 2006, for a review). However, although there has been little agreement in the past regarding the exact time course of the flow of information as part of lexical access processes within the speech production system, we will show that in more recent years some issues have been clarified. In fact, the field started out with a discrete processing view, however, over the years with dominant view has changed to a cascaded processing perspective.

2.1. Discrete models of language production

There are models that assume a strictly serial or discrete flow of information from mind to mouth (Bock, 1995), that is, from the preparation of concepts to the execution of articulation (Bloem & La Heij, 2003; Butterworth, 1989; Garrett, 1980; 1988; Levelt, 1989; 1999; 2001; Levelt et al., 1999; Roelofs, 2018). These serial or discrete processing models assume that semantic concepts to be spoken about activate a network of relatively closely related lexical candidates (e.g., all the words belonging to a particular semantic category such as furniture or body parts). However, just a single lexical candidate is selected and taken to further activate and encode the corresponding word form at the phonological level. In lexical access processes, it is important to differentiate between the activation and selection of a lexical item. The selection of a lexical candidate, that is, the target lexical item, is determined by the speaker's intention of what they want to say.

Discrete models claim that the processes involved in word retrieval during lexical access operate in strict temporal succession (Levelt et al., 1999). To illustrate the processes of lexical access in language production models we take the example of a picture naming task, for instance, naming the picture of a cow. According to discrete theories of lexical access, the picture of a cow activates the concept COW. However,

many other concepts that belong to the same semantic category node (FARM) ANIMAL such as horse, goat, sheep, pig, etc. are activated as well via the spreading of activation of a semantic category. Activated concepts automatically activate their corresponding lexical entries – also called lemmas (Levelt et al., 1999). However, Levelt et al. (1999, p. 15) argued that only selected lemmas will become phonologically activated. This proposal became the primary distinguishing feature of their theoretical approach. They argued that it seems counterproductive to activate the word forms of all activated lemmas, even those that have not been selected because the phonological word form activation of non-selected lexical items may potentially interfere with the morpho-phonological encoding of the target lexical item. That is why discrete models of lexical access predict that the labels of semantic category members of the target picture, that is, horse, goat, sheep, pig, etc., are not phonologically encoded - they are not selected for production.

2.2. Experimental support for discrete models of speech production

Evidence from speech errors (e.g. Fromkin, 1971; 1973; Garrett, 1975; 1980; 1982), tip-of-the-tongue states (e.g. Brown & McNeill, 1966; Vigliocco, Antonini, & Garrett, 1997), speakers with aphasia (Goodglass, Kaplan, Weintraub, & Ackerman, 1976), electrophysiological covert naming studies (Schmitt, Münte, & Kutas, 2000; Van Turenout, Brown, & Hagoort, 1997) and speeded naming paradigms (Levelt, Schriefers, Vorberg, Meyer, Pechmann, & Havinga, 1991; Schriefers, Meyer, & Levelt, 1990) suggested that in general lexical-semantic processing precedes phonological processing in speech production. For instance, based on a corpus of word and sound exchanges, Garrett (1982) found that word exchanges generally take place between syntactic phrases and the words involved in an exchange usually belong to the same syntactic category. Sound exchanges, however, generally take place within a syntactic phrase and the sounds involved often occur in words of different syntactic categories. Garrett (1982) proposed that word exchanges occur before sound exchanges, that is, when speakers are planning the ordering of words in the syntactic phrases of an utterance and syntactic category information about words is necessary (the functional level in Garrett's terms). Sound exchanges occur later, that is, during phonological encoding, when the planning scope is smaller and syntactic category information is no longer relevant (the positional level).

Evidence for discrete lexical processing also comes from experiments, for instance, from the famous study using the picture-word interference (PWI) paradigm that Levelt and colleagues published in 1991. In Experiment 6 of that study, native Dutch participants were presented with a series of pictures and requested to name them. An auditory stimulus, that is, either a word or a non-word, was presented 73 ms after picture onset in about one-third of the trials. In these cases, participants were requested to refrain from naming the picture and instead make a button-press lexical decision to the auditory stimulus. Each auditory word stimulus belonged to one of the following four conditions: for instance, when the picture was a desk, the auditory word was bureau ('desk') in the identical condition, stoel ('chair') in the semantic condition, stoep ('pavement') in the phonological condition, and muts ('cap') in the unrelated condition. In an earlier experiment of the same study, the authors showed that phonologically related words such as buurman ('neighbour') were responded to faster in lexical decision than unrelated words. This demonstrated the phonological activation of the picture's label (i.e., bureau), which primed the phonologically related word buurman. The phonological condition in Experiment 6 of their study, however, included a word (i.e., stoep 'pavement') that was not phonologically related to the target object 'desk' itself but rather to a semantic alternative of it, that is, the semantic category member stoel ('chair') in this case. If these

semantic category members such as *stoel* ('chair') become phonologically activated upon perceiving the target picture *bureau* ('desk'), then a phonologically related word like *stoep* should yield a phonological priming effect that is reflected in lexical decision latencies – just like for the pair *bureau* – *buurman*. The results of Experiment 6, however, showed that there was no effect for words in the phonological condition. However, there was a priming effect in the identical and the semantic condition, proving that semantic alternatives were activated to some extent. Levelt et al. (1991) argued based on their results that non-selected semantic alternatives do not become phonologically encoded in speech production, supporting their discrete processing view.

In a critique of Levelt et al. (1991), Dell and O'Seaghdha (1991, 1992) argued that the methodology may not have been sensitive enough to monitor phonological activation of multiple lexical candidates. They claimed that semantic alternatives (e.g., *stoel* 'chair') may receive less activation than the target (e.g., *bureau* 'desk'), and therefore a phonologically related word to the semantic alternative like *stoep* ('pavement') may receive even less activation. According to Dell and O'Seaghdha, that is why the effect on a mediated phonologically related word like *stoep* for the target picture *bureau* might be difficult to detect.

In another seminal study, Schriefers et al. (1990) provide additional evidence supporting discrete lexical access. They asked Dutch participants to name a set of pictures (e.g., *bureau* 'desk') and presented auditory distractors at three different points in time, so-called stimulus onset asynchronies (SOAs), relative to the presentation of the target picture, that is, 150 ms before picture onset (-150 ms), simultaneous with picture onset (0 ms), and 150 ms after picture onset (+150 ms). At the earliest SOA, semantically related distractors (e.g., *stoel* 'chair') exerted inhibition on the naming process relative to an unrelated distractor (e.g., *fles* 'bottle'), while at the later SOAs 0 ms and +150 ms, but not at the earliest SOA of -150 ms, phonologically related distractors (e.g., *buurman* 'neighbour') facilitated picture naming. Schriefers et al. argued that their findings of early influence of semantic distractors and a later effect of phonological distractor words should be taken as support for the general notion that semantic processing precedes phonological processing (however, Jescheniak & Schriefers, 2001 and Starreveld, 2000 obtained early phonological effects). More important, however, for the support of discrete models is the fact that the semantic inhibition and the phonological facilitation effects did not overlap in time. This suggests that before the system can begin to phonologically encode words, processing at the lexical-semantic level must be completed. Although the findings of Schriefers et al. are consistent with discrete models, in which lexical access proceeds in two serially ordered stages which do not affect each other — retrieval of lexical-semantic representations followed by retrieval of phonological word forms — more recent data have shown that the timing of semantic and phonological effects may be more complex (e.g., Damian & Martin, 1999; Starreveld & La Heij, 1996), and that the temporal sequencing of these effects may in fact be overlapping. Damian and Martin (1999), for instance, did find both semantic interference and phonological facilitation at SOAs -100 ms and 0 ms, that is, semantic and phonological effects overlapped in time.

