Riding the lexical speedway: a critical review on the time course of lexical selection in speech production

Kristof Strijkers[1]* and Albert Costa[2]

[1] Departamento de Psicología Básica, Universitat de Barcelona, Barcelona, Spain
[2] Departamento de Tecnología, Institució Catalana de Recerca i Estudis Avançats, Universitat Pompeu Fabra, Barcelona, Spain

Speech requires time. How much time often depends on the amount of labor the brain has to perform in order to retrieve the linguistic information related to the ideas we want to express. Although most psycholinguistic research in the field of language production has focused on the net result of time required to utter words in various experimental conditions, over the last years more and more researchers pursued the objective to flesh out the time course of particular stages implicated in language production. Here we critically review these studies, with particular interest for the time course of lexical selection. First, we evaluate the data underlying the estimates of an influential temporal meta-analysis on language production (Indefrey and Levelt, 2004). We conclude that those data alone are not sufficient to provide a reliable time frame of lexical selection. Next, we discuss recent neurophysiological evidence which we argue to offer more explicit insights into the time course of lexical selection. Based on this evidence we suggest that, despite the absence of a clear time frame of how long lexical selection takes, there is sufficient direct evidence to conclude that the brain initiates lexical access within 200 ms after stimulus presentation, hereby confirming Indefrey and Levelt's estimate. In a final section, we briefly review the proposed mechanisms which could lead to this rapid onset of lexical access, namely automatic spreading activation versus specific concept selection, and discuss novel data which support the notion of spreading activation, but indicate that the speed with which this principle takes effect is driven by a top-down signal in function of the intention to engage in a speech act.

Keywords: language production, lexical selection, time course, behavioral chronometry, ERPs, MEG, review

INTRODUCTION

Speaking is one of the most practiced psycho-motor skills in humans. In fact, we are so well practiced in it that it seems we produce language almost effortlessly. However, even in order to utter a single word, a speaker must carry out and orchestrate a number of mental operations such as retrieve and select the concept she/he intends to express, translate it into the appropriate word and sounds, and prepare the articulatory apparatus involving over 200 muscles (e.g., Dell, 1986; Levelt, 1989; Caramazza, 1998; Levelt et al., 1999). Moreover, the space to choose the words from is large with the storage of thousands of different lexical entries linked to different meanings. Despite the sophisticated nature underlying speech production, we are able to utter about three words per second while solely producing about one error every 1000 words (e.g., Levelt, 1989).

These observations call into mind two questions which will be the focus of the current review, namely "exactly how fast do we access words from the mental lexicon?" and "what is the mechanism that makes rapid lexical selection possible?". Answering the first question is our main objective, and to do so we will walk through the chronometric evidence on speech production in order to determine whether we have a reliable estimate of when the brain engages in lexical selection and how long the process lasts. To this end, we will critically review the studies that constitute the basis of the temporal model of speech production proposed by Indefrey and Levelt (2004; see also Indefrey, 2011). Next, some more recent explorations into the time course of lexical selection will be discussed and contrasted with Indefrey and Levelt's work. To advance upon the conclusion of this part, we will argue that there is sufficient evidence gathered by now indicating that lexical access in picture naming initiates within 200 ms after stimulus onset, just as predicted by the meta-analysis of Indefrey and Levelt (2004). At the same time, we will argue that a similar reliable estimate for the duration of lexical selection is still lacking. Finally, in a last and more concise part, some recent findings concerning the mechanism responsible for this rapid initiation of lexical access will be reported, addressing the second question.

HOW FAST DOES THE BRAIN ACCESS WORDS FROM THE LEXICON?

INDEFREY AND LEVELT'S TEMPORAL MODEL OF SINGLE WORD PRODUCTION

In general, psycholinguists agree that uttering a word entails at least four mental operations prior to actual articulation (see **Figure 1**; e.g., Fromkin, 1971; Garrett, 1975; Bock, 1982; Dell, 1986; Levelt, 1989; Dell and O'Seaghdha, 1992; Roelofs, 1992, 1997; Caramazza, 1997; Levelt et al., 1999). The first one, conceptual processing, refers to the retrieval of the semantic information behind the idea

FIGURE 1 | Simple and schematic model of object naming.
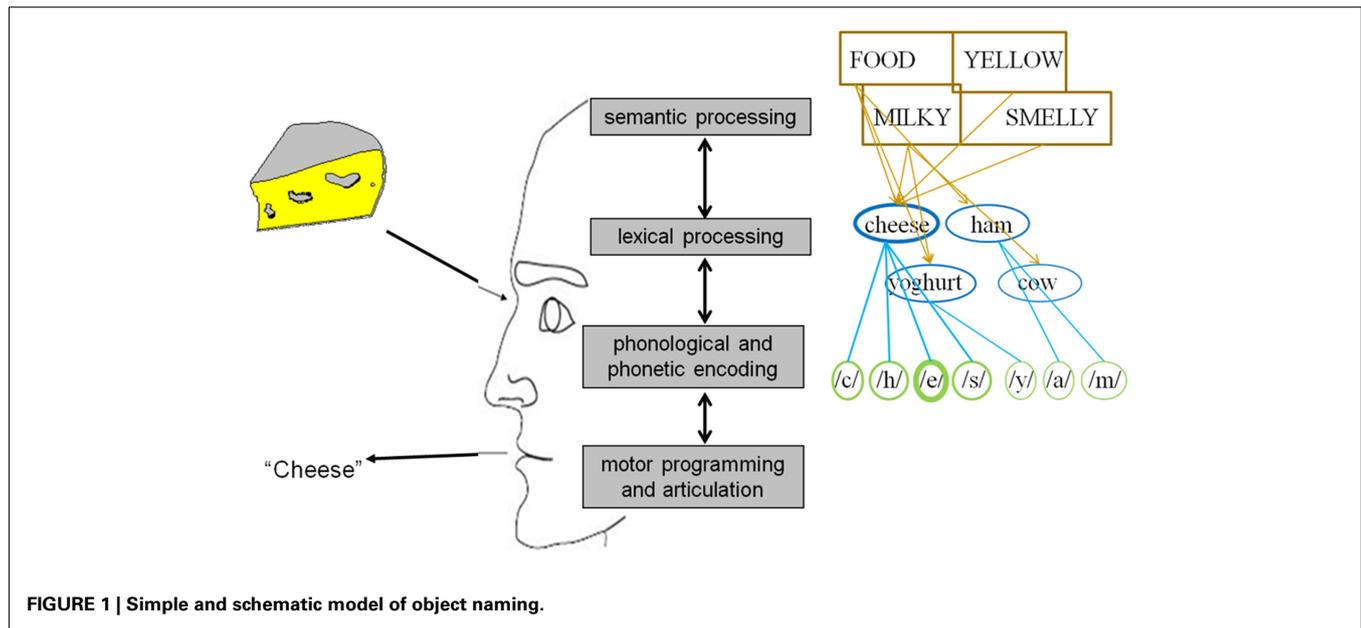
we want to express which the parser can map onto a specific word. The second operation is lexical selection, which refers to the search and selection within the mental lexicon of the syntactical and morphological properties of words. Thirdly, phonological encoding can be defined as the retrieval of the sounds corresponding to the word(s) we want to utter. And finally, during motor preparation an articulatory program is specified which enables us to eventually produce the intended word. In 2004 Indefrey and Levelt (henceforth referred to as I&L) published a comprehensive meta-analysis regarding the spatial and temporal signatures of these core operations involved in single word production. For our purposes we will leave the brain sources aside and focus on the chronometry. In their study, which can be considered as the first systematic exploration of the time course of language production, the authors integrated all temporal evidence available at that moment. As a result of this analysis the following map emerged for an average naming latency of 600 ms: object recognition and conceptual processing: 0–175 ms; lexical selection (lemma retrieval): 175–250 ms; morphological encoding (lexeme retrieval): 250–330 ms; syllabification (post-lexical phonological encoding): 330–455 ms; motor programming and articulation onset: 455–600 ms.

A clear strength of this meta-analysis is that it bridged recent neuroscientific evidence with an important psychological model of speech production (Roelofs, 1992, 1997; Levelt et al., 1999). This offers a framework to explore concrete predictions of when the brain will engage in a certain type of linguistic operation. However, without reducing the merit of I&L's work, one must be cautious how to use this map. For instance, various recent studies have explicitly relied on I&L's meta-analysis to interpret temporal data in terms of processing stages (e.g., Christoffels et al., 2007; Habets et al., 2008; Laganaro et al., 2009, 2010; Dell'acqua et al., 2010; Aristei et al., 2011). Nevertheless, when doing so, one must recall that the temporal map of I&L is still hypothetical. This is because of two independent reasons: (1) On the methodological side, most of the evidence on which I&L's work is based stems

from studies which did not involve speech production. Furthermore, as we will discuss later, the complexity associated with the methodology of these empirical studies allows for several alternative interpretations, some of which may have little interest from the perspective of language production. (2) On the theoretical side, both the temporal map of I&L and some of the chronometric studies on which it is based used the speech production model of Levelt et al. (1999) as a starting point to funnel the assignment of the temporal estimates. While this is a valid and productive strategy to provide a working frame concerning the time course of the various mental operations involved in producing speech, it restricts the estimates to a single theoretical view. Therefore, I&L's approach needs to be complemented by explicit estimates stemming from more transparent designs for speech production and from phenomena which are not necessarily bound to a single theory. In what follows, we will first critically review the empirical basis of I&L's meta-analysis, which is structured according to methodology. Afterward, we will discuss more recent experimental work examining the time course of lexical selection and assess how well the novel temporal data corresponds to the estimates of I&L.

THE EMPIRICAL BASIS OF I&L META-ANALYSIS
Behavioral chronometry
The first studies which can be considered to be informative regarding the time course of lexical selection are the picture–word interference (PWI) experiments with varying stimulus-onset-asynchrony (SOA). In these experiments, participants have to name pictures and either prior to, simultaneously with or after picture presentation they hear or see a word which must be ignored. The crucial manipulation is that the distractor can have a linguistic relationship with the target picture. Typically what is observed is that semantically related distractors (e.g., *cat* for the target *DOG*) result in interference compared to unrelated distractors (e.g., *chair*), while phonologically related distractors (e.g., *doll*

for the target *DOG*) give rise to facilitation effects (e.g., Lupker, 1979, 1982; Glaser and Düngelhoff, 1984; Mägiste, 1984; Glaser and Glaser, 1989). These results are thought to reveal lexical competition for the semantic interference (SI) effect due to spreading activation and speeded word form encoding for the phonological facilitation (PF) effect due to parallel activation of overlapping phonemes. Schriefers et al. (1990) explored the sequence in time of SI and PF effects by varying the SOA of the auditory distractors. They detected that at a SOA of −150 ms there were SI effects but no PF effects. At SOA's of 0 and +150 ms there were no SI effects, but PF effects. In sum, although no exact temporal information can be derived, one may argue that a maximal lag of 150 ms passes between the onset of lexical selection and phonological encoding.

A more precise estimate was obtained from the follow-up study conducted by Levelt et al. (1991). The authors measured lexical decision latencies on the *distractors* as independent measure. That is, participants were placed in a dual-task situation in which they had to utter the names of pictures on every trial, but on one-third of all trials they had to perform a lexical decision for an acoustic *distractor* word (or non-word) which appeared after picture presentation at varying SOAs. Just as in Schriefers et al.'s (1990) study, effects of semantic relatedness were found for short SOAs (47 ms), but not for intermediate (73 ms) or long SOAs (107 ms). In contrast, effects of phonological relatedness were present across the board. The authors constructed a mathematical model of the dual-task situation to extract which temporal settings for the lexical and the phonological stage would best mimic the experimental data. A model with a time window of 115 ms for lexical selection and 270 ms for phonological encoding gave the most approximate rendition of the behavioral data (see also Roelofs, 1997). Converging evidence for this duration estimate was demonstrated by Roelofs (1992) in a similar modeling exercise aimed to reproduce the PWI data reported by Glaser and Düngelhoff (1984). A latency of 100–150 ms for lexical selection produced an accurate replication of the SI effects.

