The spatial and temporal signatures of word production components: a critical update

Peter Indefrey[1,2]*

[1] Institut für Sprache und Information, Heinrich Heine Universität Düsseldorf, Düsseldorf, Germany
[2] Donders Institute, Centre for Cognitive Neuroimaging, Radboud University Nijmegen, Nijmegen, Netherlands

In the first decade of neurocognitive word production research the predominant approach was brain mapping, i.e., investigating the regional cerebral brain activation patterns correlated with word production tasks, such as picture naming and word generation. Indefrey and Levelt (2004) conducted a comprehensive meta-analysis of word production studies that used this approach and combined the resulting spatial information on neural correlates of component processes of word production with information on the time course of word production provided by behavioral and electromagnetic studies. In recent years, neurocognitive word production research has seen a major change toward a hypothesis-testing approach. This approach is characterized by the design of experimental variables modulating single component processes of word production and testing for predicted effects on spatial or temporal neurocognitive signatures of these components. This change was accompanied by the development of a broader spectrum of measurement and analysis techniques. The article reviews the findings of recent studies using the new approach. The time course assumptions of Indefrey and Levelt (2004) have largely been confirmed requiring only minor adaptations. Adaptations of the brain structure/function relationships proposed by Indefrey and Levelt (2004) include the precise role of subregions of the left inferior frontal gyrus as well as a probable, yet to date unclear role of the inferior parietal cortex in word production.

Keywords: language production, word production, picture naming, neuroimaging

INTRODUCTION: RECENT DEVELOPMENTS IN NEUROCOGNITIVE WORD PRODUCTION RESEARCH

The goal of neurocognitive word production research is twofold: to understand how the processes represented in functional models of word production are implemented in the brain and to improve functional models by testing their predictions at the brain level. In the first decade of neurocognitive word production research the predominant approach was brain mapping. Researchers investigated the regional cerebral brain activation correlated with word production tasks, such as picture naming and word generation, compared to more or less low-level control tasks. As is the case in most new research fields, the aim of this approach was not so much to test specific hypotheses but to gain a first insight into the behavior of the system under investigation, in this case the neural system supporting word production. This research yielded a wealth of data about which brain regions respond to tasks that were considered word production tasks, such as picture naming and verb or noun generation. Other tasks that were not typically used to study word production nonetheless involve word production components, such as word and pseudoword reading. Indefrey and Levelt (2000, 2004, see also Indefrey, 2007) conducted comprehensive meta-analyses of word production studies that had used the mapping approach. They first analyzed the tasks with respect to lead-in processes preceding word production and core word production processes as assumed by psycholinguistic models of word production. For the identification of candidate brain

regions subserving these components they then followed a simple (some may say "simplistic") heuristic principle: "*If, for a given processing component, there are subserving brain regions, then these regions should be found active in all experimental tasks sharing the processing component, whatever other processing components these tasks may comprise. In addition, the region(s) should not be active in experimental tasks that do not share the component.*" (Indefrey and Levelt, 2000). These analyses yielded a set of candidate areas corresponding to certain word production components but, of course, the validity of the identification of any of these areas depends on the validity of the underlying task analysis. Thus, essentially, these analyses generated a set of hypotheses that needed confirmation (or falsification) from independent data. A first kind of hypothesis-testing was performed in Indefrey and Levelt (2004), combining the resulting spatial information on potential neural correlates of component processes of word production with an independent estimate of the time course of word production components provided by behavioral and electrophysiological studies. If, they reasoned for example, the left posterior superior temporal gyrus (STG) was involved in word form retrieval then the time course of its activation in picture naming should fall in the interval of 250–330 ms after picture onset suggested by chronometric studies for word form retrieval. Data from the few available magnetoencephalographic (MEG) studies on picture naming that provided both spatial and temporal information confirmed the proposed assignment of component processes to brain areas in

that they were largely compatible with the predicted time windows of activation.

The resulting spatiotemporal model of word production does not only predict time windows of activation but also modulatory effects of psycholinguistic variables on the activation of specific brain regions at a specific time. In recent years, neurocognitive word production research has seen a major change toward a hypothesis-testing approach. This approach is characterized by the design of experimental variables modulating single component processes of word production and testing for predicted effects on spatial or temporal neurocognitive signatures of these components. This change has been accompanied by an impressive broadening of the spectrum of measurement and analysis techniques. Both in functional magnetic resonance imaging (fMRI) and in electroencephalography (EEG) methods have been developed that allow for overt speaking during experiments (for fMRI see, e.g., de Zubicaray et al., 2001; Grabowski et al., 2006; Christoffels et al., 2007b; Heim et al., 2009b; Hocking et al., 2009; for EEG see, e.g., Christoffels et al., 2007a; Koester and Schiller, 2008; Costa et al., 2009; Strijkers et al., 2010). Overt pronunciation provides on-line voice onset time and error data and, hence, some confirmation that a targeted psycholinguistic effect was indeed present in a neurocognitive experiment, thus increasing the likelihood that an observed hemodynamic or electrophysiological effect is indeed due to the same variable that causes an effect in the corresponding psycholinguistic experiment. On-line behavioral data can, furthermore, be used as predictors for the analysis of the neuroimaging data.

Secondly, the number of studies that used techniques that provide both spatial and temporal neurocognitive data increased over the last years. In addition to MEG studies (e.g., Sörös et al., 2003; Hultén et al., 2009) and the use of intracranial electrophysiology in neurosurgical patients (Sahin et al., 2009; Edwards et al., 2010) the main development has been the use of transcranial magnetic stimulation (TMS) as a tool for temporarily stimulating or interfering with neuronal activity in specific brain regions at a specific time (Schuhmann et al., 2009; Acheson et al., 2011). Similar to electrocortical stimulation and lesion-symptom mapping, TMS has the potential to provide evidence as to the functional necessity of a targeted brain area. This evidence is thus complementary to fMRI or positron emission tomography (PET) data that inform about the involvement of brain areas in cognitive processes but not their necessity.

A third recent development in fMRI research on word production is the use of repetition suppression or adaptation paradigms (e.g., Graves et al., 2008; Peeva et al., 2010). In the standard fMRI approach a neuronal population involved in a certain cognitive process (e.g., lexical word form retrieval in word production) is identified by subtracting the brain activation of a control condition that does not (or to a smaller extent) contain that cognitive process (for example by using the production of pseudowords that are not lexically stored). In many cases, however, finding the right control condition is extremely difficult, because in addition to the process of interest there are other unavoidable differences between the active condition and the control condition (pseudowords also differ from words in that they have no meaning). The repetition suppression paradigm, by contrast, exploits the fact that

the activation of just that neuronal population that is involved in the process of interest tends become smaller the more often that process is repeated. Experimenters can use repetition suppression to target neuronal populations subserving very specific cognitive components. The study of Peeva et al. (2010) is a nice example of this approach. In one condition, they repeated bisyllabic pseudowords (e.g., fublo, blofu, fublo...) consisting of two constant syllables. In another condition they kept the repetition of phonemes constant but varied syllable structure (lofub, fublo, lofub . . .). As a result the activation of neuronal populations interested in the specific syllables "fu" and "blo" is suppressed over time in the first but not the second condition. In their study the left ventral premotor cortex showed this behavior so it could be concluded that this region contains neurons representing complete syllables.

Finally new analysis techniques for measuring anatomical connections (diffusion tensor imaging, DTI) and modeling the interaction between brain areas (structural equation modeling, SEM; dynamic causal modeling, DCM, independent component analysis, ICA) have begun to be applied to word production (Saur et al., 2008; Tourville et al., 2008; Eickhoff et al., 2009; Heim et al., 2009a; van de Ven et al., 2009). To understand how the processes represented in functional models of word production are implemented in the brain, such approaches - together with methods providing combined spatial and temporal information - are needed to test theoretical assumptions about directions of information flow and interactions between processing components.

In the next two sections I will briefly recapitulate the cascade of processing components involved in word production and their estimated time windows. In subsequent sections I will then discuss the neural correlates of each processing component as presented in Indefrey and Levelt, 2004, henceforth I&L) and the more recent evidence about the neural implementation of these components.