Support for discrete models of speech production also came from a series of electrophysiological studies by Van Turenout, Brown and Hagoort (1997, 1998). In these studies, participants named pictures in Dutch. In one quarter of the trials, a frame appeared around the picture after 300 ms. In that case, participants were requested to refrain from naming the picture and instead engage in a double decision task. A semantic categorization task determined participants' response hand, for instance, respond with the right hand if the picture is animate (e.g., tiger, spider), and with the left hand if it is

inanimate (e.g., table, scissors). However, the participants were told to only press a button with the corresponding hand if the picture name starts with a /t/ (as in tiger or table) but not if the picture name starts with an /s/ (as in spider or scissors). The electroencephalogram (EEG) was measured above the motor cortex from electrodes C3 and C4. Results showed a lateralized readiness potential (LRP), a derivative from the ERP, in the go-condition, that is, when participants pressed the button because the picture name started with a /t/. Interestingly, however, the development of an LRP was initially also visible in the no-go-condition, that is, when the picture name started with an /s/. After 40 ms, the LRP in the no-go-condition diverted from the LRP in the go condition and went back to baseline. This has been interpreted as follows: first, participants make the semantic animateness decision, chose their response hand and prepare their motor response (button press). However, in the no-go-condition, no button press is required, and therefore motor preparation is stopped, that is, at the point in time when the phonology of the target becomes available. Van Turennout et al. (1997) claimed based on their results that the phonological make-up of a lexical item becomes available about 40 ms after aspects of its meaning have been processed.

In a follow-up experiment, Van Turennout et al. (1997) showed that when the phonological decision is based on the final sound of the objects' name, the LRP goes back to baseline after 120 ms. This led them to claim that the process of phonological encoding of the picture names used in their experiments takes about 80 ms. In a further LRP study, the same authors showed that lexico-syntactic information of a word (e.g., grammatical gender) becomes available approximately 40 ms before speakers have access to its phonological information (Van Turennout et al., 1998). More electrophysiological studies followed (e.g., Schiller, Bles, & Jansma, 2003; Schmitt et al., 2000) contributing to the time course of lexical processing calculated and proposed by Indefrey and Levelt (2004). Excellent reviews of discrete models of language production are provided by Levelt (1992, 1999) and Roelofs (2000).

The task used in the Van Turennout et al. (1997) study has been criticized for not directly reflecting the speech production process but only by proxy because it involves a judgement about language information, i.e., a metalinguistic decision. In fact, participants were asked to carry out meta-linguistic tasks such as phoneme monitoring, grammatical gender decision, or semantic categorization. Other studies employed delayed naming to avoid EEG artefacts. Strijkers and Costa (2011, 2016) therefore questioned in how far these tasks resemble speech production under normal conditions. They also wondered about the influence of other cognitive processes such as inhibition (of immediate naming) and monitoring (of meta-linguistic decisions) in these tasks on the time course of speech planning.

Discrete models of speech production imply a certain time course of processing and based on electrophysiological data (EEG/ERP and MEG) models of the exact time course of the processes involved in speech production and their durations have been proposed (Indefrey & Levelt, 2004; Indefrey, 2011). Costa, Strijkers, Martin and Thierry (2009) estimated lexical selection to start around 200 ms after picture onset in a picture naming task. Strijkers, Costa and Thierry (2010) report data from L1 and L2 speakers providing a similar estimation for the onset of lexical access employing word frequency and cognate status (see also Christoffels, Firk, & Schiller, 2007). The reported temporal onset falls into the time window that was previously inferred and claimed by Indefrey and Levelt (2004) as the time window for lexical access. Costa et al. (2009) estimate that the process of lexical selection lasts for about 180 ms and suggest that the timing of this process is largely independent of lexical parameters such as word length and lexical

frequency (see also Strijkers et al., 2010). These lexical parameters, however, determine the duration of later lexical and post-lexical processes such as phonological encoding, and finally determine naming latencies.

Despite these seemingly consistent data regarding the temporal onset of some lexical processes, a word of caution may be in place. Munding, Dubarry and Alario (2016), for instance, conducted a meta-analysis of word production involving 14 MEG studies. They report relatively early onsets of activity (approximately 180 – 250 ms post stimulus onset) for visual and semantic processing as well as for lexical access, and somewhat later (around 300 – 600 ms, sometimes up to 1,000 ms post-stimulus onset) for phonological processing, but also motor function including motor planning and articulatory processes. In other words, while they were able to see this division into earlier and later activity, any further subdivision did not seem possible due to largely overlapping patterns of brain activity in time. Therefore, we must be cautious with the detailed time course provided by Indefrey (2011; see also Indefrey & Levelt, 2004) because several studies have reported, for instance, articulatory or motor activity in word naming much earlier than the onset of 600 ms post stimulus estimated by Indefrey. Munding et al. (2016) therefore doubt that the MEG data are in line with a strictly serial or discrete model of word production. However, in all fairness, the authors concluded that “it is not yet possible to conclusively discard or accept any one current model on the basis of the currently assessed evidence” (Munding et al., 2016, p. 459).

Similarly, Strijkers and Costa (2016) warn to be cautious with the time course of neural activation during speech production. For example, for some effects, like the lexical frequency effect or the semantic interference effect, it is still debated when they arise during lexical access. Therefore, such studies may not be used to support time course models because it is not entirely clear where in the process of speech production such effects arise.

Finally, studies that explored distinct spatial and temporal components of word production within the same participants, such as the repetitive TMS study conducted by Schuhmann, Schiller, Goebel and Sack (2012) stimulating three different brain areas at different points in time during overt picture naming, do not completely support the picture proposed by Indefrey and Levelt (2004).

2.3. Cascaded and interactive models of speech production

The view that language production proceeds following strict serial temporal order regarding the processes between semantic and phonological encoding has been the dominant view for a long time. However, based on speech error data from patients with aphasia, Dell (1986, 1988) already put forward his interactive theory (cascaded activation plus feedback) and model of speech production early on. Moreover, since the second half of the 1990's approximately, the strict discrete view has also been challenged with chronometric data from healthy speakers, which led to the postulation of so-called cascaded theories of language production.

These models assume that the time course of processes is not discrete, but rather in cascade or even interactive (Caramazza, 1997; Dell, 1986, 1988; Dell, Nozari, & Oppenheim, 2014; Goldrick, 2006; Humphreys, Riddoch, & Quinlan, 1988; MacKay, 1987; Mahon et al., 2007; Oppenheim et al., 2010; Rapp & Goldrick, 2000; Schwartz, Dell, Martin, Gahl, & Sobel, 2006; Stemberger, 1985). Cascaded and interactive models of lexical access assume that all activated (but not-yet-selected) word candidates in the

lexicon activate their corresponding phonological forms prior to the selection of a lexical candidate as the target item. Due to their architecture, activation in the lexicon cascades from levels that are relatively higher in the system, such as lexical-semantic processing, to levels that are relatively lower in the system, such as phonological-phonetic processing. Importantly, lower-level processing can start as early as some activation from higher levels is available, even without lexical selection being completed. To illustrate this by the example of naming the picture of a cow, cascaded theories predict that not only the target lexical item cow is phonologically encoded, all activated category members as well, that is, goat, sheep, pig, etc. because these theories do not assume that only one lexical item is selected and further encoded.