Taken together, the chronometric data reviewed above converge to a temporal separation of ∼ 100–150 ms between the initial access to lexical representations and the start of word form retrieval. However, the mathematical models constructed to estimate the latency of lexical selection assumed *a priori* discrete delays (e.g., Levelt et al., 1991; Roelofs, 1997). Consequently, demonstrating with such model that parameter estimates of 100–150 ms closely match the behavioral data, does not mean that other types of architectures would fail to reconstruct the RT patterns (cascaded networks; e.g., Dell and O'Seaghdha, 1991, 1992). A second concern relates to the questionable PWI-like paradigms employed. PWI paradigms have a longstanding tradition in psycholinguistic research, but the typical direction of the SI and PF effects seems to vary strongly in function of strength and type of linguistic relationship being manipulated as well as the modality in which targets and distractors are presented (e.g., Starreveld and La Heij, 1995, 1996; Alario et al., 2000; Bloem and La Heij, 2003; Bloem et al., 2004; Costa et al., 2005; Mahon et al., 2007; Janssen et al., 2008). These findings resulted in skepticisms regarding the typical lexical locus of the effects (e.g., Miozzo and Caramazza, 2003; Costa et al., 2005; Finkbeiner and Caramazza, 2006; Finkbeiner et al., 2006;

Mahon et al., 2007; Janssen et al., 2008; but see e.g., Abdel Rahman and Melinger, 2009a,b). To sum up, there are both methodological and theoretical concerns surrounding the time course of lexical selection taken from PWI paradigms, which at least calls for the collection of additional evidence.

Button-press ERPs in language production tasks
More detailed data on the temporal progression of linguistic processing was gathered by studies employing the ERP technique. Prior to reporting some of these data, let us first dedicate a few words to the time course of conceptual processing which has been used to infer the onset of lexical access (Indefrey and Levelt, 2004). Many neurophysiological studies on object categorization converge that within 100–150 ms after object presentation a critical category has been activated (e.g., Thorpe et al., 1996; VanRullen and Thorpe, 2002; Johnson and Olshausen, 2005; Kirchner and Thorpe, 2006; Hauk et al., 2007; Liu et al., 2009). These results form the basis of Indefrey and Levelt's (2004) median estimate of 175 ms for conceptual identification and, consequently, the start of lexical access. However, the time needed to identify a super-ordinate category as *animal* may have an earlier onset compared to the identification of a basic-level concept (*lexical concept*) such as *dog*, *cat*, or *horse*[1]. In fact, ERP studies requiring more subtle semantic analyses of objects report time courses which are notably later compared to the above (∼250 ms; e.g., Holcomb and McPherson, 1994; Doniger et al., 2000; Eddy et al., 2006; Sitnikova et al., 2006). This does not mean that lexical access cannot initiate around 175 ms (e.g., Strijkers et al., 2011), but may rather index that the system is subject to cascading allowing activity to spread to a subsequent stage prior to selection at the previous stage[2]. If so, it is unlikely that by 175 ms conceptual processing is finished. More importantly, given that the onset of lexical access stems from object categorization data, there is much room from improvement by gathering estimates which are more closely related to lexical selection.

Now let us turn to those studies assessing the time course of linguistic processing with the ERP technique. Since the EEG signal is

---

[1]Indefrey and Levelt (2004) foresaw this potential problem and argued the following: given that 300 ms passed between Thorpe et al.'s (1996) point of category identification (150 ms) and the average button-press, those 300 ms were necessary to prepare and execute the motor response. When comparing to a similar picture recognition task with a similar button-press latency (439 ms) but where object identification was refined to a *lexical* concept (Jescheniak and Levelt, 1994), Indefrey and Levelt (2004) argued that subtracting the same 300 ms necessary for motor preparation and execution signifies that access to the basic-level concept occurred within the 150 ms time frame. However, to which extent 300 ms can be considered a reliable estimate of motor preparation and execution (note, for instance, that Thorpe et al. themselves estimate motor preparation to take about 150 ms) and to which extent that estimate can be paralleled between the two tasks is debatable. Animal models of object categorization show that motor preparation and execution takes about 70–120 ms (e.g., Thorpe and Fabre-Thorpe, 2001). For humans this latency should be longer, however the prolongation is estimated to fall within the range of 50 ms or less and not three or four times that amount (e.g., Thorpe and Fabre-Thorpe, 2001).

[2]Note that the theoretical framework underlying I&L's meta-analysis, namely Levelt et al.'s (1999) speech production model (see also Roelofs, 1992, 1997), does assume that conceptually driven lexical access is a cascaded process up to lemma selection. Thus, although the argument made is not problematic for the theory itself, the estimate of 175 ms for concept selection could occur later.

sensitive to motor-related activity, the initial approach for applying EEG measures to the field of language production was by designing tasks which relied on button-presses rather than actual speech. The first adaptations of EEG to button-press tasks exploring the time course of semantic, lexical, and phonological retrieval were conducted by van Turennout et al. (1997, 1998). These authors explored the presence of a brain potential linked to response preparation called the lateralized readiness potential (LRP; e.g., Coles, 1989; Miller et al., 1992) in several production-like experimental conditions. Participants were presented with colored line drawings of objects and were instructed to perform a syntactical–phonological categorization task (1998; semantic–phonological categorization task, 1997) prior to the articulation of a noun phrase that described the picture (e.g., the red bear). Classification consisted of a combined go/no-go and left/right button-press task. In the first experiment a phonological dimension defined the go/no-go response and a syntactical dimension (1998; semantic dimension, 1997) defined the response hand and vice versa for Experiment 2. With this elegant design, the authors observed that the LRP was present on both go and no-go trials when the decision was based on phonological information (e.g., does the picture name starts with the phoneme/b/), while the LRP was only present for go trials when the decision was based on syntactic information (e.g., does the picture name have feminine gender, 1998; semantic information, e.g., is it animal, 1997). The authors argued that the LRP asymmetry came about because syntax (semantic information, 1997) is available prior to phonological information during speech production: when the go/no-go decision is based on phonology, the hand response for syntax (semantics, 1997) is nonetheless prepared resulting in an LRP for both go and no-go trials. In contrast, when the go/no-go decision is based on syntax (semantics, 1997) no LRP for no-go trials is observed, because syntax (semantics, 1997) is retrieved before phonology and the decision that a no-go trial is present can be made prior to the activation of phonology.

Aside from providing evidence for the semantics/syntax to phonology sequentiality, the authors explored the latencies between the linguistic operations. By comparing the go with no-go LRPs in their first experiment (where syntax/semantics resulted in an LRP on both go and no-go trials) they could estimate the length of the time interval where syntactical/semantic information of the noun was available, but not yet the phonological information (the moment that information becomes accessible, go and no-go LRPs should diverge). In this way, the authors observed that: (a) syntax became available 40 ms before phonological information (van Turennout et al., 1998) and (b) 120 ms passed between semantic activation and the onset of phonological encoding (van Turennout et al., 1997). Since then, relatively well corresponding latencies between conceptual and phonological retrieval have been found using the same rationale but tracing the peak latency of the N200 component linked to response inhibition (e.g., Kok, 1986; Eimer, 1993) instead of or in combination with the LRP (e.g., Schmitt et al., 2000; Rodriguez-Fornells et al., 2002). In addition, the same N200 approach has been applied to estimate that conceptual activation precedes syntactical activation by 70–80 ms (e.g., Schmitt et al., 2001; Schiller et al., 2003). Taken together all latencies of the reported dual-task button-press ERP studies, the following picture

emerges: from input to concept: 175 ms; from concept to syntax: 75 ms; from syntax to first phoneme: 40 ms; and from concept to first phoneme: 115 ms (see **Figure 2A**).

The results obtained from these ERP studies have some clear advantages over the behavioral chronometry. For one, given the use of a fine-grained temporal measure as ERPs the latencies are more specific. Furthermore, the input (a single picture) is less convoluted compared to the primed or simultaneous presentation of pictures and words as in the PWI paradigm. Consequently, these ERP data have been given much theoretical weight in the field of language production. Nevertheless, and besides the limitation of the absence of actual speech, there are some factors we must take into account when assessing the reliability of these latency estimates. One such important factor seems to be task difficulty. For instance, Abdel Rahman and Sommer (2003), Abdel Rahman et al. (2003) explored whether the duration of semantic processing (conceptualized through the difficulty of the semantic categorization task) would affect the time frame of phoneme retrieval. Employing the same paradigm as that of van Turennout et al., 1997; semantic versus phonological classification), they managed to replicate the original results when the semantic classification was easy. But when the semantic classification was hard (herbivore versus omnivore), they failed to observe a no-go LRP for phonological information. Furthermore, when the go/no-go decision was based on the semantic dimension, the LRP onset latency for go trials (phoneme identification) was unaffected by the difficulty of
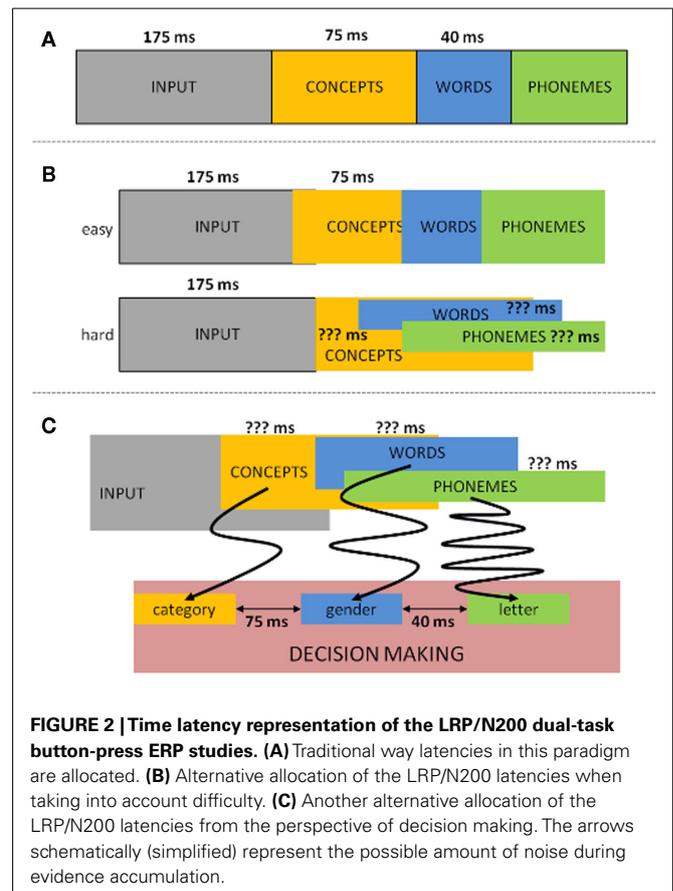

FIGURE 2 | Time latency representation of the LRP/N200 dual-task button-press ERP studies. (A) Traditional way latencies in this paradigm are allocated. **(B)** Alternative allocation of the LRP/N200 latencies when taking into account difficulty. **(C)** Another alternative allocation of the LRP/N200 latencies from the perspective of decision making. The arrows schematically (simplified) represent the possible amount of noise during evidence accumulation.

semantic processing, indicating that semantic retrieval occurred in parallel with phonological retrieval (Abdel Rahman and Sommer, 2003; Abdel Rahman et al., 2003). In other words, when taking into account the factor task difficulty the time course of conceptualization can actually overlap with that of phonological encoding (see **Figure 2B**), which shows that comparing the time course between these different dimensions may not be as straightforward as originally thought.

But one can question whether the paradigm in general is sensitive to real-time language components. The tasks employed are complex requiring two different meta-linguistic decisions linked to different responses. One cannot exclude that the extra orthogonal cognitive operations necessary to perform these tasks affect the natural time course of speech production. In fact, when considering the onsets of the LRP and N200 effects, these do not seem to fit the temporal frame of the supposed underlying processes. In van Turennout et al. (1998) study the syntax and phonology LRPs starts at 370 and 410 ms after picture presentation respectively. Similarly, N200 peak latencies for conceptual, syntactical, or phonological go/no-go difference waves occur around (averaged over studies) 410, 500, and 520 ms after stimulus onset respectively. These ERP onsets do not seem to correspond with the time frames of the linguistic processes in which they are supposed to tap into, but rather seem to correspond with response selection processes. If so, how representative are the latencies for the word production components? This question can be countered on different grounds. One is by assuming that the individual latencies reflect the on-line linguistic process, but that they are delayed in time due to additional processing (visual/conceptual or non-speech related demands) induced by the complex task. In that case, the ERP onsets of a linguistic dimension are not informative for the time course of natural speech, but the distance between ERP latencies still is (e.g., Schmitt et al., 2001).