COMPONENT PROCESSES OF WORD PRODUCTION

Models of language production (Garrett, 1980; Stemberger, 1985; Dell, 1986; Butterworth, 1989; Levelt, 1989; Caramazza, 1997; Dell et al., 1997; Levelt et al., 1999) agree that there are processing levels of meaning, form, and articulation. Speaking normally starts by preparing a preverbal conceptual representation (message). To describe a football game, the sports commentator must, for example, conceptualize events ("Ronaldo was replaced *before* the team scored the first goal." "The team scored the first goal *after* Ronald had been replaced.") and spatial configurations ("The defender was standing *behind* Ronaldo." "Ronaldo was standing *in front* of the defender.") in a particular order. These planning processes are called linearization and perspective taking (Levelt, 1989). The speaker must also take into account the audience's knowledge of the world and whether or not Ronaldo was mentioned earlier when referring to him ("Ronaldo," "he," "the Brazilian"). RONALDO and BRAZILIAN are both lexical concepts, that is, concepts for which there are words. Assuming that the speaker has decided that the concept BRAZILIAN is the appropriate one, the corresponding word "Brazilian" must be selected. It is known that at this stage semantically related lexical entries such as "South American" or "Argentinean" are also activated. Occasionally one of them will

be erroneously selected, which a listener may notice as a speech error (Argentinean) or not (South American). Levelt et al. (1999) assume this selection process to take place in a part of the mental lexicon (lemma level) that is linked to the conceptual level and contains information about the grammatical properties of words, such as word class or grammatical gender. It is only after the selection of a lemma that its corresponding sound properties (lexical phonological code, a sequence of phonemes) are retrieved at the word form level and fed into a phonological encoding process. In the case of single word utterances, this process mainly combines the retrieved phonemes into syllables and assigns a stress pattern. The output of phonological encoding is an abstract phonological representation (phonological word) containing syllables and prosodic information. In the process of phonetic encoding this representation is translated into an abstract articulatory representation, the articulatory score. For frequent syllables, articulatory representations may be retrieved from a store (syllabary). Finally, the abstract articulatory representation is realized during articulation by coordinating and executing the activation of the speech musculature.

THE TIME COURSE OF THE COMPONENT PROCESSES IN WORD PRODUCTION

Based on the comparison of chronometric data from reaction time studies, modeling data, and electrophysiological studies, I&L provided the following estimates for the duration of the different processing components in the picture naming task: Conceptual preparation (from picture onset to selection of target concept) 175 ms, Lemma retrieval 75 ms, Phonological code retrieval 80 ms, Syllabification 125 ms (25 ms per phoneme), Phonetic encoding (till onset of articulation) 145 ms. I&L cautioned against a "too rigid interpretation" of these numbers for two reasons. Firstly, they pointed out that the insecurity due to the ranges from which the estimates were taken accumulated with every processing stage. Secondly, the estimate of 600 ms for the onset of articulation was based on studies using repeated naming of the same pictures. Furthermore, naming latencies depend on numerous variables such as task variations, picture context, picture quality, familiarity of the depicted object, and length or lexical frequency of the object name. As can be seen in **Table 1**, the reported naming latencies from studies providing more recent evidence on the time windows of the processing stages of language production range from 470 to over 2000 ms so that the question arises how to rescale the I&L estimates to shorter or longer naming latencies. Although there is no simple answer to this question, it can be said that a linear rescaling of the duration of all processing stages can only be the last resort and is inadequate whenever the reason for shorter or longer naming latencies can be identified. A good example in this respect is an eye-tracking study by Huettig and Hartsuiker (2008) who measured very long naming latencies of more than 2000 ms. They asked their subjects to name objects based on a question (e.g., "What is the name of the circular object?"). The objects were presented in the context of three other objects that were categorically related, form-related, or unrelated to the target object. Huettig and Hartsuiker (2008) reasoned very plausibly that in this paradigm long naming latencies arise from additional lead-in processes ("Wrapping up comprehension instruction; Inspecting

display; Determining categories; Matching target category to those of objects") and possibly prolonged conceptual processing and lemma retrieval of the target object (due to competition of related objects), but are unlikely to arise after the onset of phonological encoding of the name of the target object. As a consequence, they subtracted the unaltered estimated duration of 350 ms for form encoding processes from the naming latencies and indeed found increased fixation proportions to categorically related competitor objects (presumably indicating competition at the lemma level) up to over 1500 ms after display onset. Conversely, short naming latencies due to short target words (see Schuhmann et al., 2009) are unlikely to arise at conceptual processing or lemma retrieval stages, so that the duration estimate of these stages is best left unaltered. Yet other factors, such as lexical frequency may themselves affect both lemma and form processing stages and are known to be correlated with conceptual factors such as item familiarity, so that indeed all processing stages could be affected, thus justifying a rescaling of the durations of all component processes (although, of course, a linear rescaling would not take into account a differential impact of lexical frequency on specific processes). Keeping these considerations in mind, we can now assess in how far the data obtained in the studies listed in **Table 1** can improve the I&L estimates for the different processing stages of word production.

CONCEPTUAL PREPARATION

The I&L estimate for the duration of conceptual preparation until the selection of a target concept was based on data showing the availability of information about whether a picture showed an animal or not (Thorpe et al., 1996, around 150 ms; Schmitt et al., 2000, around 200 ms). Two more recent studies (Rodriguez-Fornells et al., 2002; Guo et al., 2005) also used paradigms in which go/nogo responses were contingent on an animal/object decision. They report slightly later onsets of the N200 nogo responses (approximately 260 and 200 ms) representing upper boundaries for the availability of "animal" information. Zhang and Damian (2009a) used a living/non-living decision with an N200 response starting around 200 ms. In a study by Hanulová et al. (2011) N200 nogo responses were contingent on a man-made/natural decision, showing an N200 onset latency of around 300 ms. Note that the latency of N200 responses includes the time needed for the decision to withhold the button press, so that the information on which this decision is based probably is available slightly earlier and more truly reflected in the time point at which differences between ERP waveforms corresponding to the different levels of a conceptual variable emerge. Habets et al. (2008) asked their subjects to describe a sequence of events either in their natural temporal order using the temporal conjunction "after" or in reversed order using "before." Deciding on a particular linearization (Levelt, 1989) of events to be named is an essential aspect of the conceptual preparation stage and accessing the lemmas of the words "before" or "after" depends on that decision. Habets et al. (2008) found a difference between "before" and "after" ERP waveforms starting around 180 ms after picture onset. An earlier effect has been reported by Abdel Rahman and Sommer (2008) who taught their subjects novel names for novel objects but either provided additional (rather complex) conceptual information or not. Even though later naming of such novel objects took much longer (around 1200 ms)

Table 1 | Overview of recent studies providing evidence about the time course of processing stages in word production.

Study	Manipulation	Onset/time window of effect (ms)	Voice onset time (ms)	Processing stage
Abdel Rahman and Sommer (2003), ERP: LRP and N200 latencies	Easy (size) compared to hard (diet) semantic decision and phonological (vowel/consonant onset) decision in dual task	Exp. 1: LRP easy 35 earlier than LRP hard, no nogo LRP hard, Exp. 2: N200 easy 28 ms earlier than N200 hard	No overt naming	Conceptual preparation/phonological code retrieval
Abdel Rahman and Sommer (2008), ERP: waveform difference	Conceptual knowledge in novel object naming	120	~1200	Conceptual preparation
Aristei et al. (2011), ERP: waveform difference	Effects of categorically and associatively related distractors, Effect of semantic blocking	Distractor effect: 200, blocking effect: 250, interaction: 200	~770	Lemma retrieval
Camen et al. (2010), ERP: temporal, segmentation analysis	Gender monitoring, phoneme monitoring, 1st syllable, 2nd syllable	Gender: 270–290, 1st: 210–290, 2nd: 480	No overt naming	Lemma retrieval/phonological code retrieval
Cheng et al. (2010), ERP: waveform difference	High vs. low name agreement	100–150, 250–350, >800	No overt naming	Conceptual preparation/lemma retrieval
Costa et al. (2009), ERP: waveform difference	Cumulative semantic interference	200–380	~840	Lemma retrieval
Christoffels et al. (2007a), ERP: waveform difference	Cognate effect	275–375	~720	Lemma retrieval/form encoding
Guo et al. (2005), ERP: N200 peak latencies	Semantic (animal vs. object) and phonological (onset consonant) decision	Semantic: 307, phonological: 447	Delayed naming	Lemma retrieval/phonological code retrieval
Habets et al. (2008), ERP: waveform difference	Conceptual linearization	180	~1360	Conceptual preparation
Hanulová et al. (2011), ERP: N200 latencies	Semantic (man-made vs. natural) and phonological (onset consonant) decision	Semantic: 307, phonological: 393	No overt naming	Lemma retrieval/phonological code retrieval
Laganaro et al. (2009b), ERP: temporal, segmentation analysis	Lexical frequency (healthy controls), semantically impaired anomia vs. control, phonologically impaired anomia vs. control	270-330, semantic: 100-310, phonological: 390-430	Delayed naming	Conceptual preparation/phonological code retrieval
Laganaro et al. (2009a), ERP: temporal, segmentation analysis	Semantically impaired anomia vs. control, phonologically impaired anomia vs. control	Semantic: 90–200, phonological: 340–430	Delayed naming	Conceptual preparation/phonological code retrieval
Morgan et al. (2008), RT	Facilitation of phonologically related probe naming	No effect at 150, 350	~800	Form encoding
Rodriguez-Fornells et al. (2002), ERP: N200 latencies	Semantic (animal vs. object) and phonological (vowel vs. consonant) decision	Semantic: 264, phonological: 456	No overt naming	Conceptual preparation/phonological code retrieval