2.4. Experimental evidence for cascaded and interactive models of language production

Cascaded and interactive models of language production have also been supported by experimental evidence. For instance, Peterson and Savoy (1998) aimed to reveal the phonological activation of alternative lexical candidates and conducted a series of PWI experiments investigating near-synonyms. In one of their experiments, they regularly presented a written word in the middle of the picture at one of several positive SOAs (+ 100 ms to + 400 ms). Whenever a word appeared, participants were required to refrain from picture naming and name the word. Their experimental items consisted of word-naming trials in which written words were presented with so-called ambiguous pictures, that is, pictures that have two, nearly synonymous labels. Take, for instance, the picture of a piece of furniture on which two or three people can sit comfortably, usually found in living rooms. In an independent picture naming experiment of the same study, a different group of participants called this picture with its dominant name couch in 84 % of the cases and in 16 % of the cases with its secondary name sofa. The written words that appeared on half of the trials had a phonological relationship to the dominant name (e.g., count) of the depicted object (e.g., couch) or with the alternative name (e.g., soda) or they were not phonologically related (e.g., horse). Peterson and Savoy (1998) asked the question whether or not there was a priming effect (reflected in word naming latencies) for words like soda that had a phonological relationship with the alternative name of the picture (e.g., sofa). Importantly, different outcomes are predicted here by discrete and cascaded models of lexical access. Both models predict phonological priming for words (e.g., count) that are phonologically related to the dominant label of the ambiguous picture because the dominant name couch is the label that the majority of the participants will select on most trials. Preparing the label couch activates segments that overlap with the written word count, and hence count will be named faster than an unrelated control word. However, since non-selected lexical items (e.g., sofa) do not become phonologically activated according to discrete theories, they do not predict an effect for soda. On the contrary, cascaded models predict priming not only for words that are related to the dominant label, but also words that are related to the secondary label (e.g., sofa). This is due to the assumption that even non-selected lexical candidates become phonologically activated. Peterson and Savoy (1998) demonstrated significant phonological priming effects for both types of words, for instance, count and soda, compared to unrelated control items. Their result suggests that several labels of a picture can become phonologically activated during the lexicalization of a picture.

Similarly, Jescheniak and Schriefers (1998) reported phonological effects in a PWI study in German. Their experimental trials consisted of pictures that had ambiguous labels in German (e.g., Schäfer - Hirte; both meaning 'shepherd'). In the experiment, auditory distractor words that were either phonologically related to the dominant label of

the picture (e.g., Schädel 'skull') or the secondary label of the picture (e.g., Hirn 'brain') were presented. Their findings showed that picture naming of the dominant label (e.g., Schädel) was reliably faster for phonologically related distractors (e.g., Schäfer), while distractors (e.g., Hirte) that were phonologically related to the secondary label (e.g., Hirn) inhibited picture naming with the dominant label. Presumably, both alternative labels were activated, and the distractor word Hirte enhanced the activation of the competitor Hirn due to phonological overlap and slowed down the response Schädel. According to Jescheniak and Schriefers (1998) this result supports cascaded models of language production that allow for phonological activation of multiple lexical candidates and contradicts discrete models that do not allow this possibility.

Levelt et al. (1999) argued that Peterson and Savoy's (1998) and Jescheniak and Schriefers' (1998) findings may be accommodated in their discrete theory by making an additional assumption. Arguably, ambiguous, near-synonym pictures are exceptional cases in the languages such as English and German. Therefore, by assuming the system may under exceptional circumstances select multiple lexical candidates that are appropriate for a given situation and phonologically encode those lexical candidates, the findings can in principle be accommodated by discrete models. Word pairs such as couch and sofa or Schäfer and Hirte may be one such exceptional case, sub- and superordinates such as tulip and flower may be another. That is, Levelt et al. (1999) suggested that under certain rather exceptional situations two lexical nodes may be selected and both may activate their respective phonological representations. In fact, Roelofs (1992) made a similar proposal to account for blending errors, that is, speech errors that are the result of blending two words, such as "terrible", a blend between terrible and horrible.

The question then became whether phonological activation only occurs in those exceptional cases or whether it is a more general phenomenon in speech production. In their own exploration of those issues, Costa, Caramazza and Sebastián-Gallés (2000) employed cognates to test the generalizability of Jescheniak and Schriefers' (1998) and Peterson and Savoy's (1998) findings. To refer to the concept of a cat, Spanish-Catalan bilinguals would use the word gato in Spanish and gat in Catalan, that is, these two words are called cognates due to their substantial form overlap, whereas the Spanish word mesa and the Catalan word taula (both meaning table) are non-cognates. The cognate status of words is an important dimension regarding naming performance in other circumstances as well. Kohnert (2004) showed that an English-Spanish bilingual aphasic speaker named objects more accurately when their label was a cognate (e.g., lamp – lampara) than when it was not. Gollan and Acenas (2004) reported fewer TOT states for items corresponding to cognates in the two languages of the bilingual populations they tested. Costa, Santesteban and Caño (2005) proposed that the most straight-forward way to account for these different cognate effects is to assume that lexical and sub-lexical levels of representation are connected by interactive processing.

Under the assumption that different languages share the conceptual level and concepts activate words in both languages of a bilingual speaker in parallel, Costa et al. (2000) predicted that bilinguals would name cognates faster than non-cognates, whereas monolinguals – for whom this difference is irrelevant – would name both types of targets equally fast. The reason Costa et al. (2000) made this prediction is that upon being represented with the picture of a cat, Spanish-Catalan bilingual speakers would activate two lexical nodes, that is, gato and gat. According to cascaded models, but not discrete models of lexical processing, both lexical nodes should activate their phonological segments. Since cognates share a substantial amount of phonological segments, these segments should be activated more and hence the target should be named faster than is

the case for non-cognates, such as mesa and taula for table. Their results demonstrated that Spanish-Catalan bilinguals indeed named objects that correspond to cognates in Spanish and Catalan faster in Spanish than non-cognates. Monolingual speakers of Spanish did not show any difference between cognate and non-cognate targets, as expected. In a follow-up experiment, Costa et al. (2000) demonstrated that the effect is obtained in both languages of a bilingual speaker. This effect has become known as the cognate facilitation effect. In their conclusion, Costa et al. argue that their findings demonstrate the generalizability of Jescheniak and Schriefers' (1998) and Peterson and Savoy's (1998) results because bilingual speakers, presumably more than half of all the speakers world-wide (Grosjean, 2010), have two or more words in different languages for many concepts. Therefore, their data provide strong support for cascaded theories of lexical access. In a still more recent research, Costa et al. (2017) proposed a different theoretical perspective on these findings providing an alternative to the co-activation principle described above. Rather than invoking distinct lexica for each language with phonological activation occurring in parallel, this novel proposal highlights the role of the L1 in shaping the lexical phonological knowledge when learning an L2.

Another attempt to distinguish between discrete and cascaded models has been made by Cutting and Ferreira (1999) through naming homophonic pictures such as BALL. Ball may refer to a sport utility (ball_{toy}) or, alternatively, to a formal dancing event (ball_{social event}), that is, the word ball is a homophone: it refers to two different meanings whose labels sound identical. In one of their experiments, participants were presented at SOA -150 ms (i.e., shortly before the onset of the target picture) with auditory distractor words that were either semantically related to the depicted meaning (e.g., game – ball_{toy}) or the non-depicted meaning of the homophone (e.g., dance – ball_{social event}), or completely unrelated (e.g., hammer). The question was whether participants named the picture of a toy ball faster when the distractor word was related to the non-depicted meaning of the homophonic picture (i.e., dance related to [dance] ball) than when the distractor was unrelated. Moreover, Cutting and Ferreira's (1999) also investigated the effect of a distractor word semantically related to the depicted meaning (i.e., frisbee related to [toy] ball) on the naming latencies of the same picture at the same SOA relative to the unrelated condition. If phonological effects were not observed in the same time window as semantic effects, according to Cutting and Ferreira this may be taken as evidence for semantic and phonological processing having different time courses, supporting discrete models. However, if phonologically related distractors affect picture naming latencies at the same time window as semantically related distractors, that may be taken as evidence for cascaded processing because semantic and phonological processing overlap in time.