For this argument to work, we have to assume, for instance (comparing with the study of Schmitt et al., 2001), that visual and conceptual activation takes a minimum of 300 ms longer than in the typical go/no-go categorization experiments (e.g., Thorpe et al., 1996). However, the rationale that the dual-task situation and the more difficult conceptual classification will cause the brain to recognize an object 300 ms slower does not seem tenable. This is because ERP studies with more complex designs (including dual-task situations), more detailed visual input, and more subtle semantic encoding demands have shown that conceptual effects occur 100–150 ms later and not twice the amount (e.g., Holcomb and McPherson, 1994; Johnson and Olshausen, 2003; Schendan and Kutas, 2003; Eddy et al., 2006; Sitnikova et al., 2006). Another possibility is that the delay stems from other, speech unrelated operations triggered by the complex dual-task, which would constantly add time on any given trial (Schmitt et al., 2001). If so, it becomes probable that this *other* process will interfere and/or interact with the processes of interest. That is, as argued above, it is unlikely that the constant surplus of the speech unrelated operation induced by the task will take effect before the processes of interest. Similarly, attributing all of the additional cost to processes after the targeted linguistic operations is unlikely; otherwise, the LRP/N200 onsets should not be delayed to begin with. Consequently, at least part of the excessive cost has to occur during the

speech production stages under investigation. If so, the LRP/N200 peak latencies cannot be considered reliable time frames between components of word production.

A different possibility to explain the late onsets of the N200/LRP data is that they reflect the duration constant between linguistic operations which pop-up during response selection. But also such argumentation is hard to maintain: deciding upon a response is computed by the noisy accumulation of input information (e.g., Ratcliff, 1978; Luce, 1986; Usher and McClelland, 2001; Bogacz, 2007). The noise, which occurs at several levels of processing, means that sometimes response decisions are made prior to the completion of input recognition and sometimes well after, depending on a given context, task, instructions, input, etc. (e.g., the famous speed–accuracy trade off). Thus, if the N200/LRP peak latencies reflect real-time temporal differences between two linguistic operations which become manifest during response selection, this means that the amount of noise at the various levels of processing always has to be the same in these paradigms, regardless whether we are deciding upon a semantic, syntactic, or phonemic dimension. A first indication that such premise is probably wrong is the observation that the speed of a binary decision is influenced by the repertoire of all possible outcomes, even when the alternative possibilities stored in the repertoire are irrelevant for the task (e.g., Costa et al., 1998). The more possible outcomes in the response repertoire, the slower the binary decision will be. Thus, a binary decision for gender, where the whole possible response repertoire compromises 2 or 3 outcomes, will have less noise compared to a binary decision based on phonemes, where the whole scope of possible outcomes is easily 10 times higher. Consequently, response selection should be later for phonology than for syntax, but the difference in latency is not necessarily representative for the amount of time that elapses between the lower-level linguistic calculations, given that the meta-linguistic decision is slower (more noisy) in one case than in the other. Similarly, the semantic categorization of objects (the input in these tasks) as animals, tools, plants, and so forth is the natural way our brain organizes conceptual knowledge (e.g., Warrington and Shallice, 1984; Warrington and McCarthy, 1987; Caramazza and Shelton, 1998; Caramazza and Mahon, 2003; Martin, 2007). Categorizing objects as starting with phoneme *x*, *y*, or *z*, is not. The frequency imbalance between the input–response associations predicts faster decision latencies for deciding upon the semantic category of an object compared to deciding upon the first phoneme of an object's name. In other words, assuming that latency differences between different lower-level identification processes (semantics, syntax, or phonemes) fully correspond to the latency differences at the response decision level is not straightforward.

When taking these concerns into account, yet another picture of time course can be constructed to describe the data (**Figure 2C**). While we do not deny that at the time the N200/LRP designs were elegant ways of trying to get insights into the time course of speech production, it is difficult to be sure whether these latencies are representative for language production. Put differently, **Figures 2B,C** are equally plausible outcomes as **Figure 2A**. Note furthermore that the objections raised are not restricted to the particular studies reported, but are problematic for any study wanting to assess the time course of linguistic operations through

a dual-task button-press methodology. This important since several recent studies still rely on these paradigms to make claims about the time course of language production (e.g., Guo et al., 2005; Zhang and Damian, 2009a,b; Camen et al., 2010; Hanulova et al., 2011). However, unless the above described methodological and theoretical concerns are empirically countered, the reliability of the time course of the components involved in word production as uncovered with these paradigms is questionable.

MEG studies of overt picture naming

In contrast to EEG, MEG was applied to overt picture naming early on. Consequently, many of the methodological concerns raised in the previous section do not apply here. In this way, the MEG studies of picture naming can be considered as the first direct probes into the time course of speech production. However, the number of studies is low (Salmelin et al., 1994; Levelt et al., 1998; Maess et al., 2002) and they have their own problems in conveying time to the components of word production.

In the first MEG study of picture naming (Salmelin et al., 1994) the most important observations were that: (a) brain activity progressed bilaterally from the occipital cortex, related to object recognition, toward the temporal and frontal lobes, related to language-specific recognition; (b) marked activity was observed in the posterior middle temporal gyrus between 200 and 400 ms, which was hypothesized to be related to word form retrieval; (c) a later broad time frame (400–600 ms) showed strongest activation around Broca's area, supposedly involved in post-lexical phonological encoding. Nevertheless, the problem with respect to time course here is the lack of any manipulation targeting a specific component of language with which to integrate the temporal information. As long as such experimental factor is not present, and without a perfect knowledge regarding the functional anatomy of the brain, it is difficult to draw concrete conclusions with respect to the time course of a particular stage of speech production. To give an example, the activation time course in Broca's area is not only associated with speech segmentation into syllables (e.g., Indefrey and Levelt, 2004; Indefrey, 2006), but also seems to play a role in lexical, syntactical, and domain–general selection processes (e.g., Friederichi, 2002; Hagoort, 2005; Sahin et al., 2009). Hence, without explicitly targeting one of these processes it becomes difficult to understand what we are looking at. The same concern can be made for the MEG study of picture naming conducted by Levelt et al. (1998). Although they did in fact manipulate lexical frequency, thus targeting (a) specific production stage(s), they were unlucky in that the frequency effect, which was present in a behavioral pilot outside the scanner, was not present for the MEG data. Therefore, they could not explore the influence of this linguistic factor and had to perform a similar analysis to the one conducted by Salmelin et al., 1994; see also Soros et al., 2003; Liljestrom et al., 2009), hereby suffering from the same problems for assigning time course to particular speech production stages.

More successful was the lexical manipulation by Maess et al. (2002) in their MEG study. In a picture naming task the authors contrasted lists which consisted of homogeneous semantic categories (e.g., dog, cat, horse, pig, etc.) with heterogeneous lists (e.g., dog, chair, tomato, dress, etc.). Previous studies have shown that homogeneous lists result in slower naming latencies compared to the heterogeneous lists and these SI effects are supposed to arise at the lexical level (e.g., Kroll and Stewart, 1994; Damian et al., 2001). Maess et al. (2002) used this blocked SI effect to track both the temporal and spatial source of semantically driven lexical access. Between and within category lists produced the most pronounced deflections in brain responses between 150 and 225 ms after picture onset for left temporal regions. This data point is important since it has advantages over the other data we reported so far. First, the task is easy and relatively natural to study word production, making the interpretational power stronger compared to the complex designs of the PWI and the N200/LRP paradigms. Second, time course is derived from the data and not placed upon the data. That is, they explored a linguistic phenomenon thought to arise at a particular stage of processing (e.g., Kroll and Stewart, 1994; Damian et al., 2001) during a simple and well-established overt picture naming task in combination with a temporal sensitive measure. Hence, this data set was the first one to offer direct evidence on the time course of lexical selection and, although replication with other manipulations are necessary, especially given the current debate surrounding SI effects (see above), 150 ms is the most reliable estimate concerning the onset of lexical access we have seen so far.

ERPs recorded during delayed naming

To conclude this section we will briefly pay some attention to ERP measures of delayed naming. Although these data do not form part of Indefrey and Levelt (2004) meta-analysis (in part because most of them were collected afterward), this paradigm suffers to some extent from similar problems for conveying time course as the above methodologies. To move away from the contested button-press paradigms, some studies introduced delayed naming ERP paradigms to more closely resemble actual speech but avoid potential motor contamination of the electrophysiological signal (e.g., Jackson et al., 2001; Jescheniak et al., 2002; Laganaro et al., 2009, 2010). In this manner, Laganaro et al., 2009; see also Laganaro et al., 2010; Laganaro and Perret, 2011) demonstrated that patients with lexico-semantic aphasia elicited ERP deflections compared to a healthy control group in the range of 100–250 ms, whereas patients with lexico-phonological aphasia produced ERP deviations around 300–450 ms. In sum, the data show a highly interesting relationship between type of aphasia and time course, and appear to correspond well with Indefrey and Levelt's (2004) temporal map. Nevertheless, concerning concrete time course there are some problems. First, it is still an open issue whether the data from brain-damaged speakers as point of comparison to infer the time course of linguistic processing (especially the duration) in healthy individuals is transparent. Second, it is difficult to relate the time course to a specific process based on the rather broad anomic classifications. For instance, the time window between 100 and 250 ms could relate to concepts, lexical representations, or both. A final concern, which is relevant for all delayed naming ERP studies, is a potential temporal confound caused by response inhibition. While some studies comparing overt versus delayed (and covert) naming demonstrated a degree of correspondence in that motor-related activity does not come into play until rather late in the course of processing (e.g., Eulitz et al., 2000; Laganaro and Perret, 2011), earlier effects (related to cognitive relevant brain

activity) may display some variations between both conditions (e.g., Eulitz et al., 2000; Laganaro and Perret, 2011). To give one example, Laganaro and Perret (2011) reported age of acquisition (AoA) effects starting around 220 ms and around 330 ms during immediate naming, but found no AoA effects in the ERPs during delayed naming. Findings like these limit the reliability of the time course uncovered in delayed naming paradigms.

ELECTRICAL BRAIN RESPONSES DURING OVERT NAMING

The above review served the purpose to stress the need for obtaining more explicit evidence concerning the time course of lexical selection in order to assess whether the temporal estimates proposed by I&L can be maintained. In this section, we will focus on some recent studies that have tried to obtain such more explicit insights by applying a fine-grained technique as ERPs during overt speech. Concerning potential methodological problems of combining overt speech with the ERP technique, various studies have demonstrated that cognitive relevant brain activity can be obtained under these conditions. Besides the MEG studies of overt naming highlighting the likely success-ratio for doing the same with EEG, in the last 5 years several studies which combined EEG recordings with overt naming have successfully replicated the presence of well-known ERP components (e.g., N2 and N400; e.g., Christoffels et al., 2007; Ganushchak and Schiller, 2008a,b; Habets et al., 2008; Koester and Schiller, 2008; Chauncey et al., 2009; Verhoef et al., 2009, 2010; Dell'acqua et al., 2010; Strijkers et al., 2010; Aristei et al., 2011). These ERP studies can roughly be placed in two classes: those who interpret a certain effect relying on the previous estimates of I&L (e.g., Christoffels et al., 2007; Habets et al., 2008; Cheng et al., 2010; Dell'acqua et al., 2010; Aristei et al., 2011; Laganaro and Perret, 2011), and those who assessed the reliability of those estimates (e.g., Koester and Schiller, 2008; Chauncey et al., 2009; Costa et al., 2009; Sahin et al., 2009; Strijkers et al., 2010, 2011).