(Continued)

Table 1 | Continued

Study	Manipulation	Onset/time window of effect (ms)	Voice onset time (ms)	Processing stage
Schiller et al. (2003), ERP: N200 peak latencies	Metrical (stress on first or second syllable) vs. syllabification (consonant in first or second syllable) decision	Metrical: 255, syllabic: 269	No overt naming	Phonological encoding/syllabification
Schiller (2006), ERP: N200 peak latencies	Decision on lexical stress on, 1st syllable, 2nd syllable	1st: 475, 2nd: 533	∼800	Phonological encoding/syllabification
Strijkers et al. (2010), ERP: waveform difference	Cognate status; frequency	Cognate status: 200, frequency: 172	∼700	Lemma retrieval/form encoding
Zhang and Damian (2009a), ERP: N200 latencies	Decision on semantics (animacy) and orthography (left/right structure character in Chinese	Semantics: onset around 200, peak 373; orthography: onset around 350, peak 541	No overt naming	Conceptual preparation/orthographic code retrieval
Zhang and Damian (2009b), ERP: N200 latencies	Decision on segments and tones in Chinese	Segments: onset 283–293 peak 592, Tones: onset 483–493 peak 599	No overt naming	Phonological code retrieval/phonological encoding

than the typical durations found for familiar objects, Abdel Rahman and Sommer (2008) found that the presence of conceptual information affected the ERP waveforms already around 120 ms. The authors interpret this early effect as reflecting an influence of conceptual knowledge on perceptual analysis and object recognition. Results from an ERP study comparing the naming of pictures with high and low name agreement (Cheng et al., 2010) also show an early effect (100–150 ms) probably due to object recognition difficulty being one source of low name agreement.

Two recent studies by Laganaro et al. (2009a,b) compared electrophysiological picture naming responses between two groups of anomic patients and healthy controls. The ERP waveforms of anomic patients with a semantic impairment (as assessed in independent testing) differed significantly from the ERP waveforms of healthy controls in a time window between 90 and 310 ms, whereas anomic patients with a phonological impairment showed differences in a later time window corresponding to the form encoding stage (see below).

In sum, more recent studies reported slightly later availability of a type of conceptual information ("animal" or "animate") that is likely to be relevant for subsequent lemma retrieval. The median estimate of all five studies showing ERP effects related to the availability of this kind of information is 200 ms, i.e., 25 ms later than the I&L estimate for the duration of conceptual preparation. Effects presumably related to perceptual processes were earlier (100–150 ms). The availability of other types of conceptual information, such as "man-made" or "natural" may take longer. However, as shown by Abdel Rahman and Sommer (2003), the relatively late availability of more peripheral conceptual information, such as the kind of food an animal prefers, does not delay lemma and word form retrieval, suggesting that conceptual processing

continues to run in parallel with subsequent processing stages. At present we simply do not know which kind of conceptual information is necessary and sufficient for lemma retrieval. It is plausible to assume that the activation of a target concept "dog" includes "animal" and "animacy" information, but not necessarily "typical food" information. The latter may only be retrieved on demand.

LEMMA RETRIEVAL

The I&L estimate for the duration of lemma retrieval was based on mathematical modeling of the semantic interference effect (Levelt et al., 1991; Roelofs, 1992), suggesting a lemma retrieval duration of 100–150 ms and electrophysiological data by Schmitt et al. (2001) suggesting that grammatical gender information is available about 75 ms later than conceptual information about the physical weight of a depicted item. Given that gender may be retrieved subsequent to lemma activation, the latter data point was considered an upper boundary for lemma retrieval. Recently an absolute measure of the onset of the availability of gender information has been reported in an ERP study by Camen et al. (2010) using a technique (temporal segmentation analysis) that analyzes identities and differences between topographic scalp distributions over time. Around 270–290 ms after picture onset the authors found a scalp distribution difference between (French) picture names that matched or didn't match a pre-specified grammatical gender. Considering that gender availability is an upper boundary for lemma retrieval, this time fits well both with the I&L estimate for lemma selection (250 ms) and even better under the assumption of a slightly later start of lemma retrieval due to longer conceptual preparation (see previous section).

The time course of lemma retrieval has also been studied by manipulating the degree of lexical competition which according

to Levelt (1989) and Levelt et al. (1999) takes place at the lemma level. Costa et al. (2009) used a cumulative semantic interference paradigm. Picture naming latencies increased with the number of preceding items from the same semantic category and so did the amplitude of the corresponding ERP waveforms in a time window between 200 and 380 ms. In a complex design combining semantic blocking and picture-word interference, Aristei et al. (2011) found increases in naming latencies of category coordinates for both manipulations. They also found corresponding ERP effects starting around 200–250 ms.

In sum, more recent evidence suggests that lemma retrieval should start around 200 ms (onset of competition effects) and the lemma should be selected before 270–290 ms (gender available). These data are compatible with the I&L estimate of 75 ms for lemma retrieval duration.

Three other studies (Christoffels et al., 2007a; Laganaro et al., 2009b; Strijkers et al., 2010) investigated the time course of the electrophysiological effects of lexical frequency and cognate status (the target language name of the depicted object sounds/doesn't sound similar in another language spoken by the subject) manipulations. Given that these effects can in principle arise at different processing levels (for a discussion see Hanulová et al., 2011) these studies did not test the I&L time course estimates but rather used them to obtain evidence as to the processing stage affected by lexical frequency and cognate status. Strijkers et al. (2010) found early effect onsets (170–200 ms) and a correlation of ERP amplitude with voice onset time suggesting that both variables influence lemma retrieval. By contrast, a frequency effect reported by Laganaro et al. (2009b) and a cognate effect observed by Christoffels et al. (2007a) were in later time windows (frequency effect: 270–330 ms; cognate effect: 275–375 ms) better compatible with an influence of these factors at a word form encoding stage (see below).

PHONOLOGICAL CODE RETRIEVAL

The I&L estimate for the duration of phonological code retrieval was based on the difference between the lateralized readiness potential (LRP) onsets for a grammatical gender compared to a first phoneme decision in a study by van Turennout et al. (1998). There are now a number of studies providing absolute time information in the form of latencies of the onset of N200 nogo responses for decisions on the first phoneme of a depicted object. Note that phoneme monitoring probably taps into a syllabified representation, because reaction times depend on syllable position (Wheeldon and Levelt, 1995). This means that the availability of information about the first phoneme strictly speaking marks the beginning of the phonological encoding (syllabification) process, which, however, does not have to wait until all lexically specified phonemes have been retrieved. In addition to their data on semantic decisions discussed above, Rodriguez-Fornells et al. (2002), Guo et al. (2005), and Hanulová et al. (2011) also provided N200 onset times for first phoneme decisions around 460, 400, and 390 ms. Zhang and Damian (2009b) report an N200 onset around 290 ms. Camen et al. (2010) using temporal segmentation analysis observed much earlier ERP effects related to a decision on the first phoneme (210–290 ms). The median estimate of the five studies is 390 ms. N200 peak differences

between the availability of semantic and first phoneme information taken from Rodriguez-Fornells et al. (2002), Guo et al. (2005), and Hanulová et al. (2011) were approximately 170, 140, and 90 ms. Schmitt et al. (2000) reported a N200 peak difference of 90 ms. Adding the estimate of 200 ms for conceptual preparation, these numbers suggest an availability of first phoneme information between 290 and 370 ms (median 310 ms) after picture onset.

Evidence for the retrieval of word form information that does not rely on phoneme monitoring comes from a study by Morgan et al. (2008). In their study, participants named two depicted objects but on some trials the first object to be named was replaced by a written word either 150 or 350 ms after picture onset. Participants were instructed to name the word in these trials. Naming of the word was facilitated when it was phonologically related to the name of the replaced object (e.g., object "bed," word "bell") but only when the object had been seen for 350 ms. The initial phonemes of the object's name, thus, had not yet been retrieved after 150 ms but had been retrieved after 350 ms.