Cutting and Ferreira's (1999) findings revealed the naming of homophonic target picture names (e.g., ball) was faster – relative to an unrelated condition – when distractor words (e.g., dance) were semantically related to the non-depicted meaning of the homophonic target picture (e.g., ball_{toy}). The authors assumed that distractor words that were related to the non-depicted meaning (e.g., dance) of the homophonic label (e.g., ball) activated a cohort of meaning-related word forms (including ball_{social event}), which activated their corresponding lexical representations. These lexical representations in turn activated their corresponding word forms. This results in the homophonic word form ball receiving activation from two sources, that is, from the selected ball_{toy} and the non-selected ball_{social event}. Cutting and Ferreira (1999) argued that phonological processing was affected by semantically processed stimuli even though these stimuli are not semantically similar to the target. In the same time window, classic semantic interference effects (e.g., frisbee – ball_{toy}) were observed. Thus, phonological and semantic processing

of non-target lexical items seems to occur in a similar time window. According to Cutting and Ferreira, this suggests that semantic and phonological processes operate with overlapping time courses, supporting cascaded models of lexical access. However, Levelt et al. (1999) suggested an alternative account: the distractor word *dance* may have co-activated its semantic associate *ball_{social event}* both semantically and phonologically in the perceptual network upon processing the distractor word. The activated word form *ball* in the perceptual network may then directly pre-activate its corresponding word form in the production network, leading to faster picture naming latencies of the target *ball_{toy}*.

More experimental evidence for cascaded models of lexical processing comes from a study employing a picture–picture interference paradigm by Morsella and Miozzo (2002). The results of this cross-linguistic study demonstrated that irrelevant pictorial stimuli become phonologically encoded in a picture naming task. Two overlapping pictures, one in green and one in red, were presented, and participants were asked to engage in naming the green pictures as fast and as accurately as possible while ignoring the red ones. The labels of the green and red pictures were either unrelated (e.g., *BED_{green}* – *HAT_{red}*) or phonologically related (e.g., *BED_{green}* – *BELL_{red}*). Results revealed that naming latencies were significantly faster for the phonologically related pairs compared to the unrelated pairs in English. A control experiment in Italian, a language in which the exact same picture pairings were phonologically unrelated, showed no naming difference between the two conditions.

Morsella and Miozzo (2002) argued that cascaded models of lexical access are supported by their finding. These models hold that lexical nodes that have not been selected, that is, the label of the red distractor picture, would nevertheless activate their corresponding phonological representations. That way, the red picture label *BELL* may activate its phonological representation and the corresponding segments /b/, /ɛ/, /l/. When the speaker phonologically encodes the target label *BED*, some of its segments, that is, /b/ and /ɛ/, also receive activation from the distractor picture label. Therefore, their selection and encoding are facilitated, and ultimately the target word is produced faster than in a condition where the two picture labels are not phonologically unrelated. Cascaded, but not discrete models of lexical access, can accommodate Morsella and Miozzo's (2002) findings because they assume phonological activation of non-selected lexical candidates. Navarrete and Costa (2005) convincingly replicated and extended Morsella and Miozzo's (2002) findings with a different paradigm in Spanish (but see e.g., Jescheniak, Oppermann, Hantsch, Wagner, Mädebach and Schriefers, 2009 who failed to replicate Morsella and Miozzo's results; see also Jescheniak, Matushanskaya, Mädebach and Müller, 2014 as well as Matushanskaya, Mädebach, Müller and Jescheniak, 2016). For reviews of cascaded/interactive models of language production see Rapp and Goldrick (2000) and Vigliocco and Hartsuiker (2002).

This discussion made clear that ample evidence for discrete theories of lexical access on the one hand and for cascaded/interactive theories on the other hand is available. However, the extreme discreteness assumption that was originally part of of *WEAVER++* model (Levelt et al., 1999; Roelofs, 1997) has been dropped by Roelofs (2008) in what he calls limited cascading (see also Roelofs & Ferreira, 2019).

3. Lexical selection

Language production comprises the cognitive processes of activating and selecting the target word from the mental lexicon. The lexical selection process has often been assumed to be competitive, that is, multiple activated word candidates compete to be selected such that the activation of a candidate works against the selection of other candidates by delaying it or making it proportionately less likely. Our semantic-conceptual system activates a network of lexical entries that are semantically related to each other. Semantic speech errors in healthy individuals (Fromkin, 1971; Garrett, 1975) and in aphasic speakers (Caramazza & Hillis, 1990; Alario, Schiller, Domoto-Reilly & Caramazza, 2003) may be taken to support this assumption. If it is the case that more lexical candidates are activated (or activated to a higher degree), speakers need more time to select the target lexical item as compared to a situation with fewer competitors (or lower activation levels of the competitors) because the competition is harder to resolve. The target lexical candidate usually has a) the highest activation level and b) a sufficiently large difference in activation to the lexical item with the next-highest activation level at a particular point in time, and thus will be selected by the speaker. Lexical selection by competition has been operationalized in computational models such as WEAVER++ (Roelofs, 1997) by means of the so-called Luce ratio (Luce choice axiom; Luce, 1959). This measure computes the probability of a certain lexical item to be selected by dividing its level of activation by the level of activation of all words in the lexicon. If other lexical nodes are (highly) activated during the process of lexical selection, the selection of the target lexical node will be delayed.

To test the lexical-selection-by-competition hypothesis, researchers have employed a paradigm that was derived from the Stroop task (Stroop, 1935; for a comprehensive review see MacLeod, 1991) and already mentioned in the previous section. In the so-called picture-word interference paradigm (PWI), to-be-named pictures are accompanied by visually superimposed or auditorily presented interfering stimuli, so-called distractor words (Glaser, 1992; Glaser & Dünghoff, 1984; Glaser & Glaser, 1989; Posnansky & Rayner, 1977; Rosinski, Golinkoff, & Kukish, 1975). When target object and distractor word share the same semantic category (picture: BED, distractor: table), naming latencies are usually longer than when they do not (picture: BED, distractor: apple). This is called the semantic interference effect and has become a very well investigated effect. A recent meta-analysis by Bürki, Elbuy, Madec and Vasishth (2020) of more than 150 studies on semantic interference conducted over a period of the last 40 years demonstrated that the vast majority of the (published) studies in their meta-analysis showed a semantic interference effect with an average size of slightly more than 20 ms. However, the authors also concluded that, based on the available evidence, the nature of the effect cannot be fully determined and nor can the effect be used to model lexical access. However, this is fairly recent knowledge and historically the discussion unfolded in different directions, that we now describe.

Lexical selection by competition has been challenged by a different proposal, that is, the so-called Response Exclusion Hypothesis (REH; Mahon et al., 2007). The REH proposes that lexical selection is independent of the number and activation level of potential competitors. Furthermore, according to the REH, the effect of the spreading of activation is always facilitative, that is, there are no inhibitory consequences due to competition. Instead, lexical selection and potential delays in producing a word occur at a later (post-lexical) stage, that is, at the stage where production-ready representations enter an output or response buffer in which they will reside, ready to be produced. The distractor words in the PWI paradigm are assumed to have privileged access to the articulators and enter the output or response buffer prior to the picture; consequently, the

response buffer must be cleared before the picture's name (i.e., the target) can be produced. The response buffer is assumed to be sensitive to semantic information to be able to decide whether or not a particular distractor word is an appropriate response in each trial (e.g., the verb 'sleep' is not an appropriate response when the task is object naming, e.g., 'bed', and not action naming). If a distractor is not response-relevant, the output buffer must be cleared of non-response items. For instance, when representations of the target and the distractor correspond to different semantic categories, the distractor can be discarded as a possible response candidate early based on categorical information (Costa, Alario & Caramazza, 2005). Mahon and colleagues suggest that the semantic interference effect arises at a post-lexical processing level and reflects how fast a production-ready representation (e.g., the distractor word) can be excluded from the response buffer as a relevant response to the target picture (for additional experimental support of their argument, see e.g., Finkbeiner & Caramazza, 2006).