For our purposes, only the latter ERP studies are relevant. Nevertheless, concerning the former class of ERP studies, the conclusions related to time course should be treated with caution. Let us give one concrete example: Aristei et al. (2011) traced in the ERPs when the effects of SI (and facilitation) occurred during overt object naming, with the objective to uncover whether SI effects reflect lexical competition. They observed that SI elicited ERP modulations starting around 250 ms after picture onset. Based on I&L's meta-analysis the authors concluded that SI falls within the temporal frame assigned to lexical selection, hence supporting the notion of lexical competition (see also e.g., Abdel Rahman and Melinger, 2011; Roelofs et al., 2011). However, given that we do not know the exact functionality underlying the negative deflection around 250 ms, alternative accounts for the ERP data are conceivable. For instance, if we assume that conceptual processing is still ongoing the moment the brain starts accessing lexical information, an assumption shared by most speech production models (e.g., Dell, 1986; Caramazza, 1997; Levelt et al., 1999), the negative deflection around 250 ms may relate to recognition (visual and semantic) processes instead of lexical processes (see also Janssen et al., 2011). The point we wish to make, regardless which interpretation is more parsimonious in the above example, is that unless we obtain more explicit insights on the time

course of word production components and a better understanding of the ERP characterization underlying it, temporal studies relying on indirect estimates are restricted in the scope of their interpretational power.

Exploring the temporal estimates explicitly

In this final section of part 1 we will report some recent work which, according to us, is in a better position to provide explicit temporal data relevant for word production. The rationale used in these studies is identical to and stems from the strategy adopted in the previously discussed MEG study by Maess et al. (2002). That is, a combination of a fine-grained temporal measure such as ERPs (or MEG, intracranial recordings) during immediate naming, while manipulating variables thought to affect a particular word production stage. Such approach has certain advantages: for one, the measure of interest stems from a naming response making the data more transparent to speech production, alleviating previous methodological concerns expressed against the indirect chronometric evidence. Second, the inference for assigning time course to word production components is based on psycholinguistic phenomena which are not necessarily bound to a particular theoretical framework, alleviating the theoretical concern we discussed above. Notwithstanding the advantages, this approach does have its own problems to explore the time course: probably the biggest set-back is that one has to assume *a priori* that the locus where a variable is thought to exert its effect is uncontested. However, for the majority of psycholinguistic phenomena no such uniform source exists. Luckily, there are ways to reduce the impact of this concern: first and foremost (at least in the beginning), one can manipulate different psycholinguistic phenomena within the same experiment, which allows for additional control to assess the origin of the tested effects. Second, and as consequence of the latter, if one knows which specific expression in time (e.g., ERP morphology) is sensitive to which linguistic operation, one can use such marker as a tool in future endeavors to assess the nature underlying a certain manipulation. By doing so, one can explore whether the accumulation of data gathered from different perspectives converges to the same time frame.

In that manner, Strijkers et al. (2010) traced the ERP onset of two such variables, the lexical frequency and the cognate effects, during simple picture naming[3]. While lexical frequency is known

---

[3] Indefrey (2011; see also Hanulova et al., 2011) argued that the study by Strijkers et al. (2010) does not offer independent evidence concerning the time course of lexical selection, based on the argument that psycholinguistic phenomena can have different potential loci. We disagree with this assessment. As extensively discussed by Strijkers et al. (2010; see also the current section), such argumentation would have been correct if only one ambiguous variable was manipulated. In contrast, Strijkers et al. (2010) *a priori* manipulated three variables (all for which both empirical data and theory exists which argues that they affect lexical selection) to see whether they would converge to the same temporal effect, in which case only one parsimonious source remains, hereby providing independent evidence concerning the onset of lexical access (see Strijkers et al., 2010; the current section). If not, Indefrey's (2011) concern holds for any study manipulating psycholinguistic variables. However, this is not the case, since Indefrey (2011) does rely on studies which solely manipulate one and arguably more ambiguous variables (e.g., SI in Aristei et al., 2011; name agreement in Cheng et al., 2010; etc.) as independent sources of evidence concerning time course.

to correlate with conceptual attributes[4], posing difficulties in terms of interpretation, the manipulation of cognate status could serve as a control since it has no obvious relationship with semantic variables (cognates are words with formal overlap in two languages)[5]. Strijkers and colleagues observed that ERPs elicited by high frequency items started to diverge from low frequency items around 180 ms after picture onset with the latter displaying more positive going amplitudes compared to the former (P2; see **Figure 3**). Importantly, identical results were found when comparing non-cognate versus cognate ERPs (see **Figure 3**). Interestingly, Christoffels et al. (2007) actually found the same result for cognate status in their study (personal communication). Given the overlap between the ERP signatures of the frequency and cognate effects, Strijkers et al. (2010) concluded that the early modulations could not sprout from conceptual differences but instead had to be located during the onset of lexical access[6]. A similar result was obtained by Sahin et al. (2009) relying on a different paradigm and technique. Local field potentials (LFP) from depth electrodes placed in Broca's area of three pre-operative epileptic patients were recorded while they engaged in a sentence completion task. Around 200 ms after target presentation low frequency words elicited more positive going LFPs compared to high frequency words. Just as Strijkers et al. (2010), they concluded that lexical access in speech production initiates within 200 ms after target presentation.

In sum, the results from these different studies with distinct experimental contexts converge to the same time course: lexical selection, at least as identified through picture or word naming, initiates within 200 ms after stimulus presentation. Nevertheless, some potential caveats must be recognized. One disadvantage is that they all contrasted between-stimuli manipulations. Comparing distinct physical items with electrical recordings can elicit distinct brain responses independent of the actual manipulation (but see Footnote 5). Another nuisance relates to the possibility that, if the brain is a highly interactive device, an imbalance between items at the lexical level (e.g., lexical frequency, cognate status) may provoke over time a similar imbalance at the

---

[4]It is also often debated which component of word retrieval is affected by lexical frequency, lemmas, or lexemes (e.g., Jescheniak and Levelt, 1994; Caramazza et al., 2001). Clearly, the variable can only be indicative of the onset of lexical access, if it is sensitive to the initial transition from concepts to lexical knowledge. Nevertheless, by now this is much less an issue since the vast majority of researchers agree that lexical frequency affects the lexicon across the board and compelling evidence exists backing up this notion. In addition, by manipulating cognate status as well, the authors could control for this potential, though unlikely, confound.

[5]This was confirmed by independent ratings of the materials used in the experiments. Cognates and non-cognates were indistinguishable in terms of familiarity, typicality, imageability, and visual complexity. In addition, the inter-stimulus perceptual variability was calculated between high and low frequency items and cognate and non-cognate items. None of the comparison resulted in differences, excluding visual factors as potential source of the effects.

[6]Furthermore, for both the frequency and cognate effects the authors demonstrated significant positive correlations between the P2 amplitude and the naming latencies, but, importantly, not between the onsets of the frequency and cognate effects in the ERPs (that is, the splitting point latencies indexing the onset of ERP divergences nor the peak latencies correlated with naming speed or effect size of the naming latencies). This indicates that the onset of both effects come about during initial lexical activation (the transition from concepts to lexical representations) and not after lexical activation at the moment of selection.
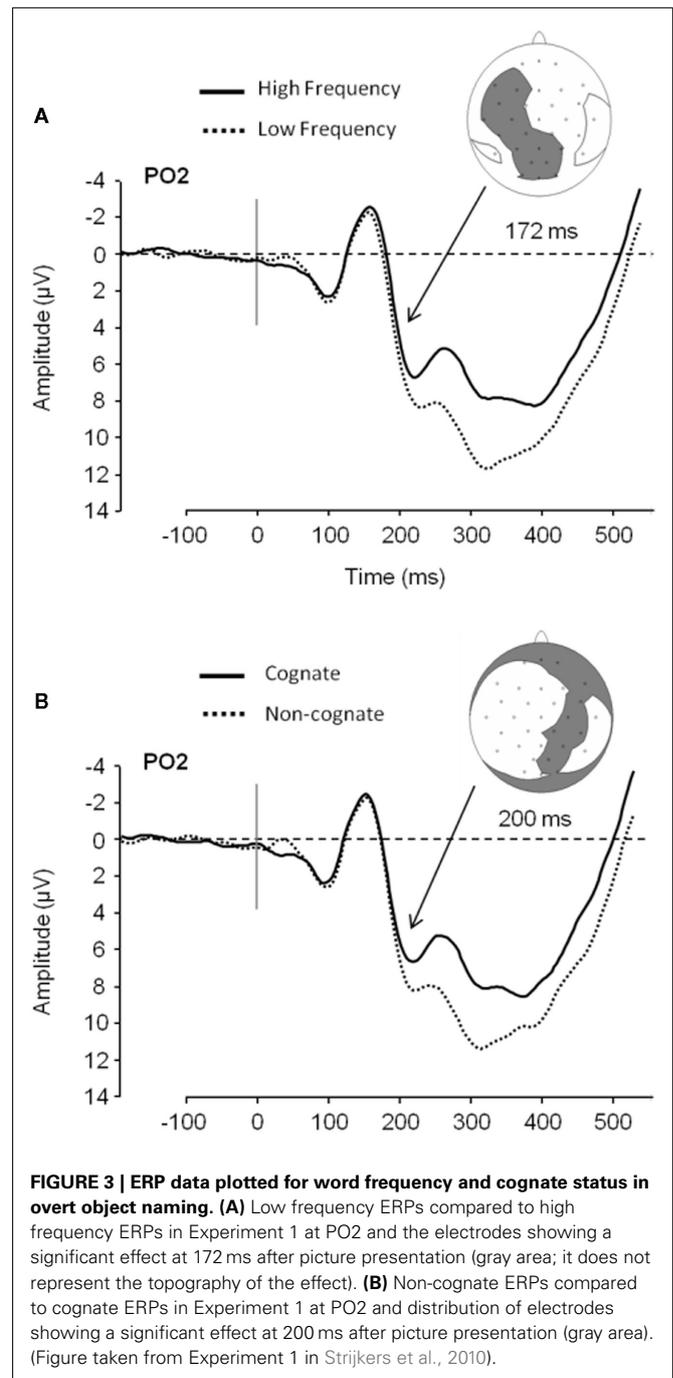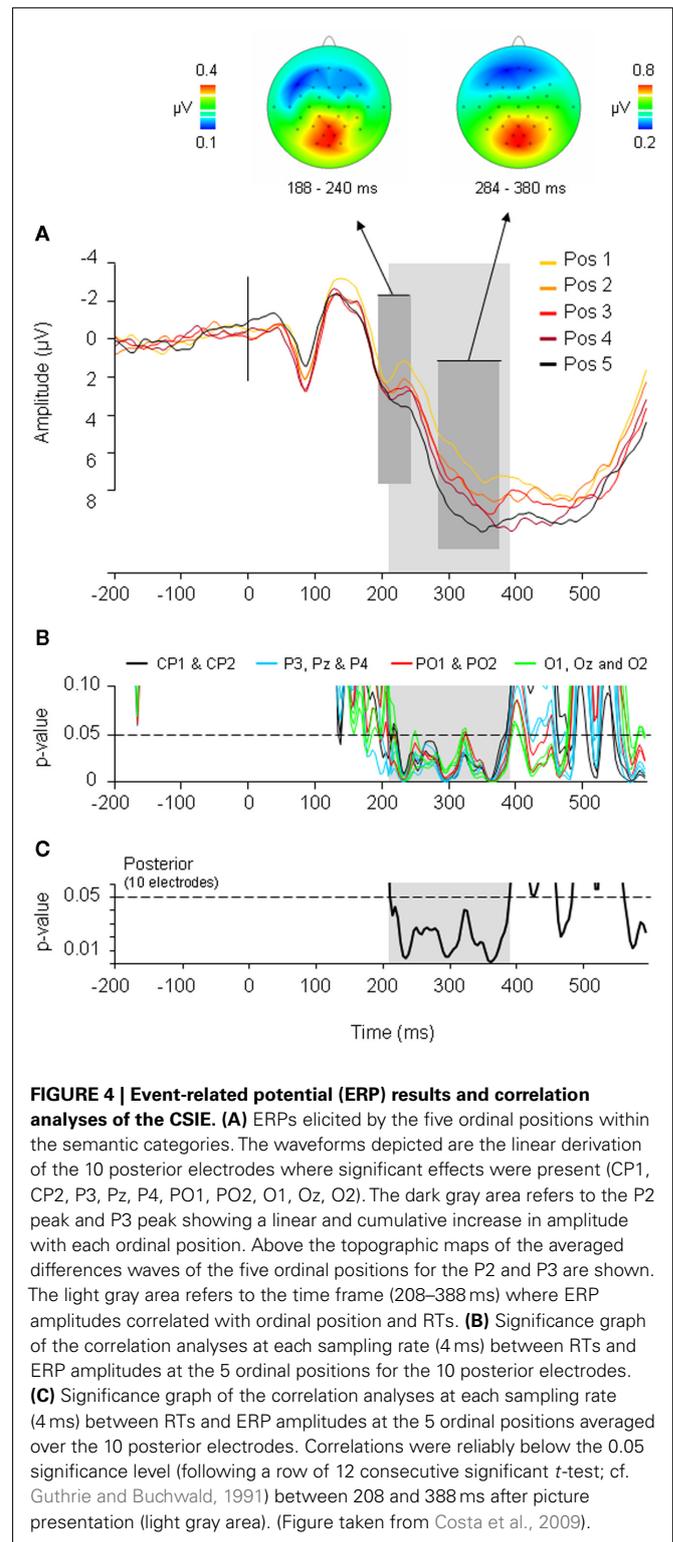