In sum, assuming that phonological code retrieval starts around 275 ms (200 ms conceptual preparation + 75 ms lemma retrieval) N200 onset data of recent studies suggest a longer duration ($390 - 275 = 115$ ms) than estimated by I&L. Based on the temporal difference in the availability of semantic and first phoneme information as measured by peak rather than onset latencies of the N200, the estimated duration would be shorter ($310 - 275 = 35$ ms). It seems, therefore, that the I&L estimate of 80 ms is a reasonable figure. Nonetheless, its interpretation as the "duration" of phonological code retrieval should probably be reconsidered. Given that it is solely based on measures of the availability of first phoneme information, this estimate is much more appropriately interpreted as the time interval from the beginning of phonological code retrieval to the beginning of phonological encoding. There is no reason to assume that phonological encoding waits until all phonemes have been retrieved. So phonological code retrieval may well go on after phonological encoding has started.

PHONOLOGICAL ENCODING (SYLLABIFICATION AND METRICAL RETRIEVAL)

The I&L estimate for the duration of phonological encoding (syllabification) was based on phoneme monitoring reaction times in picture naming (Wheeldon and Levelt, 1995) and the time difference between LRP onsets for decisions on the first and the last phoneme of the picture name reported by van Turennout et al. (1997). Wheeldon and Levelt (1995) measured a 125 ms difference between RTs for the first and the last phoneme of bisyllabic words with on average six phonemes. van Turennout et al. (1997) measured a corresponding 80 ms difference for words with on average 4.5 phonemes. The I&L estimate (25 ms/phoneme) was therefore slightly too long and should be corrected to 20 ms/phoneme. More recently Schiller (2006) measured an N200 peak difference of 58 ms between a lexical stress decision on the first and the second syllable of bisyllabic words, confirming earlier data by Wheeldon and Levelt (1995) and van Turennout et al. (1997) that suggested a phonological encoding duration of around 55 ms per syllable. Starting from the estimate of 355 ms for the beginning of

phonological encoding, the process should last 100–120 ms for a bisyllabic word of 5–6 phonemes, ending around 455–475 ms.

Independent evidence about the duration of phonological code retrieval and encoding comes from the clinical studies of Laganaro et al. (2009a,b). As mentioned above, they also studied differences between the ERP waveforms of anomic patients with a phonological impairment and healthy control subjects. As the predominant naming errors of these patients were phonological paraphasias and neologisms it is plausible to assume that their impairment was related to phonological code retrieval or encoding problems. Differences in the patients' waveforms were observed between 340 and 430 ms which is compatible with the estimated time windows of phonological code retrieval and encoding.

PHONETIC ENCODING

The I&L estimate for the duration of phonological code retrieval until the initiation of articulation was based on the difference between an average voice onset time for the undistracted naming of repeatedly presented pictures (600 ms) and the end of the phonological encoding operation (455 ms). Given that the more recent data used above to update the estimates on the preceding processing stages come from studies using repeated picture presentation, the I&L estimate still seems adequate. It should be noted, though, that the estimate of 455 ms is an upper boundary, because phonetic encoding may start as soon as the first syllable has been phonologically encoded.

INTERIM SUMMARY

Except for small adaptations of the duration of conceptual preparation and syllabification the time windows for the processing stages of word production estimated in I&L have largely been confirmed in more recent studies and hence are now based on a broader data base. **Table 2** presents an updated version of the estimated onset times and durations. It should be noted that the available estimates of onsets and durations of component processes do not provide conclusive evidence for or against serial or cascaded transitions between subsequent operations. Non-overlapping time windows should, therefore, not be interpreted as indicating strictly serial processing stages. Specifically, as discussed in the corresponding sections above, the estimates for the onsets

Table 2 | Estimated onset times and durations for operations in spoken word encoding.

Operation	Onset (ms)	Duration (ms)
Conceptual preparation	0	200+
Lemma retrieval	200	75*
Form encoding		
Phonological code retrieval	275*	
Syllabification	355	20 per phoneme, 50–55 per syllable
Phonetic encoding	455*	
Articulation	600	

+Continues after relevant conceptual information for lexical access has become available, *upper boundary.

of phonological code retrieval and phonetic encoding are upper boundaries based on the evidence for the duration of the preceding stages and, hence, do not preclude earlier onsets. By contrast, onset of lemma retrieval and duration of conceptual preparation were estimated independently and lemma retrieval seems to begin at about the same time as relevant conceptual information (animacy/animal) becomes available. It is still conceivable, however, that more specific conceptual information needed to select a particular lemma among a number of competitors comes in later.

BRAIN AREAS INVOLVED IN WORD PRODUCTION

Brain activation studies on language production using the mapping approach have mainly used a limited set of tasks, namely picture naming, word generation, and word or pseudoword reading. These tasks differ with respect to the cognitive processes preceding word production as such, which have been termed lead-in processes by Indefrey and Levelt (2000). Picture naming but not the other tasks, for example, involves visual object recognition. Reading tasks involve visual word recognition through grapheme-to-phoneme recoding or accessing a visual input lexicon. The lead-in processes of word generation include recognition of the stimulus words and various cognitive processes from association to visual imagery, even the retrieval of whole episodes from long-term memory. I&L further assumed that word production tasks also differ with respect to the point at which they enter the cascade of the core processes of word production. While in picture naming and word generation the result of the lead-in processes is a concept for which the appropriate lemma is then retrieved, this is not the case for the reading tasks where the activation of lemma and conceptual representations are part of word recognition rather than production, so that the flow of activation is reversed compared to tasks that start out from a conceptual representation. The pronunciation of written pseudowords, finally, is a production task that enters the cascade of word production processes after the lexical stages. For lack of a lexical entry, a phonemic representation of a written pseudoword is created by grapheme-to-phoneme conversion. This phonemic representation can then be fed into the syllabification process and the subsequent phonetic and articulatory stages.

Taken together, the properties of the tasks are such that (due to the lead-in processes) no single task allows for the identification of neural correlates of all and only the core word production components that have been psycholinguistically identified. Core word production components are on the other hand shared between tasks, so that their neural correlates may be identified as common activation areas across tasks. The latter consideration served as the guiding principle for the meta-analysis of word production experiments conducted by I&L. They first identified a set of reliably activated regions for each of the four tasks described above (picture naming, word generation, word reading, and pseudoword reading). In the next step, they identified sets of regions that were possibly related to one or more processing components of word production by analyzing which reliable areas were shared by tasks that shared certain processing components. Picture naming and word generation differ in their lead-in processes but share the whole cascade of word production components from lemma

retrieval onward. The set of regions that were reliably reported for both tasks consisted of the left posterior inferior frontal gyrus (IFG), the left precentral gyrus, the supplementary motor area (SMA), the left mid and posterior parts of the STG and middle (MTG) temporal gyri, the right mid STG, the left fusiform gyrus, the left anterior insula, the left thalamus, and the cerebellum (see I&L: Figure 4). According to I&L, these regions can be assumed to support the core components of word production. Note that this set of regions does not include all of the widespread areas involved in conceptual processing (posterior inferior parietal lobe, MTG, fusiform and parahippocampal gyri, dorsomedial prefrontal cortex, IFG, ventromedial prefrontal cortex, and posterior cingulate gyrus) identified in a recent comprehensive meta-analysis by Binder et al., 2009, see also Schwartz et al., 2009, for an excellent discussion of the clinical evidence on regions involved in conceptual processing). Tasks like picture naming and word generation probably activate quite different concepts and, hence, quite plausibly only enter a common pathway from the point of concept-based lexical retrieval onward.

To find out which of these regions are in fact necessary for word production (rather than just somehow involved), Indefrey (2007) compared them with brain areas in which transient lesions induced by electrocortical or TMS stimulation reliably interfered with picture naming across seven studies (Ojemann, 1983; Ojemann et al., 1989; Schäffler et al., 1993; Haglund et al., 1994; Malow et al., 1996; Hamberger et al., 2001; Stewart et al., 2001). The result showed that all of the core areas (except possibly the left motor cortex) seemed to be necessary for word production. In addition there seemed to be additional necessary areas in the inferior parietal cortex that were only rarely reported in hemodynamic studies.