Mahon et al. (2007) supported their proposal by a large set of experimental data. They observed, for instance, a facilitation effect in object naming for the target – distractor combination BED – sleep (as compared to BED – shoot). The authors propose that this facilitation effect is the result of the spreading of activation from multiple semantic items (bed, sleep) and the fact that sleep was not a suitable candidate for production. However, when table was the distractor word, it would be suitable (or response-relevant) since participants' task was to name objects. Therefore, it would have remained longer in the response buffer and more difficult to be removed from it, resulting in longer naming latencies for the target BED.

The REH was supported by Dhooge and Hartsuiker (2011; see also Dhooge & Hartsuiker, 2010; 2012) who demonstrated that the distractor-frequency effect (Miozzo & Caramazza, 2003) – the finding that low-frequency distractors interfere more with picture naming than high-frequency distractor words – is likely to have a late origin during processing because it occurs in a delayed as well as in an immediate picture naming condition. In another study, Dhooge and Hartsuiker (2010) showed that the distractor-frequency effect disappeared when the distractor was presented long enough, that is, at least 200 ms, before the target picture. In that case, there is enough time to clear the response buffer before the target response is given.

It should be noted that the theoretical proposal of the REH has not remained unchallenged. For instance, in a combined conditional and free naming study, Aristei and Abdel-Rahman (2013) demonstrated that response-relevance and semantic-categorical relationship of distractor words did not contribute significantly beyond semantic similarity. They argue that lexical selection is a competitive process reflected by inhibitory effects in the PWI, however, semantic similarity goes beyond categorical relationships between words, suggesting a flexible architecture of the production system regarding semantic similarity. Taken together, the proposal of the REH led a heated and yet unresolved debate about whether or not lexical selection is by competition or by exclusion (e.g., Abdel Rahman & Melinger, 2009; Mahon, Garcea, & Navarrete, 2012; Melinger & Abdel-Rahman, 2013; Spalek, Damian, & Bölte, 2013; Starreveld, La Heij, & Verdonschot, 2013; Vieth, McMahan, & De Zubicaray, 2014).

Further evidence for the processes involved in lexical selection derives from research on bilingualism and multilingualism. Regarding bilingual speech production, one of the most intriguing research questions has been to elucidate how bilingual speakers select lexical items in the language they want to use. Although (intended) insertions from the other language of a bilingual speaker are attested (so-called code-switching), generally bilingual speakers are good at producing speech output in the intended language

or language-in-use (Costa & Santesteban, 2004). It is widely assumed that a common conceptual or semantic system activates corresponding words in both lexicons of a bilingual speaker. Two proposals have been made to account for how bilingual speakers achieve selecting words of the intended language. The first assumes that an external system proactively inhibits the lexical items of the language-not-in-use (Green, 1998). This solution assumes that lexical selection is language-non-specific, that is it considers lexical items of both languages, but because one lexicon is suppressed, a lexical item from the language-in-use is selected. The second proposal suggests that the process of lexical selection is language-specific and only considers lexical items of the intended language (Costa, Miozzo, & Caramazza, 1999).

Employing the cross-linguistic identity condition, Costa and Caramazza (1999) demonstrated that bilingual speakers select their lexical items through a language-specific lexical selection mechanism. Spanish-English bilinguals named target pictures (e.g., TABLE) in their L1 Spanish. Those pictures were accompanied by visual distractor words in either Spanish (e.g., mesa) or English (e.g., table). The language-non-specific lexical selection hypothesis predicts that the lexical selection mechanism considers all the lexical items that were activated, that is, mesa and table. Therefore, relative to an unrelated condition (e.g., dedo or finger), interference will be predicted in the identity condition because there is competition between two highly activated lexical items, namely mesa and table. In contrast, the language-specific lexical selection hypothesis predicts facilitation in the identity condition (relative to an unrelated condition) because mesa has received activation from both the target picture (TABLE) and the distractor word (either mesa or table), but only the lexicon of the language-in-use (i.e., Spanish) is considered for selection. Costa and Caramazza (1999) found clear support for the language-specific selection account, both when Spanish was the L1 and – in a second experiment employing English-Spanish bilinguals – when Spanish was the L2. In addition, the identity effect was much larger when the distractor word was in the same language (target: mesa – distractor: mesa) than when the distractor was in the other language (target: mesa – distractor: table). Presumably, this is due to an additional phonological priming effect.

The effects reported by Costa and Caramazza (1999) for Spanish-English bilinguals (high-proficient late L2 learners) were replicated by Costa et al. (1999) for high-proficient early Catalan-Spanish bilinguals, that is speakers with essentially two L1's, both in blocked (Catalan only) and mixed (Catalan and Spanish) naming conditions. This finding further supports previous results and extends them to different situations (e.g., L1 and L2 naming, blocked and mixed conditions), and they add further support to the language-specific selection hypothesis.

Costa and Santesteban (2004) reported data that are difficult to account for by the proactive lexical inhibition account (Green, 1998; see above) but in line with the language-specific selection hypothesis. In a series of language switching experiments, multilingual participants named pictures in their L1 or L2. Costa and Santesteban (2004) replicated in their first experiment the asymmetrical switch cost reported by Meuter and Allport (1999), that is switching from the L2 to the L1 is more costly than switching from the L1 to the L2. Inhibitory models have accounted for this effect by an inhibitory mechanism by which speakers try to inhibit the activation of the L1 lexical representations when the response language is their L2. However, when they need to switch back from the L2 to their L1, it takes more effort to reactively overcome this inhibition than vice versa, leading to larger and asymmetrical switch costs. In Experiments 2 and 3, high-proficient bilinguals did not show asymmetrical switch costs. This result has later been replicated with high-proficient Spanish-Basque early bilinguals

and Spanish-English late bilinguals, demonstrating that neither age-of-onset of L2 acquisition (see review in Sebastián-Gallés & Kroll, 2003) nor typological proximity (language similarity) seem to influence this effect. Spanish and Basque are typologically unrelated and the Spanish-English bilinguals acquired their L2 English at a later age (Costa, Santesteban & Ivanova, 2006). It could be argued that high-proficient bilinguals inhibit their two languages equally strong, thus, leading to symmetrical switch costs. However, it may also be the case that high-proficient bilinguals select their lexical items in a qualitatively different way compared to less proficient bilinguals. In Experiment 4, high-proficient Spanish-Catalan bilinguals named pictures in their L1 (Spanish) or their less-proficient L3 (English). According to the proactive inhibition account, switching costs should be asymmetric in this condition, however, the results showed symmetric switching costs (for a replication see Costa et al., 2006, Experiment 2). This result cannot be accounted for by the inhibitory, language-nonspecific selection account. However, Costa and Santesteban (2004) suggest that high-proficient bilinguals shift from a reliance on inhibitory control to reliance on a mechanism based on language-specific lexical selection. Importantly, they do this not only for their dominant languages but also for their less dominant languages.

However, this last conclusion does not seem to be completely correct. Costa et al. (2006) failed to replicate the “shift of mechanism”-hypothesis with high-proficient Catalan-Spanish bilinguals who switched between their less-proficient L3 (English) and their even less-proficient L4 (French). These high-proficient bilinguals demonstrated asymmetrical switch costs. The same was found when high-proficient bilinguals who were learning a third language and were in their initial stage to do so switched between their L1 and recently learned new words from the L3. Costa et al. (2006) argued that the presence or absence of asymmetrical switch costs in high-proficient bilinguals depends on how well the weak language is involved in the task. For instance, in a newly learned language or a low-proficient L4 (Experiments 3 and 4 of Costa et al., 2006), there may not be a language-specific lexicon to focus on yet. However, this account is of tentative character and more research is needed to arrive at more definitive conclusions (see, for instance, Declerck & Koch, 2022, in press, who argue that inhibitory control of the non-target language may provide a coherent framework for bilingual language production).