FIGURE 3 | ERP data plotted for word frequency and cognate status in overt object naming. (A) Low frequency ERPs compared to high frequency ERPs in Experiment 1 at PO2 and the electrodes showing a significant effect at 172 ms after picture presentation (gray area; it does not represent the topography of the effect). **(B)** Non-cognate ERPs compared to cognate ERPs in Experiment 1 at PO2 and distribution of electrodes showing a significant effect at 200 ms after picture presentation (gray area). (Figure taken from Experiment 1 in Strijkers et al., 2010).

conceptual level due to the continuous cross-talk between the two representational systems (see Strijkers et al., 2010). To counter these potential concerns, Strijkers et al. (2010) compared the ERPs between their two experiments. The only difference between them was language dominance, namely bilinguals producing speech in L1 for Experiment 1 and in L2 for Experiment 2. The between-experiment comparison resulted in the same P2 modulation, with L2 naming eliciting more positive amplitudes compared to L1 naming around 192 ms after picture onset. Given that in this case the ERPs which are contrasted stem from the same items,

removing the between-stimuli concern, and that conceptual activation should be similar between L1 and L2 naming of concrete nouns (at least for early high proficient bilinguals using both their languages daily), removing the interactivity concern, the most parsimonious explanation which remains is that the effects tested in the study of Strijkers et al. (2010) originated during the initiation of lexical access. In conclusion, these data provide an explicit confirmation of the onset of lexical access as estimated by Indefrey and Levelt (2004). In fact, despite the concerns discussed in the previous section, it is remarkable that the data stemming from a very different approach corresponds so well with the work done by Indefrey and Levelt (2004).

Following along these lines, Costa et al. (2009) wanted to generalize the ERP findings of Strijkers et al. (2010) to another type of manipulation. Furthermore, they also see whether with the same strategy the duration of lexical selection could be plotted. With this aim in mind, the ERP signature of the cumulative semantic interference effect (CSIE) was tracked. The CSIE refers to the observation that people are increasingly slower in naming objects which belong to the same semantic category as previously named objects (e.g., Brown, 1981; Howard et al., 2006). This effect has some interesting properties: first, the crucial manipulation in the CSIE paradigm is ordinal position (i.e., the position in which an item of a certain category appears) rather than within-item attributes (such as lexical frequency). This means that the different conditions contain the same physical stimuli, making the paradigm very suitable to combine with ERPs. Second, the effect has a rather uncontested lexical locus. Although it is currently debated whether or not the CSIE is an indication of lexical competition, the two formalized accounts (i.e., lexical competition and incremental learning) both assume the effects to come about during lexical selection (e.g., Howard et al., 2006; Oppenheim et al., 2007, 2010; Navarrete et al., 2010). Finally, given the linear nature of the effect one cannot only explore when the CSIE initiates, but also how long the correlation between ordinal position and the linear increase lasts.

Costa et al. (2009) observed an ERP pattern which mimicked the behavioral data: ERPs elicited by ordinal position increased cumulatively for each subsequent position. Importantly, this cumulative increase first became apparent at the same P2 peak where Strijkers et al. (2010) reported frequency, cognate, and language effects. Each subsequent item belonging to the same category as previously named items induced a positive increase in ERP amplitudes around 200 ms (see **Figure 4**). To ascertain how long the cumulative ERP deflections manifested, the amplitudes for each ordinal position were correlated ms-by-ms with the corresponding naming latencies. Significant positive correlations were observed between 208 and 388 ms (see **Figure 4**). Two main conclusions were drawn: first, previous findings showing that the brain initiates lexical access within 200 ms after picture onset (e.g., Maess et al., 2002; Sahin et al., 2009; Strijkers et al., 2010) were replicated with a different experimental setting, hereby again corroborating I&L's temporal estimate concerning the onset of lexical access. Second, the cumulative effect lasted for 180 ms. If this latency reflects the time required for an intended lexical representation to be singled out, the time course is notably longer compared to previous indirect (e.g., Levelt et al., 1991, 1998; Schmitt et al.,

FIGURE 4 | Event-related potential (ERP) results and correlation analyses of the CSIE. (A) ERPs elicited by the five ordinal positions within the semantic categories. The waveforms depicted are the linear derivation of the 10 posterior electrodes where significant effects were present (CP1, CP2, P3, Pz, P4, PO1, PO2, O1, Oz, O2). The dark gray area refers to the P2 peak and P3 peak showing a linear and cumulative increase in amplitude with each ordinal position. Above the topographic maps of the averaged differences waves of the five ordinal positions are shown. The light gray area refers to the time frame (208–388 ms) where ERP amplitudes correlated with ordinal position and RTs. **(B)** Significance graph of the correlation analyses at each sampling rate (4 ms) between RTs and ERP amplitudes at the 5 ordinal positions for the 10 posterior electrodes. **(C)** Significance graph of the correlation analyses at each sampling rate (4 ms) between RTs and ERP amplitudes at the 5 ordinal positions averaged over the 10 posterior electrodes. Correlations were reliably below the 0.05 significance level (following a row of 12 consecutive significant *t*-test; cf. Guthrie and Buchwald, 1991) between 208 and 388 ms after picture presentation (light gray area). (Figure taken from Costa et al., 2009).

2001; Indefrey and Levelt, 2004) and direct estimates (Maess et al., 2002). However, prior to assuming that this finding would be problematic for the I&L model, let us consider at least two reasons which can be put forward to account for the longer latency:

(A) if the system propagates information in a cascaded manner, the cumulative interference effect which initiates during lexical selection may spill over to phonology; (B) if the speed of lexical selection is subject to the amount of items competing at any given moment (or alternatively, amount of connection adjustments in a non-competitive account; see Oppenheim et al., 2010), longer latencies are expected in the current design compared to previous studies where only two "competing" words were contrasted (see Costa et al., 2009). Alternatively (option C), lexical selection lasts roughly 180 ms regardless the amount of competing items or the extent of cascading.

Based on the results of Costa and colleagues we cannot directly differentiate between the three alternative accounts for the 180 ms time window. Nonetheless, a comparison with the time frame identified in the MEG study of Maess et al. (2002) allows us to doubt the explanatory power of option B. Maess et al. (2002) report a latency of 75 ms (150–225 ms) where brain responses were significantly different between same and mixed semantic category conditions. Just as in Costa et al. (2009), the conditions which were contrasted consisted of five related or unrelated objects. If we assume that the semantic blocking effect is another application of the same neuronal expression as the CSIE, the longer latency reported for the CSIE cannot be explained in terms of the amount of semantically related items which are compared (option B). Regarding options A and C, it might be interesting to look again at the results of Sahin et al. (2009). Besides lexical frequency, in that study also the time course of grammatical processing (null-inflection; e.g., present tense) and phonological encoding (overt-inflection; e.g., past tense) were traced in Broca's area. The grammatical operation became apparent around 320 ms and the phonological inflection produced modulations around 450 ms after stimulus presentation. Thus, from lexical onset (∼200 ms) to grammar 120 ms passed and to phonology 250 ms elapsed. Although a direct comparison between both studies is not straightforward given the differences in tasks, measures, participants, and especially manipulated variables, using Sahin et al.'s (2009) uncovered latencies as "independent" entities of word production stages may be helpful to generate hypotheses of what the 180 ms in Costa et al. (2009) could refer to. That is, if one takes the data of Sahin et al. (2009) to reflect the first pass of information (e.g., Hagoort and Levelt, 2009), the 180 ms latency as uncovered with the CSIE would best correspond with the latencies reported by Sahin et al. (2009) if characterized as the time window to complete lexical selection (option C). If the duration estimates as identified in Broca's area reflect selection processes (e.g., Thompson-Schill et al., 1997; Schnur et al., 2009) rather than the onset of those linguistic operations, then the 180 ms latency better corresponds with the notion that the CSIE affects both lexical selection and phonological encoding (option A; see also Goldrick et al., 2009). It will be interesting to see how these and related paradigms can be exploited in the future to gain more stable knowledge on what these latency differences between the various studies reveal about the dynamics underlying speech production.

Conclusion

In this section we reviewed studies on the time course of lexical selection during overt naming. Based on these data, we believe that there is sufficient evidence gathered to respond to our initial question posed in this review, namely "exactly how fast does the brain engages in lexical access?": the brain engages in lexical access within 200 ms after stimulus presentation. One of the important reasons why we can be confident about the onset latency of lexical access is the accumulation of evidence over variables and paradigms (see **Table 1**). Of course, it does not mean that every modulation around 200 ms in a production task necessarily indexes the initiation of lexical access. But, there are ways to narrow down the probability that one is dealing with a lexical source. To name one, the above studies also revealed an ERP signature which seems to be sensitive to the difficulty of lexical selection, namely the P2 with more positive brain responses for the more difficult lexical conditions (e.g., Christoffels et al., 2007; Costa et al., 2009; Strijkers et al., 2010, 2011).

Finally, returning to the work which stimulated the time course research in the field and lays at the basis of the current review, namely Indefrey and Levelt's (2004) temporal model, two main conclusions can be drawn: first and foremost, all studies which explored the onset of lexical selection in more natural experimental settings provided evidence which corroborates I&L's estimate. Second, concerning the latency of lexical selection, less definite statements can be made. The few studies which were able to look at the duration of lexical selection (Maess et al., 2002; Costa et al., 2009; Sahin et al., 2009) displayed varying estimates roughly ranging between 75 and 180 ms. Based on these data, if we were to apply a similar averaging approach as Indefrey and Levelt (2004), the estimate of 75 ms should be raised to ∼130 ms. However, before doing so, we should understand why the different durations between studies emerged as well as collect more explicit data on the time frame of lexical selection. In this aspect, we reiterate that, as long as we do not have more reliable evidence, caution has to be exerted when interpreting temporal data solely in function of I&L's model (and, although the current review focuses on lexical selection, this also holds for the subsequent production stages).

BY WHICH MECHANISM IS THE LEXICON ACCESSED SO RAPIDLY?

Having established that within 200 ms after stimulus presentation information is transmitted from the conceptual level to the lexicon, in this final and more concise section we will address by what means such rapid onset of lexical access is achieved. First

Table 1 | Overview of lexical variables and their time course reported in ERP studies of immediate naming.

Lexical phenomenon	Onset (ms)	Study
Lexical frequency	∼200	Sahin et al. (2009)
	∼178	Strijkers et al. (2010)
Cognate status	∼192	Strijkers et al. (2010)
	∼190	Christoffels et al. (2007)
Semantic interference	∼150	Maess et al. (2002)
	∼208	Costa et al. (2009)
L1 versus L2 Naming	∼195	Strijkers et al. (2010)
	∼190	Christoffels et al. (2007)

we will describe the theoretical considerations made by speech production models concerning this issue and then present a recent study which, by exploiting the temporal information reviewed above, offers an explanation which currently no speech production model integrates. In this manner, besides addressing the functional mechanism behind the time course on lexical selection, this section highlights how descriptive temporal information on speech production can be exploited to address cognitive questions from a novel perspective.