The strategy of across task comparisons can be taken even further to identify neural correlates of single processing components. I&L exploited the fact that reading tasks only recruit subsets of the core processes of word production. Hence core regions that are also reliably found for reading tasks should be related to the subset of shared processing components rather than to the components that are not shared. Based on such comparisons, they suggested candidate areas for different processing stages from lemma retrieval to phonetic encoding and articulation. I will in the following present their tentative assignments of regions to these processing stages and discuss whether they are compatible with more recent evidence from studies that were designed to target specific processing components. **Table 3** lists data from eight studies providing both spatial and temporal evidence about brain activation during picture naming and provides a median estimate for "peak" activations based on reported peak latencies and the centers of reported time intervals for those regions that have been reported by at least two studies. Regions and median latencies are also shown in **Figure 1**.

CONCEPTUALLY DRIVEN LEXICAL (LEMMA) SELECTION

I&L observed in their meta-analysis that in contrast to all other brain regions found for picture naming and word generation the left MTG was not reliably reported in reading studies (see also Turkeltaub et al., 2002) and reasoned that this region might be related to a processing component also lacking in reading: the retrieval of lexical entries (lemmas) based on concepts speakers want to express. I&L's conjecture was, thus, based on negative evidence and needed further support.

It is, therefore, fortunate, that one of the first studies using a hypothesis-testing approach in an fMRI experiment on word production (de Zubicaray et al., 2001) investigated the neural correlates of lemma selection. They used a semantic picture-word interference paradigm. In this paradigm, competition at the lemma level is induced by presenting semantically related distractors during picture naming, for example the word "pear" when the picture shows an apple. de Zubicaray et al. (2001) found (among other regions) stronger left mid MTG activation for semantic distractors compared to neutral distractors (rows of "X"s), confirming the predicted role of this area. Interestingly, de Zubicaray et al. (2001) also found stronger left posterior STG activation, which they interpreted as evidence for additional competition at the word form level. Such a finding would constitute a serious challenge for a sequential model such as Levelt et al. (1999) which assumes word form activation to take place after lemma selection, i.e., after the competition at the lemma level has been resolved. Note, however, that the control condition of this fMRI study did not involve distractor words but a non-lexical distractor, so that the additional posterior superior temporal activation might have reflected word reading rather than semantic competition. In a more recent study, de Zubicaray et al., 2006, see, however, Abel et al., 2009, for a negative finding in left mid MTG in a similar contrast) ruled out this possible concern by showing increased mid STG and posterior STG activation for the naming of target words that were preceded by the production of semantically related nouns compared to target words that were preceded by unrelated nouns. In this study, they, furthermore, showed additional activation of the anterior cingulate (see also Hirschfeld et al., 2008), and inferior prefrontal regions suggesting the involvement of top-down control processes in the naming of target words preceded by distractors. The latter observation suggests the possibility that the observed STG activation might reflect a top-down influence on self-monitoring activity rather than phonological competitor activation.

Lexical selection in word production was also targeted in an MEG study on picture naming (Maess et al., 2002) by use of the semantic category interference paradigm. In this paradigm, the naming of objects in blocks comprising other objects of the same semantic category is slowed down compared to the naming of objects in semantically heterogeneous blocks. One account for this effect assumes enhanced competition from conceptually similar preceding items (Damian et al., 2001) and, hence, predicts stronger neural activation in a region subserving lemma selection. For subjects showing the behavioral effect, Maess et al. (2002) found significant activation differences between the same-category and the different-category conditions in the left mid MTG and posterior STG in an early (150–225 ms post-stimulus) and a late time window (450–475 ms post-stimulus). Since the available chronometric data on picture naming suggest a time window between 175 and 250 ms for lemma selection (see above), these data are compatible with a role of the left mid MTG in this process. Using a similar paradigm in an arterial spin labeling fMRI study with overt naming, Hocking et al. (2009) observed hippocampal and left mid to posterior STG activation but no mid MTG activation in the semantic blocking condition (see

Table 3 | Overview of studies providing spatial and temporal evidence about brain activation in picture naming.

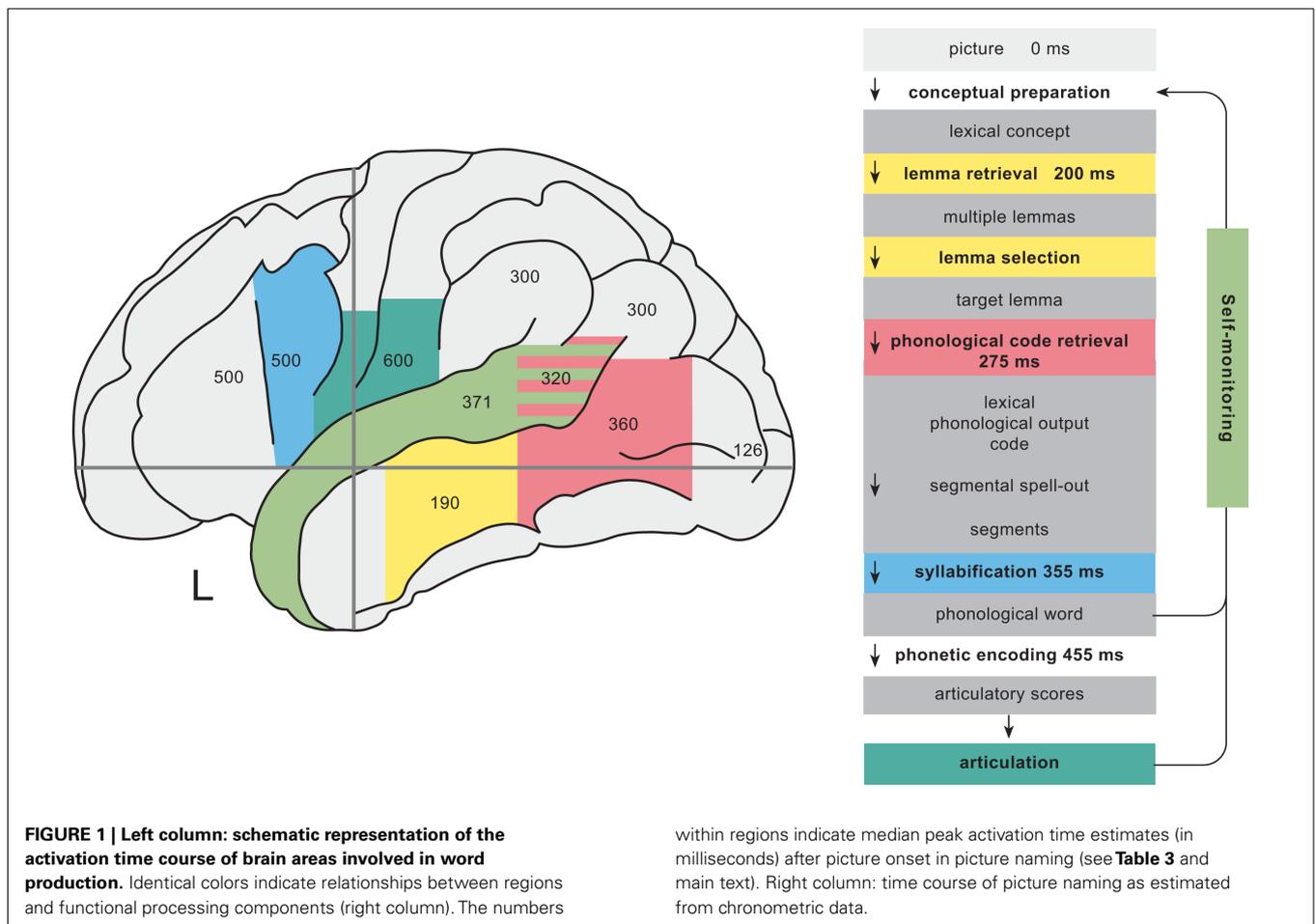| Study Year Method				Salmelin 1994 MEG	Levelt 1998 MEG	Maess 2002 MEG PCA	Sörös 2003 MEG	Vihla 2006 MEG peak	Hultén 2009 MEG peak	Schuhmann 2009 TMS	Acheson In press TMS	Median
Frontal	R	Posterior	GFi	400–600				400	730			500
		Motor	VentGPrc	400–600			400–800	400	730			550
			SMA	400–600								
	L	Posterior	GFi	400–600			200–800	400	600	300–350		500
		Motor	VentGPrc	400–600			400–800	400	600			600
			SMA	400–600								
Temporal	R	Mid	GTs	300–600								
		Posterior	GTs				200–400					
			GTm	200–400								
	L	Anterior	GTs				400–800					
		Mid	GTs	200–400	275–400		400–800	371				371
			GTm			150–225		371			0–200	190
		Posterior	GTs		275–400		200–400		420			320
			GTm	200–400					420			360
			GTi						420			
Parietal	R	Sensory	VentGPoc			400–600						
		Posterior	Lpi				200–400	280	300			300
			Gsm		150–275		200–400	280	300			280
			Ga	200–400			200–400					300
	L	Sensory	VentGPoc			400–600						
		Posterior	LPi				200–400	280	300			300
			Gsm				200–400	280	300			300
			Ga	200–400			200–400					300
Occipital	R			0–200	0–275		0–400	117	100			117
	L				0–275		0–800	117	100			126