4. Summary and conclusions

We presented discrete and cascaded/interactive theories of speech production and discussed experimental evidence presented for and against them roughly over the years when Albert Costa contributed to these debates. We also discussed lexical access in language production extensively, an area that is still much investigated and debated. Our discussion has shown that a strict serial model seems difficult to rhyme with the available evidence that points in the direction of cascaded activation during lexical access in speech production.

Albert Costa and his colleagues have worked on issues related to lexical access, and he significantly contributed to this area of research. Through his work on bilingualism, Albert Costa added an important dimension to the discussion. Costa’s language-specific lexical selection hypothesis gained a lot of experimental support, however, as he himself admitted, there are still some loose ends and more work to be done in the future.

References

- Abdel Rahman, R., & Melinger, A. (2009). Semantic context effects in language production: a swinging lexical network proposal and a review. *Language and Cognitive Processes*, 24, 713-734. <http://dx.doi.org/10.1080/01690960802597250>
- Alario, F.-X., Schiller, N. O., Domoto-Reilly, K., & Caramazza, A. (2003). The role of phonological and orthographic information in lexical selection. *Brain and Language*, 84, 372-398. [http://dx.doi.org/10.1016/S0093-934X\(02\)00556-4](http://dx.doi.org/10.1016/S0093-934X(02)00556-4)
- Aristei, S., & Abdel-Rahman, R. (2013). Semantic interference in language production is due to graded similarity, not response relevance. *Acta Psychologica*, 144, 571-582. <http://dx.doi.org/10.1016/j.actpsy.2013.09.006>
- Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: implications for models of lexical access in language production. *Journal of Memory and Language*, 48, 468-488. [http://dx.doi.org/10.1016/S0749-596X\(02\)00503-X](http://dx.doi.org/10.1016/S0749-596X(02)00503-X)
- Bock, K. (1995). Sentence production: From mind to mouth. In J. L. Miller & P. D. Eimas (Eds.), *Handbook of perception and cognition (2nd ed.)*, Vol. 11. *Speech, language, and communication* (p. 181-216). New York: Academic Press.
- Bock, K., & Levelt, W. J. M. (1994). Language production: grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945-984). New York: Academic Press.
- Browman, C. P., & Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3, 219-252.
- Brown, R., & McNeill, D. (1966). The 'tip-of-the-tongue' phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 5, 325-337. [http://dx.doi.org/10.1016/S0022-5371\(66\)80040-3](http://dx.doi.org/10.1016/S0022-5371(66)80040-3)
- Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical Representation and Process* (pp. 108-135). Cambridge, MA: MIT Press.
- Bürki, A., Elbuy, S., Madec, S., & Vasissth, S. (2020). What did we learn from forty years of research on semantic interference? A Bayesian meta-analysis. *Journal of Memory and Language*, 114, 104125. <http://dx.doi.org/10.1016/j.jml.2020.104125>
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14, 177-208. <http://dx.doi.org/10.1080/026432997381664>
- Caramazza, A. & Hillis, A. E. (1990). Where Do Semantic Errors Come From? *Cortex*, 26, 95-122. [http://dx.doi.org/10.1016/S0010-9452\(13\)80077-9](http://dx.doi.org/10.1016/S0010-9452(13)80077-9)
- Costa, A., Alario, F.-X., & Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychonomic Bulletin and Review*, 12, 125-131. <http://dx.doi.org/10.3758/BF03196357>
- Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish-English and English-Spanish

- bilinguals. *Bilingualism: Language and Cognition*, 2, 231-244. <http://dx.doi.org/10.1017/S1366728999000334>
- Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1283-1296. <http://dx.doi.org/10.1037/0278-7393.26.5.1283>
- Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, 41, 365-397. <http://dx.doi.org/10.1006/jmla.1999.2651>
- Costa, A., Pannunzi, M., Deco, G., Pickering, M. J. (2017). Do bilinguals automatically activate their native language when they are not using it? *Cognitive Science*, 41, 1629-1644. <http://dx.doi.org/10.1111/cogs.12434>
- Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly-proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50, 491-511. <http://dx.doi.org/10.1016/j.jml.2004.02.002>
- Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language*, 94, 94-103. <http://dx.doi.org/10.1016/j.bandl.2004.12.002>
- Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 1057-1074. <http://dx.doi.org/10.1037/0278-7393.32.5.1057>
- Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences*, 106, 21442-21446. <http://dx.doi.org/10.1073/pnas.0908921106>
- Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 318-344. <http://dx.doi.org/10.1037/0278-7393.25.2.318>
- Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 345-361. <http://dx.doi.org/10.1037/0278-7393.25.2.345>
- Declerck, M., & Koch, I. (2022). The concept of inhibition in bilingual control. *Psychological Review*. Advance online publication. <https://doi.org/10.1037/rev0000367>
- Deese, J. (1984). *Thought into speech: The psychology of a language*. Englewood Cliffs, NJ: Prentice-Hall.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283-321. <http://dx.doi.org/10.1037/0033-295X.93.3.283>
- Dell, G. S. (1988). The retrieval of phonological forms in production: tests of predictions from a connectionist model. *Journal of Memory and Language*, 27, 124-142. [http://dx.doi.org/10.1016/0749-596X\(88\)90070-8](http://dx.doi.org/10.1016/0749-596X(88)90070-8)

- Dell, G. S., Nozari, N., & Oppenheim, G. M. (2014). Word production: Behavioral and computational considerations. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds.), *The Oxford Handbook of Language Production* (pp. 88–104). Oxford, New York: OUP.
- Dell, G. S., & O’Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: a comment on Levelt et al. (1991). *Psychological Review*, 98, 604-614. <http://dx.doi.org/10.1037/0033-295X.98.4.604>
- Dell, G. S., & O’Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, 42, 287-314. [http://dx.doi.org/10.1016/0010-0277\(92\)90046-K](http://dx.doi.org/10.1016/0010-0277(92)90046-K)
- Dhooge, E., & Hartsuiker, R. J. (2010). The distractor frequency effect in picture-word interference: Evidence for response exclusion. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 36, 878-891. <http://dx.doi.org/10.1037/a0019128>
- Dhooge, E., & Hartsuiker, R. J. (2011). The distractor frequency effect in a delayed picture–word interference task: Further evidence for a late locus of distractor exclusion. *Psychonomic Bulletin & Review*, 18, 116-122. <http://dx.doi.org/10.3758/s13423-010-0026-0>
- Dhooge, E., & Hartsuiker, R. J. (2012). Lexical selection and verbal self-monitoring: Effects of lexicality, context, and time pressure in picture-word interference. *Journal of Memory and Language*, 66, 163-176. <http://dx.doi.org/10.1016/j.jml.2011.08.004>
- Finkbeiner, M., & Caramazza, A. (2006). Now you see it, now you don’t: On turning semantic interference into facilitation in a Stroop-like task. *Cortex*, 42, 790-796. [http://dx.doi.org/10.1016/S0010-9452\(08\)70419-2](http://dx.doi.org/10.1016/S0010-9452(08)70419-2)
- Fitch, W. T. & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, 1316, 87-104. <http://dx.doi.org/10.1111/nyas.12406>
- Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 47, 27-52. <http://dx.doi.org/10.2307/412187>
- Fromkin, V. A. (1973). *Speech Errors as Linguistic Evidence*. The Hague: Mouton.
- Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation*, vol. 9 (pp. 133-177). San Diego: Academic Press. [http://dx.doi.org/10.1016/S0079-7421\(08\)60270-4](http://dx.doi.org/10.1016/S0079-7421(08)60270-4)
- Garrett, M. F. (1976). Syntactic processes in sentence production. In R. J. Wales & E. Walker (Eds.), *New Approaches to Language Mechanisms* (pp. 231-256). Amsterdam: North Holland.
- Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language Production: Speech and Talk*, vol. 1 (pp. 177-220). New York: Academic Press.
- Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19-76). Orlando/London: Academic Press.