LEXICAL ACCESS: A MATTER OF SPREADING ACTIVATION OR CONCEPT-DRIVEN SELECTION?

A widely endorsed notion on how lexical access in speech production engages, is through semantically driven spreading activation (e.g., Dell, 1986; Dell and O'Seaghdha, 1991, 1992; Roelofs, 1992, 1997; Caramazza, 1997; Dell et al., 1997; Levelt et al., 1999). Spreading activation refers to the automatic transmission of information between strongly connected representations (e.g., Collins and Loftus, 1975). For speech production this implies that the moment a concept becomes active it automatically triggers the corresponding lexical representation(s) regardless of whether a speaker intends to utter the word in question or not. Thus, according to the models embracing the principle of spreading activation, the lexicon receives rapid input from the *activated* semantic system due to the strong links between them. Note that these models are not free from intention, but any modulations to the activation levels of lexical representations in function of a speaker's *linguistic intentions*[7] take effect reactively; that is, after the lexicon has received some input from the active semantic representations in a feedforward manner (e.g., Dell, 1986; Roelofs, 1992, 1997, 2003; Caramazza, 1997; Levelt et al., 1999). Most of the evidence supporting this property comes from contextual effects in picture naming where distractors, items a speaker does not want to verbalize, affect the speed of target naming (e.g., Glaser and Glaser, 1989; Levelt et al., 1991; Roelofs, 1992, 2003, 2006, 2008; Peterson and Savoy, 1998; Cutting and Ferreira, 1999; Costa and Caramazza, 2002; Morsella and Miozzo, 2002; Navarrete and Costa, 2005; Meyer and Damian, 2007). Indeed, these and related effects have been interpreted to reveal parallel activation of both the speech intended and non-intended lexical information, confirming the role of spreading activation. However, not all researchers concur with the interpretation given to these contextual effects and refute the notion of automatic spreading activation (e.g., Levelt, 1989; Bloem and La Heij, 2003; Bloem et al., 2004). Instead, they argue that the intention to verbalize an item has to be specified at the conceptual level in order to access the linguistic system. According to concept selection models, fast lexical access is not achieved through automaticity but rather through specificity, namely the fine-grained propagation of activity between those activated concepts one intends to verbalize and their corresponding lexical representations. In contrast to spreading activation models of lexical access, according to the latter only speech intended information will get processed linguistically. In sum, although both classes of

models assume that initial lexical access occurs in a feedforward manner from concepts to words, the way this feedforward propagation takes place is distinct; being automatic in one case and concept-specific in the other.
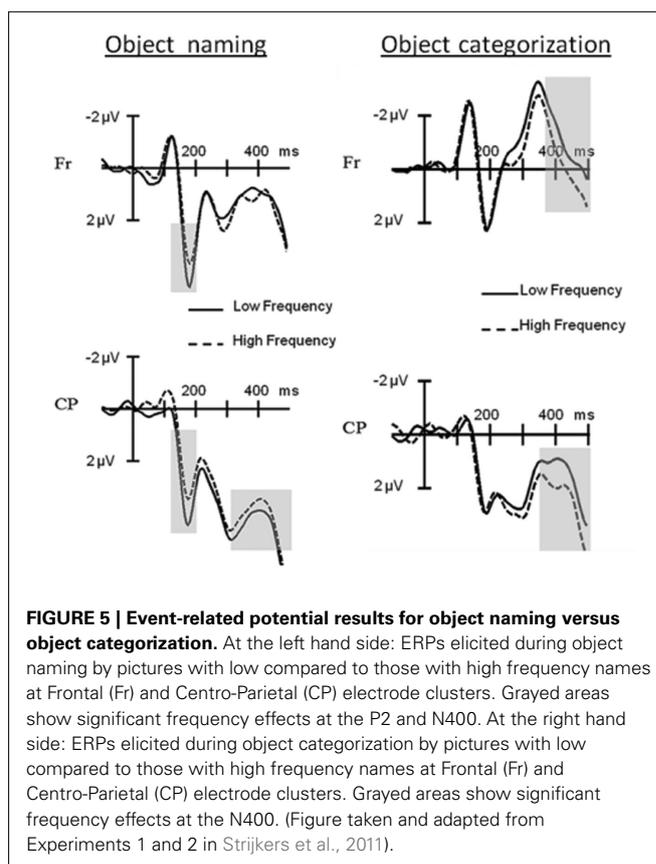
RAPID LEXICAL SELECTION THROUGH TOP-DOWN PROACTIVE FACILITATION

There are many studies in the literature exploring the linguistic effects of stimuli we do not intend to utter during picture naming (i.e., PWI and picture–picture tasks) and a few studies exploring the same for non-verbal paradigms (e.g., Glaser and Glaser, 1989; Levelt et al., 1991; Roelofs, 1992, 1997, 2003, 2006, 2008; Meyer et al., 1998, 2007; Cutting and Ferreira, 1999; Costa and Caramazza, 2002; Jescheniak et al., 2002; Morsella and Miozzo, 2002; Navarrete and Costa, 2005; Bles and Jansma, 2008) Nevertheless, regardless of whether these studies reported linguistic effects for speech unintended information, most of them remain silent of *when* the intention to speak takes effect (see Strijkers et al., 2011). Recently, Strijkers et al. (2011) explored the issue. Besides wanting to contrast the two theoretical constructs specified above, they entertained a third possibility: in vision science it has been demonstrated that top-down signals related to attention, intention, and context can trigger task-relevant representations prior to the feedforward sensory-driven activity (e.g., Desimone and Duncan, 1995; Luck et al., 1997; Kastner et al., 1999; O'Craven and Kanwisher, 2000; Bar, 2003; Bar et al., 2006; Gilbert and Sigman, 2007; Peelen et al., 2009). Relying on these advances, Strijkers et al. (2011) hypothesized that the intention to speak may pre-activate the lexical system in a top-down fashion prior to the feedforward flow coming from the conceptual system. In order to test the three possibilities they exploited the P2 modulation (descriptively labeled pP2), previously shown to be sensitive to lexical variables in picture naming tasks (e.g., Costa et al., 2009; Strijkers et al., 2010). Specifically, the word frequency effect was compared in an overt picture naming task versus a non-verbal semantic categorization task. Given that previous studies have shown that during semantic categorization the basic-level concept associated with the input becomes activated (e.g., Grill-Spector and Kanwisher, 2005; Eddy et al., 2006), contrasting the ERP response to lexical frequency between both tasks allowed Strijkers and colleagues to explore the temporal role of the intention to speak. This is because in one case participants had the conscious intention to verbalize the activated concepts (picture naming) while in the other case they did not (picture categorization). This set-up resulted in the following predictions: (a) if initial lexical access is achieved through spreading activation regardless of a speaker's intention, then for both the verbal and non-verbal task an early pP2 frequency effect should be observed; (b) if lexical access takes place through concept selection, then only a pP2 frequency effect should be found for the naming task and no word frequency modulation should be elicited in the non-verbal task; and finally; (c) if initial access to the lexicon occurs through top-down pre-activation, then both tasks should display a word frequency effect, but a pP2 deflection should only be present for the task where participants have the intention to speak (picture naming), while in the non-verbal task the frequency effect should be qualitatively different (due to the absence of intention).

---

[7]As opposed to semantic intentions or the conceptualization of the message a speaker wishes to communicate (which naturally has to be specified prior to accessing the linguistic system).

The results supported the last option: in the picture naming task the ERPs elicited by low frequency items started to diverge from the ones elicited by high frequency items 156 ms after picture onset, with more positive amplitudes for the former especially at the posterior electrodes. A finding which replicated the pP2 frequency effect (see **Figure 5**; Strijkers et al., 2010). In contrast, in the categorization task, no ERP differences for lexical frequency were observable at the pP2. Instead, the effect came about 200 ms later resulting in a typical N400 modulation (see **Figure 5**). To the extent that lexical frequency is sensitive to the first pass activation of lexical items, as indexed by most recent behavioral, neurophysiological, speech-error, and patient data (e.g., Caramazza et al., 2001; Navarrete et al., 2006; Almeida et al., 2007; Graves et al., 2007; Kittredge et al., 2007; Knobel et al., 2008; Strijkers et al., 2010), the *qualitatively* different ERP onset of this variable indicates that the brain's rapid engagement in lexical access is driven by the conscious intention to perform a speech act (the pP2) rather than only through automatic feedforward spreading activation from concepts to words (the N400). In addition, this intention-driven access is generated by a top-down signal rather than a feedforward signal from selected semantic representations as specified in concept selection models (e.g., Bloem and La Heij, 2003; Bloem et al., 2004). This is because a word frequency effect was still apparent in the non-verbal semantic task. Thus, two main conclusions could be drawn from these findings: first, activated concepts trigger lexical knowledge regardless of whether one has the intention to utter the name of that concept. In other words, these data offered support for those models of lexical access which embrace the principle of spreading activation (e.g., Dell, 1986; Dell and O'Seaghdha, 1992; Roelofs, 1992, 1997; Caramazza, 1997; Dell et al., 1997; Levelt et al., 1999). Second, in contrast to the specifications of current speech production models, if there is intention to engage in a speech act, initial access to the lexicon is facilitated in a proactive manner. Put differently, spreading activation models of lexical access need to be complemented by a top-down mechanism capable of proactively tuning the lexical system in function of a speaker's intention. It is important to point out that spreading activation models of lexical access are formalized from the perspective that there is intention to speak. Hence, the data amplify them to a broader context of information processing rather than falsifying them. Similarly, the results of Strijkers et al. (2011) remain silent about the functionality of reactive goal-directed mechanisms during speech production. That is, reactive and proactive goal-directed mechanisms are not mutually exclusive: when the intention to speak is specified, it is perfectly possible that selection of target-relevant lexical information is ensured through "verification of production-rules" or "lexical activation boosters" as described in certain models (e.g., Roelofs, 1992, 1997, 2003; Levelt et al., 1999; Gordon and Dell, 2003; Dell et al., 2008; Oppenheim et al., 2010). Notwithstanding, the data by Strijkers et al. (2011) does open the possibility that issues such as goal-directed linguistic selection may not solely be achieved through reactive mechanisms, but timing differences induced by proactive top-down processing may become an important factor as well; depending on the exact functionality and scope of the proactive top-down influences identified.

Strijkers et al. (2011) proposed two (not mutually exclusive) tentative accounts of how the intention to speak can bring about

FIGURE 5 | Event-related potential results for object naming versus object categorization. At the left hand side: ERPs elicited during object naming by pictures with low compared to those with high frequency names at Frontal (Fr) and Centro-Parietal (CP) electrode clusters. Grayed areas show significant frequency effects at the P2 and N400. At the right hand side: ERPs elicited during object categorization by pictures with low compared to those with high frequency names at Frontal (Fr) and Centro-Parietal (CP) electrode clusters. Grayed areas show significant frequency effects at the N400. (Figure taken and adapted from Experiments 1 and 2 in Strijkers et al., 2011).

such proactive tuning of the lexicon (based on general notions lend from vision science). One is by assuming that when a speaker intends to produce verbal output the activation level of the whole lexico-semantic network/pathway is enhanced (for similar proposals in vision see e.g., Desimone and Duncan, 1995; Luck et al., 1997; Kastner et al., 1999; O'Craven and Kanwisher, 2000; Gilbert and Sigman, 2007; Peelen et al., 2009). As a consequence of this proactive enhancement of speech-relevant pathways, the stimulus-driven access to words will be facilitated. In this case, current spreading activation models of lexical access only require slight modifications. For instance, in the WEAVER++ model this can be achieved by allowing a (general) production rule to take effect prior to spreading activation (e.g., Levelt et al., 1999; Roelofs, 2003) or in Dell's speech production model by "jolting" information related to output goals prior to entering the semantic layer (e.g., Dell and O'Seaghdha, 1992).