For MEG studies, locations and time windows or peak activation times of MEG sources or a spatiotemporal principal component (Maess et al., 2002) are given, for TMS studies the stimulation area and the effective stimulation interval. The Median is calculated from the centers of time windows and the peak activation times. The abbreviations of gyri follow Talairach and Tournoux (1988) except for SMA, supplementary motor area. Ga, angular gyrus; GFi, inferior frontal gyrus; GPoC, postcentral gyrus; GPrC, precentral gyrus; Gsm, supramarginal gyrus; GTs, GTm, superior and middle temporal gyrus; LPi, inferior parietal lobule.

also Heim et al., 2009b). These authors, therefore, doubt that the slowing down of naming responses in the semantic blocking paradigm is due to lexical competition and attribute the posterior STG activation found in both studies to increased demands on self-monitoring.

Whereas attempts to modulate mid MTG activation by inducing competition have yielded only weak support for an involvement of this region in lemma retrieval, more convincing evidence comes from a recent clinical study using a voxel-based lesion-symptom mapping approach. Analyzing data from 64 aphasic patients, Schwartz et al. (2009) found a significant association between semantic errors in language production and lesions in three cortical areas: the left prefrontal lobe, the left posterior MTG, and the left anterior to mid MTG. Importantly, after factoring out non-verbal conceptual deficits this result only held up for the left anterior to mid MTG lesions, allowing the authors to conclude that the left anterior temporal lobe plays a "specific and necessary role" for mapping concepts to words. This study, thus, provides much stronger support than previous studies for an involvement of anterior to mid temporal regions in lemma selection, which had been called into question by some authors (e.g., Wilson et al., 2009).

Two further recent studies have increased the available evidence on the activation time course of the left mid MTG in picture naming (**Table 3**). Vihla et al. (2006) report activation of the left temporal cortex including the mid MTG starting around 250 ms and peaking around 370 ms, that is, later than Maess et al., 2002; 150–225 ms). Acheson et al. (2011) found a significant facilitating TMS effect on response latencies when the left mid MTG but not when the left posterior STG was stimulated between 100 ms before and 200 ms after picture onset. The median peak estimate of the three studies (190 ms) is slightly too early for the revised estimate of 200–275 ms for lemma retrieval. Note, however that the range across the three studies is considerable.

FIGURE 1 | Left column: schematic representation of the activation time course of brain areas involved in word production. Identical colors indicate relationships between regions and functional processing components (right column). The numbers within regions indicate median peak activation time estimates (in milliseconds) after picture onset in picture naming (see **Table 3** and main text). Right column: time course of picture naming as estimated from chronometric data.

Sahin et al. (2009) found activation of Broca's area around 200 ms in a recent word production study using intracranial electrodes in neurosurgical patients. Considering that this activation was sensitive to word-frequency and that I&L suggested a time window for lemma selection between 175 and 250 ms, the authors (and also Hagoort and Levelt, 2009) interpreted this result as indicating a role of Broca's area in lexical access in word production. Note, however, that the tasks used by Sahin et al. (2009) involved the presentation of a written target word rather than a picture, so that the activation observed 200 ms after the target word most likely reflected lexical access in word reading (i.e., from a graphemic code) rather than the concept-based lemma access in word production.

In sum, data on the time course of left mid MTG activation are to date largely compatible with the assumption that this region is involved in conceptually driven lemma retrieval and incompatible with an involvement of this region in a later processing stage, for example phonological retrieval. If one accepts an involvement of this region in lemma retrieval based on the clinical evidence alone (to avoid circularity), the time course data may also be seen as problematic for interactive models assuming feedback from a phonological processing stage to lemma retrieval (cf. Dell et al., 1997). Predictions of modulation due to enhanced competition for lexical selection have

been confirmed in some semantic interference studies, but not convincingly in semantic blocking studies. Insofar as effects of enhanced competition have been found, the data also suggested that competition might affect later processing stages (phonological code retrieval, see next section). These observations are not in accordance with a strictly serial view of the transition from the lemma to the word form level, but this matter is far from settled because an alternative interpretation of these findings as reflecting increased self-monitoring activity is possible.

PHONOLOGICAL CODE (WORD FORM) RETRIEVAL

I&L proposed that left posterior superior temporal lobe might be involved in lexical phonological code (word form) retrieval because this region was reliably found in word production tasks involving the retrieval of lexical word forms but not in pseudoword reading. A more recent study by Binder et al. (2005) suggests that this area and the adjacent angular gyrus can even be deactivated for pseudoword reading compared to a fixation condition.

To date, four MEG studies (Salmelin et al., 1994; Levelt et al., 1998; Sörös et al., 2003; Hultén et al., 2009) provide timing data on the activation of posterior STG and MTG. With a median peak activation of 320–360 ms (see **Table 3**), these data are in good

accordance with an onset of word form retrieval in picture naming around 275 ms (see **Table 2**). Using indwelling electrodes in four patients with epilepsy, Edwards et al. (2010) measured cortical activation in the high gamma range (>70–160 Hz) relative to stimulus onset as well as relative to response onset during picture naming. They report activation at a posterior superior temporal electrode site bordering the parietal lobe and activation of posterior MTG starting around 300 ms after picture onset and continuing until well after response onset. Other mid and posterior STG sites showed activation only after but not before overt responses. The authors interpret the latter finding as supporting a role of the left STG in monitoring (see below) but not in lexical phonological code retrieval.

Other studies have targeted phonological code retrieval by manipulating variables such as phonological relatedness and lexical frequency, by investigating the learning of novel word forms, or by investigating word finding difficulties. In an fMRI study using the picture-word interference paradigm, de Zubicaray et al. (2002) targeted lexical word form retrieval by using distractor words that were phonologically related to the picture names. Such distractor words facilitate naming responses compared to phonologically and semantically unrelated distractor words. As predicted, de Zubicaray et al. (2002) found reduced activation in the left STG suggesting that related distractors primed a phonological representation of the target picture names (see, however, Abel et al., 2009, for activation increase in the adjacent supramarginal gyrus in a similar contrast).

Bles and Jansma (2008) manipulated the phonological relatedness of unattended distractor pictures and also found activation decreases in the left posterior STG when the distractor pictures were phonologically related. In this study different tasks were used and the effect was only observed when participants performed an offset decision task requiring the retrieval of the complete lexical phonological code of a depicted object's name.

Graves et al. (2007) manipulated three variables (lexical frequency, object familiarity, and word length) to study effects at lexical phonological, semantic, and articulatory processing stages. They found the left posterior STG to be sensitive to frequency but not the other variables. Wilson et al. (2009) manipulated the same variables and also report a posterior STG region that was activated in picture naming compared to rest and sensitive to frequency. In a subsequent study, Graves et al. (2008) used a pseudoword repetition task and found decreasing hemodynamic responses over six repetitions of pseudowords in the same region of left posterior STG as in Graves et al. (2007). Given that the pseudowords lacked any semantic content, the authors concluded that "this area participates specifically in accessing lexical phonology." This interpretation, of course, presupposes that over time the pseudowords became novel words. Gaskell and Dumay (2003) showed that pseudowords only become fully integrated in the lexicon, i.e., showing competition effects on phonologically similar words) after consolidation during a sleep phase. Davis et al. (2008) linked this behavioral effect to changes in hemodynamic activation. In their fMRI study, acoustically presented novel words showed word-like lexical competition and word-like hemodynamic activation (in mid and posterior STG) after sleep consolidation. This result might explain why other word learning studies involving training over

several days (e.g., Cornelissen et al., 2004; Grönholm et al., 2005) found stronger activation for newly learned words compared to familiar words in frontal or inferior parietal areas but not the posterior STG.

Results with respect to an involvement of the left posterior STG in *failures* of word form retrieval are mixed. Yagishita et al. (2008) asked their subjects to name famous faces during fMRI scanning. Participants experienced fewer tip-of-the-tongue (TOT) states when the first syllable of the name but not when a second or later syllable of the name was given as a phonological cue. The first-syllable condition resulted in stronger hemodynamic activation of two left mid and posterior STG regions suggesting an involvement of these regions in name retrieval.