- Garrett, M. F. (1988). Processes in language production. In E. F. J. Newmeyer (Ed.), *Language: Psychological and Biological Aspects* (pp. 69-96). Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511621062.004>
- Glaser, W. R. (1992). Picture naming. *Cognition*, 42, 61-105. [http://dx.doi.org/10.1016/0010-0277\(92\)90040-O](http://dx.doi.org/10.1016/0010-0277(92)90040-O)
- Glaser, W. R., & Döngelhoff, F.-J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 640-654. <http://dx.doi.org/10.1037/0096-1523.10.5.640>
- Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, 118, 13-42. <http://dx.doi.org/10.1037/0096-3445.118.1.13>
- Goldrick, M. (2006). Limited interaction in speech production: Chronometric, speech error, and neuropsychological evidence. *Language and Cognitive Processes*, 21, 817-855. <http://dx.doi.org/10.1080/01690960600824112>
- Goldstein, L. & Fowler, C. A. (2003). Articulatory Phonology: A phonology for public language use. In Schiller, N. O. & Meyer, A. S. (Eds.), *Phonetics and phonology in language comprehension and production* (pp. 159-208). Berlin: Mouton de Gruyter. <http://dx.doi.org/10.1515/9783110895094.159>
- Gollan, T. H., & Acenas, L. A. (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, 246-269. <http://dx.doi.org/10.1037/0278-7393.30.1.246>
- Goodglass, H., Kaplan, E., Weintraub, S., & Ackerman, N. (1976). The ‘tip-of-the-tongue’ phenomenon in aphasia. *Cortex*, 12, 145-153. [http://dx.doi.org/10.1016/S0010-9452\(76\)80018-4](http://dx.doi.org/10.1016/S0010-9452(76)80018-4)
- Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67-81. <http://dx.doi.org/10.1017/S1366728998000133>
- Griffin, Z. M., & Ferreira, V. S. (2006). Properties of spoken language production. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics*, 2nd ed. (pp. 21-59). Amsterdam: Elsevier. <http://dx.doi.org/10.1016/B978-012369374-7/50003-1>
- Grosjean, F. (2010). *Bilingual. Life and reality*. Cambridge, MA: Harvard University Press.
- Guenther, F. H. (2003). Neural control of speech movements. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production* (pp. 209-240). Berlin: Mouton de Gruyter. <http://dx.doi.org/10.1515/9783110895094.209>
- Guenther, F. H. (2016). *Neural control of speech*. Cambridge, MA: The MIT Press.
- Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, 5, 67-103. <http://dx.doi.org/10.1080/02643298808252927>

- Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, 2, 255. <http://dx.doi.org/10.3389/fpsyg.2011.00255>
- Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, 92, 101-144. <http://dx.doi.org/10.1016/j.cognition.2002.06.001>
- Jescheniak, J.D., Matushanskaya, A., Mädebach, A., & Müller, M. M. (2014). Semantic interference from distractor pictures in single-picture naming: evidence for competitive lexical selection. *Psychonomic Bulletin & Review*, 21, 1294-1300. <http://dx.doi.org/10.3758/s13423-014-0606-5>
- Jescheniak, J. D., Oppermann, F., Hantsch, A., Wagner, V., Mädebach, A., & Schriefers, H. (2009). Do perceived context pictures automatically activate their phonological code? *Experimental Psychology*, 56, 56-65. <http://dx.doi.org/10.1027/1618-3169.56.1.56>
- Jescheniak, J. D., & Schriefers, H. (1998). Discrete serial versus cascaded processing in lexical access in speech production: further evidence from coactivation of near-synonyms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1256-1274. <http://dx.doi.org/10.1037/0278-7393.24.5.1256>
- Jescheniak, J. D., & Schriefers, H. (2001). Priming effects from phonologically related distractors in picture-word interference. *Quarterly Journal of Experimental Psychology*, 54A, 371-382. <http://dx.doi.org/10.1080/713755981>
- Kohnert, K. (2004). Cognitive and cognate-based treatments for bilingual aphasia: A case study. *Brain and Language*, 91, 294-302. <http://dx.doi.org/10.1016/j.bandl.2004.04.001>
- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112-146). New York: Wiley.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, 44, 1-22. [http://dx.doi.org/10.1016/0010-0277\(92\)90038-J](http://dx.doi.org/10.1016/0010-0277(92)90038-J)
- Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, 3, 223-232. [http://dx.doi.org/10.1016/S1364-6613\(99\)01319-4](http://dx.doi.org/10.1016/S1364-6613(99)01319-4)
- Levelt, W. J. M. (2001). Spoken word production: a theory of lexical access. *Proceedings of the National Academy of Sciences*, 98, 13464-13471. <http://dx.doi.org/10.1073/pnas.231459498>
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1-75. <http://dx.doi.org/10.1017/S0140525X99001776>
- Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: a study of picture naming. *Psychological Review*, 98, 122-142. <http://dx.doi.org/10.1037/0033-295X.98.1.122>

- Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. New York: Wiley.
- MacKay, D. G. (1987). *The Organization of Perception and Action: A Theory for Language and Other Cognitive Skills*. New York: Springer.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163-203. <http://dx.doi.org/10.1037/0033-2909.109.2.163>
- MacNeilage, P. F. (2008). *The Origin of Speech*. Oxford: Oxford University Press.
- Mahon, B. Z., Costa, A., Peterson, R., Vargas, K., & Caramazza, A. (2007). Lexical selection is not by competition: a reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 503-535. <http://dx.doi.org/10.1037/0278-7393.33.3.503>
- Mahon, B. Z., Garcea, F. E., & Navarrete, E. (2012). Picture-word interference and the response exclusion hypothesis: a response to Mulatti and Coltheart. *Cortex*, *48*, 373-377. <http://dx.doi.org/10.1016/j.cortex.2011.10.008>
- Matushanskaya, A., Mädebach, A., Müller, M. M., & Jescheniak, J. D. (2016). When sufficiently processed, semantically related distractor pictures hamper picture naming: Implications for models of lexical access in speech production. *Experimental Psychology*, *63*, 307-317. <http://dx.doi.org/10.1027/1618-3169/a000340>
- Melinger, A., & Abdel Rahman, R. (2013). Lexical selection is competitive: Evidence from indirectly activated semantic associates during picture naming. *Journal of Experimental Psychology: Learning Memory and Cognition*, *39*, 348-364. <http://dx.doi.org/10.1037/a0028941>
- Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*, 25-40. <http://dx.doi.org/10.1006/jmla.1998.2602>
- Meyer, A. S. (1990). The time course of phonological encoding in language production: the encoding of successive syllables of a word. *Journal of Memory and Language*, *29*, 524-545. [http://dx.doi.org/10.1016/0749-596X\(90\)90050-A](http://dx.doi.org/10.1016/0749-596X(90)90050-A)
- Meyer, A. S. (1991). The time course of phonological encoding in language production: phonological coding inside a syllable. *Journal of Memory and Language*, *30*, 69-89. [http://dx.doi.org/10.1016/0749-596X\(91\)90011-8](http://dx.doi.org/10.1016/0749-596X(91)90011-8)
- Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in picture-word interference paradigm. *Journal of Experimental Psychology: General*, *132*, 228-252. <http://dx.doi.org/10.1037/0096-3445.132.2.228>
- Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *28*, 555-563. <http://dx.doi.org/10.1037/0278-7393.28.3.555>
- Munding, D., Dubarry, A.-S., & Alario, F.-X. (2016). On the cortical dynamics of word production: a review of the MEG evidence. *Language, Cognition and Neuroscience*, *31*, 441-462. <http://dx.doi.org/10.1080/23273798.2015.1071857>