A second potential mechanism to explain the data of Strijkers et al. (2011) is that the top-down projections are capable of making well-estimated guesses about which words are likely to be uttered. Such prediction-based top-down influences can be achieved in several ways, but one compelling account in the case of single picture input (without context) states that the brain can rapidly transmit the coarse visual information (low spatial frequencies) to the prefrontal cortex where expectations regarding picture semantics are build up in order to pre-activate potential task-relevant object representations (e.g., Bar, 2003; Bar et al., 2006). If we adopt such mechanism to object naming, the predictions in function

of picture semantics may proactively trigger a specific subset of lexical representations. If such prediction-based mechanism is indeed functional for speech production (or at the least, for object naming), amplifications of the dynamics of existing spreading activation models will be more severe. For instance, target-relevant lexical activation is initially achieved through top-down prediction and not solely through later feedforward spreading activation. Also, eventual selection of the target will not only depend on amount of activation and reactive control, but will also depend on the amount of overlap between the guesses instantiated by top-down prediction and the slower feedforward spreading activation coming from the conceptual system (cf. Bar, 2003). Future research concerning the role of proactive goal-directed influences in speech production – a topic which has received little attention in the field so far – will be important to amplify our understanding of the dynamics underlying speech production.

CONCLUSION

In this article we critically reviewed the literature on the time course of lexical access. Various points were raised throughout the article: we urged the necessity to corroborate I&L's temporal map through direct explorations into time course. One promising way for doing so is through the accumulation of chronometric evidence stemming from the exploration of a variety of psycholinguistic variables in simple overt naming paradigms combined with precise

measures such as ERPs. In that manner, several studies have been discussed which indicate that lexical access initiates within 200 ms after stimulus onset, explicitly confirming I&L's estimate. Support for I&L's duration of lexical selection was less clear given that the few studies on this topic reported different latencies, roughly between 75 and 180 ms. An early positive going ERP component (pP2) was discussed and argued to be sensitive to lexical variables. Taking advantage of this lexically *sensitive* electrophysiological marker, it was demonstrated how time course information can be used to address cognitive questions in the field from a novel perspective. In doing so, evidence was reported that the brain's rapid engagement in lexical selection is driven by top-down intention to speak. To sum up, although still a lot of work needs to be done before having a complete temporal map of language production, the advances made in recent years are substantial and continuing research along these lines should be able to address the open issues concerning time course in the near future.

ACKNOWLEDGMENTS

REFERENCES

Abdel Rahman, R., and Melinger, A. (2009a). Dismissing lexical competition does not make speaking any easier: a rejoinder to Mahon and Caramazza (2009). *Lang. Cogn. Process.* 24, 749–760.

Abdel Rahman, R., and Melinger, A. (2009b). Semantic context effects in language production: a swinging lexical network proposal and a review. *Lang. Cogn. Process.* 24, 713–734.

Abdel Rahman, R., and Melinger, A. (2011). The dynamic microstructure of speech production: semantic interference built on the fly. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 149–161.

Abdel Rahman, R., and Sommer, W. (2003). Does phonological encoding in speech production always follow the retrieval of semantic knowledge? Electrophysiological evidence for parallel processing. *Brain Res. Cogn. Brain Res.* 16, 372–382.

Abdel Rahman, R., Van Turennout, M., and Levelt, W. J. M. (2003). Phonological encoding is not contingent on semantic feature retrieval: an electrophysiological study on object naming. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 850–860.

Alario, F.-X., Segui, J., and Ferrand, L. (2000). Semantic and associative priming in picture nam-

ing. *Q. J. Exp. Psychol. A* 53A, 741–764.

Almeida, J., Knobel, M., Finkbeiner, M., and Caramazza, A. (2007). The locus of the frequency effect in picture naming: when recognizing is not enough. *Psychon. Bull. Rev.* 14, 177–1182.

Aristei, S., Melinger, A., and Abdel Rahman, R. (2011). Electrophysiological chronometry of semantic context effects in language production. *J. Cogn. Neurosci.* 23, 1567–1586.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *J. Cogn. Neurosci.* 15, 600–609.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmidt, A. M., Dale, A. M., Hämäläinen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., and Halgren, E. (2006). Top-down facilitation of visual recognition. *Proc. Natl. Acad. Sci. U.S.A.* 103, 449–454.

Bles, M., and Jansma, B. M. (2008). Phonological processing of ignored distractor pictures, an fMRI investigation. *BMC Neurosci.* 9, 20.

Bloem, I., and La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: implications for models of lexical access in language production. *J. Mem. Lang.* 48, 468–488.

Bloem, I., van den Boogaard, S., and La Heij, W. (2004). Semantic facili-

tation and semantic interference in language production: further evidence for the conceptual selection model of lexical access. *J. Mem. Lang.* 51, 307–323.

Bock, J. K. (1982). Towards a cognitive psychology of syntax: information processing contributions to sentence formulation. *Psychol. Rev.* 89, 1–47.

Bogacz, R. (2007). Optimal decision making theories: linking neurobiology with behaviour. *Trends Cogn. Sci. (Regul. Ed.)* 11, 118–125.

Brown, A. S. (1981). Inhibition in cued retrieval. *J. Exp. Psychol. Hum. Learn. Mem.* 7, 204–215.

Camen, C., Morand, S., and Laganaro, M. (2010). Re-evaluating the time course of gender and phonological encoding during silent monitoring tasks estimated by ERP: serial or parallel processing? *J. Psycholinguist. Res.* 39, 35–49.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cogn. Neuropsychol.* 14, 177–208.

Caramazza, A. (1998). The interpretation of semantic category-specific deficits: what do they reveal about the organization of conceptual knowledge in the brain? *Neurocase* 4, 265–272.

Caramazza, A., Costa, A., Miozzo, M., and Bi, Y. (2001). The specific-word frequency effect: implications for the

representation of homophones in speech production. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 1430–1450.

Caramazza, A., and Mahon, B. Z. (2003). The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends Cogn. Sci.* 7, 354–361.

Caramazza, A., and Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: the animate-inanimate distinction. *J. Cogn. Neurosci.* 10, 1–34.

Chauncey, K., Holcomb, P. J., and Grainger, J. (2009). Primed picture naming within and across languages: an ERP investigation. *Cogn. Affect. Behav. Neurosci.* 9, 286–303.

Cheng, X. R., Schafer, G., and Akyurek, E. G. (2010). Name agreement in picture naming: an ERP study. *Int. J. Psychophysiol.* 76, 130–141.

Christoffels, I. K., Firk, C., and Schiller, N. (2007). Bilingual language control: an event-related brain potential study. *Brain Res.* 1147, 192–208.

Coles, M. G. H. (1989). Modern mind-brain reading: psychophysiology, physiology, and cognition. *Psychophysiology* 26, 251–269.

Collins, A. M., and Loftus, E. F. (1975). A spreading-activation theory of semantic priming. *Psychol. Rev.* 82, 407–428.

Costa, A., Alario, F.-X., and Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture–word interference paradigm. *Psychon. Bull. Rev.* 12, 125–131.

Costa, A., and Caramazza, A. (2002). The production of noun phrases in English and Spanish: implications for the scope of phonological encoding in speech production. *J. Mem. Lang.* 46, 178–198.

Costa, A., Cutler, A., and Sebastián-Gallés, N. (1998). Effects of phoneme repertoire on phoneme decision. *Percept. Psychophys.* 60, 1022–1031.

Costa, A., Strijkers, K., Martin, C., and Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21442–21446.

Cutting, J. C., and Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 318–344.

Damian, M. F., Vigliocco, V., and Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition* 81, B77–B86.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* 93, 283–321.

Dell, G. S., Oppenheim, G. M., and Kittredge, A. K. (2008). Saying the right word at the right time: syntagmatic and paradigmatic interference in sentence production. *Lang. Cogn. Process.* 23, 583–608.

Dell, G. S., and O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: a comment on Levelt et al. (1991). *Psychol. Rev.* 98, 604–614.

Dell, G. S., and O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition* 42, 287–314.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., and Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychol. Rev.* 104, 801–838.

Dell'acqua, R., Sessa, P., Peressotti, F., Mulatti, C., Navarrete, E., and Grainger, J. (2010). ERP evidence for ultra-fast semantic processing in the picture-word interference paradigm. *Front. Psychol.* 1:177. doi:10.3389/fpsyg.2010.00177

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222.

Doniger, G. M., Foxe, J. J., Murray, M. M., Higgins, B. A., Snodgrass, J. G., and Schroeder, C. E. (2000). Activation time course of ventral visual stream object-recognition areas: high density electrical mapping of perceptual closure processes. *J. Cogn. Neurosci.* 12, 615–621.

Eddy, M., Schmid, A., and Holcomb, P. (2006). Masked repetition priming and event-related brain potentials: a new approach for tracking the time course of object perception. *Psychophysiology* 43, 564–568.

Eimer, M. (1993). Effects of attention and stimulus probability on ERPs in a go/nogo task. *Biol. Psychol.* 35, 123–138.

Eulitz, C., Hauk, O., and Cohen, R. (2000). Electroencephalographic activity over temporal brain areas during phonological encoding in picture naming. *Clin. Neurophysiol.* 111, 2088–2097.

Finkbeiner, M., Almeida, J., Janssen, N., and Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1075–1089.

Finkbeiner, M., and 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.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* 6, 78–84.

Fromkin, V. A. (1971). The non-anomalous of anomalous utterances. *Language* 47, 27–52.

Ganushchak, L. Y., and Schiller, N. O. (2008a). Motivation and semantic context affect brain error-monitoring activity: an event-related brain potentials study. *Neuroimage* 39, 395–405.

Ganushchak, L. Y., and Schiller, N. O. (2008b). Brain error-monitoring activity is affected by semantic relatedness: an event-related brain potentials study. *J. Cogn. Neurosci.* 20, 927–940.

Garrett, M. F. (1975). "The analysis of speech production," in *The Psychology of Learning and Motivation*, ed. G. H. Bower (London: Academic Press), 133–177.

Gilbert, C. D., and Sigman, M. (2007). Brain states: top-down influences in sensory processing. *Neuron* 54, 677–696.

Glaser, W. R., and Düngelhoff, F. J. (1984). The time course of picture-word interference. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 640–654.

Glaser, W. R., and Glaser, M. O. (1989). Context effects on Stroop-like word and picture processing. *J. Exp. Psychol. Gen.* 118, 13–42.

Goldrick, M., Dell, G. S., Kroll, J., and Rapp, B. (2009). Sequential information processing and limited interaction in language production. *Online letters to Science.* Available at: http://www.sciencemag.org/content/326/5951/445/reply

Gordon, J. K., and Dell, G. S. (2003). Learning to divide the labor: an account of deficits in light and heavy verb production. *Cogn. Sci.* 27, 1–40.

Graves, W. W., Grabowski, T. J., Mehta, S., and Gordon, J. K. (2007). A neural signature of phonological access: distinguishing the effects of word frequency from familiarity and length in overt picture naming. *J. Cogn. Neurosci.* 19, 617–631.

Grill-Spector, K., and Kanwisher, N. (2005). Visual recognition: as soon as you know it is there, you know what it is. *Psychol. Sci.* 16, 152–160.

Guo, T. M., Peng, D. L., Lu, C. M., and Liu, H. Y. (2005). The temporal course of semantic and phonological activation in Chinese word production: an ERP study. *Acta Psychol. Sin.* 37, 569–574.

Guthrie, D., and Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology* 28, 240–244.

Habets, B., Jansma, B. M., and Münte, T. F. (2008). Neurophysiological correlates of linearization in language production. *BMC Neurosci.* 9, 77. doi:10.1186/1471-2202-9-77

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends Cogn. Sci.* 9, 416–423.

Hagoort, P., and Levelt, W. J. M. (2009). The speaking brain. *Science* 326, 372–373.

Hanulova, J., Davidson, D. J., and Indefrey, P. (2011). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Lang. Cogn. Process.* 26, 902–934.

Hauk, O., Patterson, K., Woollams, A., Cooper-Pye, E., Pulvermuller, F., and Rogers, T. T. (2007). How the camel lost its hump: the impact of object typicality on event-related potential signals in object decision. *J. Cogn. Neurosci.* 19, 1338–1353.