In patients with temporal lobe epilepsy, Trebuchon-Da Fonseca et al. (2009) found a relationship between TOT states and reduced resting-state metabolism (measured with 18-fluoro-2-desoxy-D-glucose-PET) in the left inferior parietal lobe and the posterior superior and inferior temporal cortex. By contrast Shafto et al. (2007) found age-related word finding problems to be correlated with gray matter atrophy in the left insula but not the posterior temporal lobe.

Further recent clinical evidence for a role of the left posterior temporal lobe in phonological code retrieval comes from studies on primary progressive aphasia (PPA). Gorno-Tempini and colleagues (Gorno-Tempini et al., 2004, 2008; Henry and Gorno-Tempini, 2010; Wilson et al., 2010) described a so-called logopenic variant of PPA in which mainly gray matter in the mid to posterior temporal lobe is affected[1]. This variant is characterized by anomia and phonemic paraphasias in confrontation naming, an impairment of verbal short-term memory functions, and an absence of the phonological similarity effect on letter recall, whereas conceptual knowledge seems to be relatively unaffected (Gorno-Tempini et al., 2008).

In sum, more recent spatiotemporal data have largely confirmed that during picture naming the left posterior STG/MTG is activated in the predicted time window starting at 275 ms after picture onset. The high resolution data from Edwards et al. (2010) show, however, that within this larger area even spatially close neuronal populations may show differential activation time courses.

Studies manipulating variables affecting lexical word form retrieval have consistently found the predicted effects in the left posterior STG. It should be noted, however, that some studies also found (as yet inconsistent) effects in other brain regions such as the left IFG and the right anterior temporal cortex. Given that also the lexical integration of newly learned words seems to affect their activation of the left posterior STG, it can be concluded that this region's involvement in lexical word form storage and retrieval still has excellent empirical support.

A role of the bilateral posterior superior temporal lobes in the storage of phonological word forms accessed in speech comprehension has been proposed by Hickok and Poeppel (2000, 2004, 2007) on the basis of aphasic comprehension deficits. Wernicke's

---

[1]I would like to thank Stephen Wilson for bringing the logopenic PPA to my attention.

area may thus serve as a common store of lexical word form representations for word production and perception (see also Hocking and Price, 2009, for a lexical phonological effect on left posterior STG activation in comprehension). Most studies on word production report left, rather than bilateral, posterior temporal effects, suggesting that the production system may be more strongly lateralized than the comprehension system.

PHONOLOGICAL ENCODING

All production tasks involve the cascade of word production processes from phonological encoding (syllabification) onward. Comparisons across tasks, therefore, can no longer provide evidence with respect to possible core areas supporting syllabification. I&L reasoned that a comparison between experiments using overt articulation and experiments using covert responses might yield a distinction between syllabification and later processing stages. Syllabification is conceived of as operating on an abstract segmental representation and should be independent of overt articulation, whereas in the subsequent stages of phonetic encoding and articulation motor representations are built up and executed. These processes might be more recruited in overt responses. Corresponding areas might show stronger blood flow increases and therefore might be more easily detected and reported. The left posterior IFG (Broca's area) was the only remaining core area that was not more often reported in experiments using overt responses (see Murphy et al., 1997; Wise et al., 1999; Huang et al., 2001; and more recently Ackermann and Riecker, 2004, for the absence of Broca's area activation in direct comparisons of overt and covert responses). The somewhat indirect conclusion that Broca's area is the most likely candidate area for syllabification has more recently been challenged for different reasons. Firstly, there are good reasons to assume that Broca's area is involved in semantic processing (e.g., Binder et al., 2009). Its activation in word production could, therefore, be due to conceptual preparation rather than post-lexical phonological encoding. Secondly, the influential dual-stream model of speech processing (Hickok and Poeppel, 2000, 2004, 2007) assumes an area at the boundary between temporal and parietal lobes (area Spt) to function as a sensorimotor interface in language production. Hickok and Poeppel (2007) describe the function of area Spt as a "translation between . . . sensory codes and the motor system" and assume sensory codes to represent sequences of segments or syllables. If their view is correct, then the output of such a translation would be motor rather than phonological representations, and motor rather than phonological representations would be relayed forward to Broca's area. Consequently Broca's area would have no role in phonological encoding in word production.

Crucial evidence with respect to the first issue (conceptual processing in word production) comes from studies providing information on the time course of inferior frontal activation in picture naming. The updated chronometric data (see above) suggest a time window between 0 and 200 ms for conceptual preparation and a time window between 355 and 455 ms for syllabification. Including more recent MEG and TMS studies there are now five studies providing temporal data on left IFG activation in picture naming (Salmelin et al., 1994; Sörös et al., 2003; Vihla et al., 2006; Hultén et al., 2009; Schuhmann et al., 2009) summarized in

Table 3. None of the studies found IFG activation before 200 ms. Apart from one subject in Sörös et al. (2003), all MEG studies agree that IFG activation starts after 400 ms. Schuhmann et al. (2009) report increased naming latencies when stimulating Broca's area between 300 and 350 ms after picture onset but not before or after. They used relatively short picture names with an average naming latency of 470 ms so that an effect on phonological encoding before the predicted time window of 355–455 ms is not surprising.

These time course data suggest that whatever the role of Broca's area in conceptual processing may be, it does not seem to be relevant for the preparation of the concept that is used for accessing the lemma level in picture naming, because in this task Broca's area becomes activated too late. In fact, the median peak activation of 500 ms calculated in **Table 3** even raises the question whether Broca's area is activated in time for phonological encoding. On the one hand, the TMS results of Schuhmann et al. (2009) suggest that this depends on the naming latencies of the picture names involved and the MEG studies might have used pictures with relatively long typical naming latencies (Sörös et al., 2003, for example report an average naming latency of 1100 ms). A recent study by Papoutsi et al. (2009), on the other hand, suggests that Broca's area is not only involved in phonological but also subsequent processing stages so that at least the upper boundaries of the observed time intervals may have reflected a later processing stage.

Papoutsi et al. (2009) used a pseudoword repetition task and reasoned that syllabification should be sensitive to the amount of material to be inserted into syllables, i.e., pseudoword length, but not to the frequency of co-occurrence of phonemes in the language (biphone frequency), whereas both variables should affect phonetic encoding and articulation stages. Their hemodynamic activation results showed a dissociation between one more dorsal region in left IFG that was only sensitive to word length but not biphone frequency and, hence, compatible with a role in syllabification and another more ventral region that was sensitive to both variables and thus probably involved in a phonetic processing stage. Results by Ghosh et al. (2008) confirm a stronger activation of the pars opercularis of the left IFG for the production of bisyllables compared to monosyllables. Sahin et al. (2009) report a word length effect in the pars triangularis of the left IFG.

In sum, an involvement of the left IFG in phonological encoding is still compatible with the available spatiotemporal activation data on this region. These data seem to rule out an involvement in an earlier processing stage but certainly not in a later processing stage. Likewise, effects of manipulations of (pseudo)word length are compatible with phonological encoding and later processing stages. To date only one study (Papoutsi et al., 2009) used an experimental variable (biphone frequency) that convincingly disentangles syllabification from later phonetic and motor processing stages and confirmed a pattern predicted for syllabification in a subregion of Broca's area. Following Papoutsi et al.'s (2009) reasoning that phonetic and motor representations should be sensitive to biphone frequency, their result can also be seen as speaking against a purely motor function of Broca's area as assumed by the dual-stream model. However, clear evidence for the assembly of phonological syllable representations in Broca's area is missing.

PHONETIC ENCODING AND ARTICULATION

Of the remaining core areas, the left precentral gyrus, the left thalamus, and the cerebellum are much more frequently found in overt-response paradigms and are most likely involved in articulation. Peeva et al. (2010) found fMRI adaptation for repeated syllables in the left ventral premotor cortex, suggesting syllable-level representations in this region. The exact functional roles of the SMA and the left anterior insula in phonetic encoding or articulation are not so clear. In I&L both areas are reliably found in covert-response studies and only moderately more often in overt articulation studies. With respect to the insula, this pattern of reports is better compatible with a role in articulatory planning as suggested by Dronkers (1996) than a role in articulatory execution. Carreiras et al. (2006) also favor a role in articulatory planning based on their finding of a syllable frequency effect in the left anterior insula. By contrast, Ackermann and Riecker (2004) and Riecker et al. (2000) directly compared overt and covert responses and found insular activation only for overt responses. In another study, insular activation increased linearly with syllable repetition rate (Riecker et al., 2005). These authors suggest an articulatory coordination function for the insula. Murphy et al. (1997), by contrast, did not find articulation-related responses in the insula. Shuster and Lemieux (2005) compared the overt production of multisyllabic words to the production of monosyllabic words. Both suggested functions, articulatory planning and coordination, would predict stronger responses for multisyllabic words, but Shuster and Lemieux did not find any activation difference in the left insula. Clearly, such contradictory findings point to the need for further research in order to identify the experimental conditions under which insular activation is or is not observed.