- Navarrete, E., & Costa, A. (2005). Phonological activation of ignored pictures: further evidence for a cascade model of lexical access. *Journal of Memory and Language*, 53, 359-377. <http://dx.doi.org/10.1016/j.jml.2005.05.001>
- 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, 227-252. <http://dx.doi.org/10.1016/j.cognition.2009.09.007>
- Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 24, 539-557. <http://dx.doi.org/10.1037/0278-7393.24.3.539>
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169-226. <http://dx.doi.org/10.1017/S0140525X04000056>
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36, 329-392. <http://dx.doi.org/10.1017/S0140525X12001495>
- Posnansky, C. J., & Rayner, K. (1977). Visual-feature and response components in a picture-word interference task with beginning and skilled readers. *Journal of Experimental Child Psychology*, 24, 440-460. [http://dx.doi.org/10.1016/0022-0965\(77\)90090-X](http://dx.doi.org/10.1016/0022-0965(77)90090-X)
- Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, 107, 460-499. <http://dx.doi.org/10.1037/0033-295X.107.3.460>
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107-142. [http://dx.doi.org/10.1016/0010-0277\(92\)90041-F](http://dx.doi.org/10.1016/0010-0277(92)90041-F)
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249-284. [http://dx.doi.org/10.1016/S0010-0277\(97\)00027-9](http://dx.doi.org/10.1016/S0010-0277(97)00027-9)
- Roelofs, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: verb-particle combinations. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 24, 904-921. <http://dx.doi.org/10.1037/0278-7393.24.4.904>
- Roelofs, A. (2000). WEAVER++ and other computational models of lemma retrieval and word-form encoding. In L. R. Wheeldon (Ed.), *Aspects of Language Production* (pp. 71-114). Hove: Psychology Press/Taylor and Francis.
- Roelofs, A. (2008). Tracing attention and the activation flow in spoken word planning using eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 353-368. <http://dx.doi.org/10.1037/0278-7393.34.2.353>
- Roelofs, A. (2018). A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition*, 172, 59-72. <http://dx.doi.org/10.1016/j.cognition.2017.12.007>
- Roelofs, A., & Ferreira, V. S. (2019). The architecture of speaking. In P. Hagoort (Ed.), *Human language: From genes and brains to behavior* (pp. 35-50). Cambridge, MA: The MIT Press. <http://dx.doi.org/10.7551/mitpress/10841.003.0006>

- Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture–word interference task. *Child Development*, *46*, 247-253. <http://dx.doi.org/10.2307/1128859>
- Schiller, N. O. (2006). Phonology in the production of words. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (pp. 545-553). Amsterdam: Elsevier. <http://dx.doi.org/10.1016/B0-08-044854-2/00075-4>
- Schiller, N. O. (2020). Neurolinguistic approaches in morphology. In R. Lieber (Ed.), *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. <http://dx.doi.org/10.1093/acrefore/9780199384655.013.601>
- Schiller, N. O. (2021). The mental lexicon. In M. Aronoff (Ed.), *Oxford Bibliographies in Linguistics*. Oxford: Oxford University Press.
- Schiller, N. O., Bles, M., & Jansma, B. M. (2003). Tracking the time course of phonological encoding in speech production: An event-related brain potential study on internal monitoring. *Cognitive Brain Research*, *17*, 819-831. [http://dx.doi.org/10.1016/S0926-6410\(03\)00204-0](http://dx.doi.org/10.1016/S0926-6410(03)00204-0)
- Schiller, N. O., & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language*, *48*, 169-194. [http://dx.doi.org/10.1016/S0749-596X\(02\)00508-9](http://dx.doi.org/10.1016/S0749-596X(02)00508-9)
- Schiller, N. O., & Verdonschot, R. G. (2015). Accessing words from the mental lexicon. In J. R. Taylor (Ed.), *The Oxford Handbook of the Word* (pp. 481-492). Oxford: Oxford University Press.
- Schiller, N. O. & Verdonschot, R. G. (2019). Morphological theory and neurolinguistics. In J. Audring & F. Masini (Eds.), *The Oxford Handbook of Morphological Theory* (pp. 554-572). Oxford: Oxford University Press.
- Schmitt, B. M., Münte, T. F., & Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology*, *37*, 473-484. <http://dx.doi.org/10.1111/1469-8986.3740473>
- Schriefers, H., Meyer, A., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: picture-word interference studies. *Journal of Memory and Language*, *29*, 86-102. [http://dx.doi.org/10.1016/0749-596X\(90\)90011-N](http://dx.doi.org/10.1016/0749-596X(90)90011-N)
- Schuhmann, T., Schiller, N. O., Goebel, R., & Sack, A. T. (2012). Speaking of which: Dissecting the neurocognitive network of language production in picture naming. *Cerebral Cortex*, *22*, 701-709. <http://dx.doi.org/10.1093/cercor/bhr155>
- Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, *54*, 228-264. <http://dx.doi.org/10.1016/j.jml.2005.10.001>
- Sebastián-Gallés, N., & Kroll, J. F. K. (2003). Phonology in bilingual language processing: Acquisition, perception and production. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production* (pp. 279-317). Berlin: Mouton de Gruyter. <http://dx.doi.org/10.1515/9783110895094.279>

- Spalek, K., Damian, M. F., & Bölte, J. (2013). Is lexical selection in spoken word production competitive? Introduction to the Special Issue on lexical competition in language production. *Language and Cognitive Processes*, 28, 597-614. <http://dx.doi.org/10.1080/01690965.2012.718088>
- Starreveld, P. A. (2000). On the interpretation of onsets of auditory context effects in word production. *Journal of Memory and Language*, 42, 497-525. <http://dx.doi.org/10.1006/jmla.1999.2693>
- Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 896-918. <http://dx.doi.org/10.1037/0278-7393.22.4.896>
- Starreveld, P. A., La Heij, W., & Verdonschot, R. G. (2013). Time course analysis of the effects of distractor frequency and categorical relatedness in picture naming: an evaluation of the response exclusion account. *Language and Cognitive Processes*, 28, 633-654. <http://dx.doi.org/10.1080/01690965.2011.608026>
- Stemberger, J. P. (1985). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the Psychology of Language* (pp. 143-186). Mahwah, NJ: Erlbaum.
- Strijkers, K., & Costa, A. (2011). Riding the lexical speedway: A critical review on the time course of lexical selection in speech production. *Frontiers in Psychology*, 2, 356. <http://dx.doi.org/10.3389/fpsyg.2011.00356>
- Strijkers, K., & Costa, A. (2016). The cortical dynamics of speaking: present shortcomings and future avenues. *Language, Cognition and Neuroscience*, 31, 484-503. <http://dx.doi.org/10.1080/23273798.2015.1120878>
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662. <http://dx.doi.org/10.1037/h0054651>
- Van Turennout, M., Hagoort, P., & Brown, C. M. (1997). Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 787-806. <http://dx.doi.org/10.1037/0278-7393.23.4.787>
- Van Turennout, M., Hagoort, P., & Brown, C. M. (1998). Brain activity during speaking: from syntax to phonology in 40 milliseconds. *Science*, 280, 572-574. <http://dx.doi.org/10.1126/science.280.5363.572>
- Vieth, H. E., McMahan, K. L., & De Zubicaray, G. I. (2014). Feature overlap slows lexical selection: Evidence from the picture–word interference paradigm. *Quarterly Journal of Experimental Psychology*, 67, 2325-2339. <http://dx.doi.org/10.1080/17470218.2014.923922>
- Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, 8, 314-317. <http://dx.doi.org/10.1111/j.1467-9280.1997.tb00444.x>
- Vigliocco, G., & Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin*, 128, 442-472. <http://dx.doi.org/10.1037/0033-2909.128.3.442>