Holcomb, P. J., and McPherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain Cogn.* 24, 259–276.

Howard, D., Nickels, L., Coltheart, M., and Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition* 100, 464–482.

Indefrey, P. (2006). "A meta-analysis of hemodynamic studies on first and second language processing: which suggested differences can we trust and what do they mean?" in *The Cognitive Neuroscience of Second Language Acquisition*, eds M. Gullberg and P. Indefrey (Malden, MA: Blackwell Publishing), 279–304.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. *Front. Psychol.* 2:255. doi: 10.3389/fpsyg.2011.00255

Indefrey, P., and Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144.

Jackson, G. M., Swainson, R., Cunnington, R., and Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Biling. Lang. Cogn.* 4, 169–178.

Janssen, N., Carreiras, M., and Barber, H. A. (2011). Electrophysiological effects of semantic context in picture and word naming. *Neuroimage* 57, 1243–1250.

Janssen, N., Schirm, W., Mahon, B. Z., and Caramazza, A. (2008). The semantic interference effect in the picture-word interference paradigm: evidence for the response selection hypothesis. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 249–256.

Jescheniak, J. D., and Levelt, W. J. M. (1994). Word frequency effects in speech production: retrieval of syntactic information and of phonological form. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 824–843.

Jescheniak, J. D., Schriefers, H., Garrett, M. F., and Friederici, A. D. (2002). Exploring the activation of semantic and phonological codes during speech planning with event-related brain potentials. *J. Cogn. Neurosci.* 94, 94–103.

Johnson, J. S., and Olshausen, B. A. (2003). Time course of neural signatures of object recognition. *J. Vis.* 3, 499–512.

Johnson, J. S., and Olshausen, B. A. (2005). The earliest EEG signatures of object recognition in a cued-target task are postsensory. *J. Vis.* 5, 299–312.

Kastner, S., Pinsk, M., De Weerd, P., Desimone, R., and Ungerleider, L. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22, 751–761.

Kirchner, H., and Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: visual processing revisited. *Vision Res.* 46, 1762–1776.

Kittredge, A. K., Dell, G. S., Verkuilen, J., and Schwartz, M. F. (2007). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cogn. Neuropsychol.* 1, 1–30.

Knobel, M., Finkbeiner, M., and Caramazza, A. (2008). The many places of frequency: evidence for a novel locus of the lexical frequency effect in word production. *Cogn. Neuropsychol.* 25, 256–286.

Koester, D., and Schiller, N. (2008). Morphological priming in overt language production: electrophysiological evidence from Dutch. *Neuroimage* 42, 1622–1630.

Kok, A. (1986). Effects of degradation of visual stimuli on components of the event-related potential (ERP) in go/nogo reaction tasks. *Biol. Psychol.* 23, 21–38.

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174.

Laganaro, M., Morand, S., Michel, C. M., Spinelli, L., and Schnider, A. (2010). ERP correlates of word production before and after stroke in an aphasic patient. *J. Cogn. Neurosci.* 23, 374–381.

Laganaro, M., Morand, S., and Schnider, A. (2009). Time course of evoked-potential changes in different forms of anomia in aphasia. *J. Cogn. Neurosci.* 21, 1499–1510.

Laganaro, M., and Perret, C. (2011). Comparing electrophysiological correlates of word production in immediate and delayed naming through the analysis of word age of acquisition effects. *Brain Topogr.* 24, 19–29.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation.* Cambridge, MA: The MIT Press.

Levelt, W. J. M., Praamstra, P., Meyer, A. S., Helenius, P., and Salmelin, R. (1998). An MEG study of picture naming. *J. Cogn. Neurosci.* 10, 553–567.

Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–75.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., and Havinga, J. (1991). The time course of lexical access in speech production: a study of picture naming. *Psychol. Rev.* 98, 122–142.

Liljestrom, M., Hulten, A., Parkkonen, L., and Salmelin, R. (2009). Comparing MEG and fMRI views to naming actions and objects. *Hum. Brain Mapp.* 30, 1845–1856.

Liu, H., Agam, Y., Madsen, J. R., and Kreiman, G. (2009). Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron* 62, 281–290.

Luce, R. D. (1986). *Response Times.* New York: Oxford University Press.

Luck, S. J., Chelazzi, L., Hillyard, S. A., and Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J. Neurophysiol.* 77, 24–42.

Lupker, S. J. (1979). The semantic nature of response competition in the picture–word interference task. *Mem. Cogn.* 7, 485–495.

Lupker, S. J. (1982). The role of phonetic and orthographic similarity in picture-word interference. *Can. J. Psychol.* 36, 349–367.

Maess, B., Friederici, A. D., Damian, M., Meyer, A. S., and Levelt, W. J. M. (2002). Semantic category interference in overt picture naming: an MEG study. *J. Cogn. Neurosci.* 14, 455–462.

Mägiste, E. (1984). Stroop tasks and dichotic translation: the development of interference patterns in bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 304–315.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., and Caramazza, A. (2007). Lexical selection is not by competition: a reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 503–535.

Martin, A. (2007). The representation of object concepts in the brain. *Annu. Rev. Psychol.* 58, 25–45.

Meyer, A. S., Belke, E., Telling, A. L., and Humphreys, G. W. (2007). Early activation of object names in visual search. *Psychon. Bull. Rev.* 14, 710–716.

Meyer, A. S., and Damian, M. F. (2007). Activation of distractor names in the picture-picture interference paradigm. *Mem. Cogn.* 35, 494–503.

Meyer, A. S., Sleiderink, A. M., and Levelt, W. J. M. (1998). Viewing and naming objects: eye movements during noun phrase production. *Cognition* 66, 25–33.

Miller, J., Riehle, A., and Requin, J. (1992). Effects of preliminary perceptual output on neuronal activity of the primary motor cortex. *J. Exp. Psychol. Hum. Percept. Perform.* 18, 1121–1138.

Miozzo, M., and Caramazza, A. (2003). When more is less: a counterintuitive effect of distractor frequency in picture–word interference

paradigm. *J. Exp. Psychol. Gen.* 132, 228–252.

Morsella, E., and Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 555–563.

Navarrete, E., Basagni, B., Alario, X. F., and Costa, A. (2006). Does word frequency affect lexical selection in speech production? *Q. J. Exp. Psychol.* 10, 1681–1690.

Navarrete, E., and Costa, A. (2005). Phonological activation of ignored pictures: further evidence for a cascade model of lexical access. *J. Mem. Lang.* 53, 359–377.

Navarrete, E., Mahon, B. Z., and Caramazza, A. (2010). The cumulative semantic cost does not reflect lexical selection by competition. *Acta Psychol. (Amst.)* 134, 279–289.

O'Craven, K., and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* 12, 1013–1023.

Oppenheim, G. M., Dell, G. S., and Schwartz, M. F. (2007). Cumulative semantic interference as learning. *Brain Lang.* 103, 175–176.

Oppenheim, G. M., Dell, G. S., and 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.

Peelen, M. V., Fei-Fei, L., and Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* 460, 94–97.

Peterson, R. R., and Savoy, P. (1998). Lexical selection and phonological encoding during language production: evidence for cascaded processing. *J. Exp. Psychol. Learn. Mem. Cogn.* 24, 539–557.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59–108.

Rodriguez-Fornells, A., Schmitt, B. M., Kutas, M., and Munte, T. F. (2002). Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming. *Neuropsychologia* 40, 778–787.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42, 107–142.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition* 64, 249–284.

Roelofs, A. (2003). Goal-referenced selection of verbal action: modelling attentional control in the Stroop task. *Psychol. Rev.* 110, 88–125.

Roelofs, A. (2006). Context effects of pictures and words in naming

objects, reading words, and generating simple phrases. *Q. J. Exp. Psychol. (Hove)* 59, 1764–1784.

Roelofs, A. (2008). Attention to spoken word planning: chronometric and neuroimaging evidence. *Lang. Linguist. Compass* 2, 389–405.

Roelofs, A., Piai, V., and Schriefers, H. (2011). Selective attention and distractor frequency in naming performance: comment on Dhooge and Hartsuiker (2010). *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 1032–1038.

Sahin, N., Pinker, S., Cash, S., Schomer, D., and Halgren, E. (2009). Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science* 326, 445–449.

Salmelin, R., Hari, R., Lounasmaa, O. V., and Sams, M. (1994). Dynamics of brain activation during picture naming. *Nature* 368, 463–465.

Schendan, H. E., and Kutas, M. (2003). Time course of processes and representations supporting visual object identification and memory. *J. Cogn. Neurosci.* 15, 111–135.

Schiller, N. O., Münte, T. F., Horemans, I., and Jansma, B. M. (2003). The influence of semantic and phonological factors on syntactic decisions: an event-related brain potential study. *Psychophysiology* 40, 869–877.

Schmitt, B. M., Münte, T. F., and Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology* 37, 473–484.

Schmitt, B. M., Schiltz, K., Zaake, W., Kutas, M., and Münte, T. F. (2001). An electrophysiological analysis of the time course of conceptual and syntactic encoding during tacit picture naming. *J. Cogn. Neurosci.* 13, 510–522.

Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., and Thompson-Schill, S. L. (2009). Localizing interference during naming: convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proc. Natl. Acad. Sci. U.S.A.* 106, 322–327.

Schriefers, H., Meyer, A. S., and Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: picture-word interference studies. *J. Mem. Lang.* 29, 86–102.

Sitnikova, T., West, W. C., Kuperberg, G. R., and Holcomb, P. J. (2006). The neural organization of semantic memory: electrophysiological activity suggests feature-based segregation. *Biol. Psychol.* 71, 326–340.

Soros, P., Cornelissen, K., Laine, M., and Salmelin, R. (2003). Naming actions and objects: cortical dynamics in healthy adults and in an anomic patient with a dissociation in action/object naming. *Neuroimage* 19, 1787–1801.

Starreveld, P. A., and La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 686–698.

Starreveld, P. A., and La Heij, W. (1996). Time course analysis of semantic and orthographic context effects in picture-naming. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 896–918.

Strijkers, K., Costa, A., and Thierry, G. (2010). Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. *Cereb. Cortex* 20, 912–928.

Strijkers, K., Holcomb, P., and Costa, A. (2011). Conscious intention to speak facilitates lexical access during overt object naming. *J. Mem. Lang.* 65, 345–362.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., and Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc. Natl. Acad. Sci. U.S.A.* 94, 14792–14797.

Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of processing in the human visual system. *Nature* 381, 520–522.

Thorpe, S. J., and Fabre-Thorpe, M. (2001). Neuroscience seeking categories in the brain. *Science* 291, 5502.

Usher, M., and McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* 108, 550–592.

van Turennout, M., Hagoort, P., and Brown, C. M. (1997). Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 787–806.

van Turennout, M., Hagoort, P., and Brown, C. M. (1998). Brain activity during speaking: from syntax to phonology in 40 milliseconds. *Science* 280, 572–574.

VanRullen, R., and Thorpe, S. J. (2002). Surfing a spike wave down the ventral stream. *Vision Res.* 42, 2593–2615.

Verhoef, K. M. W., Roelofs, A., and Chwilla, D. J. (2009). Role of inhibition in language switching: evidence from event-related brain potentials in overt picture naming. *Cognition* 110, 84–99.

Verhoef, K. M. W., Roelofs, A., and Chwilla, D. J. (2010). Electrophysiological evidence for endogenous control of attention in switching between languages in overt picture naming. *J. Cogn. Neurosci.* 22, 1832–1843.

Warrington, E. K., and McCarthy, R. (1987). Categories of knowledge: further fractionations and an attempted integration. *Brain* 110, 1273–1296.

Warrington, E. K., and Shallice, T. (1984). Category specific semantic impairments. *Brain* 197, 829–854.

Zhang, Q. F., and Damian, M. F. (2009a). The time course of segment and tone encoding in Chinese spoken production: an event-related potential study. *Neuroscience* 163, 252–265.

Zhang, Q. F., and Damian, M. F. (2009b). The time course of semantic and orthographic encoding in Chinese word production: an event-related potential study. *Brain Res.* 1273, 92–105.