SELF-MONITORING

Self-monitoring involves an internal loop and an external loop. The internal loop takes as input the phonological score (the phonological word in the case of single words), i.e., the output of phonological encoding. The external loop takes as input the acoustic speech signal of the speaker's own voice (see **Figure 1**). I&L concluded an involvement of the bilateral STG in the external loop of self-monitoring based on studies showing additional bilateral superior temporal activations by distorting the subjects' feedback of their own voice or presenting the subjects with alien feedback while they spoke (McGuire et al., 1996; Hirano et al., 1997). An involvement of the bilateral STG in the internal loop of self-monitoring is suggested by data from Shergill et al. (2002) who manipulated the rate of inner speech.

More recently, Tourville et al. (2008) used feedback with a shifted first formant frequency and applied SEM to the resulting fMRI data. Their results suggest an influence of the auditory cortex on right frontal areas, which according to the authors might be involved in motor correction. Christoffels et al., 2007b, see also van de Ven et al., 2009) studied self-monitoring using verbal feedback without distortion. Their data suggest a much larger network of areas involved in self-monitoring including the cingulate cortex, the bilateral insula, the SMA, bilateral motor areas, the cerebellum, the thalamus and the basal ganglia. The SMA and/or the anterior cingulum also seem to be involved in internal speech monitoring (Möller et al., 2007).

CONCLUSION

Recent neurocognitive research has considerably increased the available evidence on the time course of component processes of word production and on the time course of activation in specific brain regions during picture naming. Furthermore this research field has moved beyond a mere mapping approach and provided highly informative data on the effects of experimental manipulations targeting specific component processes of word production. This article has attempted to evaluate and update the proposals of Indefrey and Levelt (2004) with respect to the time course of word production and with respect to the involvement of brain regions in component processes in the light of more recent evidence. The time course assumptions have largely been confirmed, requiring only some minor adaptations. For the proposed brain structure/function relationships there are varying degrees of supporting and problematic evidence, but as yet no downright falsifications. Adaptations of the original assumptions include the involvement of a more restricted dorsal area within the left IFG in syllabification, the involvement of other parts of the left inferior frontal gyrus in phonetic encoding and/or articulatory planning as well as a probable, yet to date unclear role of the inferior parietal cortex in word production.

REFERENCES

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., and Sommer, W. (2008). Seeing what we know and understand: how knowledge shapes perception. *Psychon. Bull. Rev.* 15, 1055–1063.

Abel, S., Dressel, K., Bitzer, R., Kummerer, D., Mader, I., Weiller, C., and Huber, W. (2009). The separation of processing stages in a lexical interference fMRI-paradigm. *Neuroimage* 44, 1113–1124.

Acheson, D. J., Hamidi, M., Binder, J. R., and Postle, B. R. (2011). A common neural substrate for language production and verbal working memory. *J. Cogn. Neurosci.* 23, 1358–1367.

Ackermann, H., and Riecker, A. (2004). The contribution of the insula to motor aspects of speech production: a review and a hypothesis. *Brain Lang.* 89, 320–328.

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

Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–2796.

Binder, J. R., Medler, D. A., Desai, R., Conant, L. L., and Liebenthal, E. (2005). Some neurophysiological constraints on models of word naming. *Neuroimage* 27, 677–693.

Bles, M., and Jansma, B. M. (2008). Phonological processing of ignored distractor pictures, an fMRI investigation. *BMC Neurosci.* 9, 20. doi:10.1186/1471-2202-9-20

Butterworth, B. (1989). "Lexical access in speech production," in *Lexical Representation and Process*, ed. W. Marslen-Wilson (Cambridge, MA: MIT Press), 108–135.

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.

Carreiras, M., Mechelli, A., and Price, C. J. (2006). Effect of word and syllable frequency on activation during lexical decision and reading aloud. *Hum. Brain Mapp.* 27, 963–972.

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. O. (2007a). Bilingual language control: an event-related brain potential study. *Brain Res.* 1147, 192–208.

Christoffels, I. K., Formisano, E., and Schiller, N. O. (2007b). Neural correlates of verbal feedback processing: an fMRI study employing overt speech. *Hum. Brain Mapp.* 28, 868–879.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., and Ziegler, J. C. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108, 204–256.

Cornelissen, K., Laine, M., Renvall, K., Saarinen, T., Martin, N., and Salmelin, R. (2004). Learning new names for new objects: cortical effects as measured by magnetoencephalography. *Brain Lang.* 89, 617–622.

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.

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

Davis, M. H., Di Betta, A. M., Macdonald, M. J. E., and Gaskell, M. G. (2008). Learning and consolidation of novel spoken words. *J. Cogn. Neurosci.* 21, 803–820.

de Zubicaray, G., McMahon, K., Eastburn, M., and Pringle, A. (2006). Top-down influences on lexical selection during spoken word production: a 4T fMRI investigation of refractory effects in picture naming. *Hum. Brain Mapp.* 27, 864–873.

de Zubicaray, G. I., McMahon, K. L., Eastburn, M. M., and Wilson, S. J. (2002). Orthographic/phonological facilitation of naming responses in the picture-word task: an event-related fMRI study using overt vocal responding. *Neuroimage* 16, 1084–1093.

de Zubicaray, G. I., Wilson, S. J., McMahon, K. L., and Muthiah, S. (2001). The semantic interference effect in the picture-word paradigm: an event-related fMRI study

employing overt responses. *Hum. Brain Mapp.* 14, 218–227.

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

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.

Dronkers, M. F. (1996). A new brain region for coordinating speech articulation. *Nature* 384, 159–161.

Edwards, E., Nagarajan, S. S., Dalal, S. S., Canolty, R. T., Kirsch, H. E., Barbaro, N. M., and Knight, R. T. (2010). Spatiotemporal imaging of cortical activation during verb generation and picture naming. *Neuroimage* 50, 291–301.

Eickhoff, S. B., Heim, S., Zilles, K., and Amunts, K. (2009). A systems perspective on the effective connectivity of overt speech production. *Philos. Transact. A Math. Phys. Eng. Sci.* 367, 2399–2421.

Garrett, M. F. (1980). "Levels of processing in sentence production," in *Speech and Talk (I). Language Production*, ed. B. Butterworth (New York: Academic Press), 177–220.

Gaskell, M. G., and Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition* 89, 105–132.

Ghosh, S. S., Tourville, J. A., and Guenther, F. H. (2008). A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. *J. Speech Lang. Hear. Res.* 51, 1183–1202.

Gorno-Tempini, M. L., Brambati, S. M., Ginex, V., Ogar, J., Dronkers, N. F., Marcone, A., Perani, D., Garibotto, V., Cappa, S. F., and Miller, B. L. (2008). The logopenic/phonological variant of primary progressive aphasia. *Neurology* 71, 1227–1234.

Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., Johnson, J. K., Weiner, M. W., and Miller, B. L. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Ann. Neurol.* 55, 335–346.

Grabowski, T. J., Bauer, M. D., Foreman, D., Mehta, S., Eaton, B. L., Graves, W. W., Defoe, D. L., and Bolinger, L. (2006). Adaptive pacing of visual stimulation for fMRI studies involving overt speech. *Neuroimage* 29, 1023–1030.

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.

Graves, W. W., Grabowski, T. J., Mehta, S., and Gupta, P. (2008). The left posterior superior temporal gyrus participates specifically in accessing lexical phonology. *J. Cogn. Neurosci.* 20, 1698–1710.

Grönholm, P., Rinne, J. O., Vorobyev, V., and Laine, M. (2005). Naming of newly learned objects: a PET activation study. *Brain Res. Cogn. Brain Res.* 25, 359–371.

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.

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

Haglund, M. M., Berger, M. S., Shamseldin, M., Lettich, E., and Ojemann, G. A. (1994). Cortical localization of temporal-lobe language sites in patients with gliomas. *Neurosurgery* 34, 567–576.

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

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

Hamberger, M. J., Goodman, R. R., Perrine, K., and Tamny, T. (2001). Anatomic dissociation of auditory and visual naming in the lateral temporal cortex. *Neurology* 56, 56–61.

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

Heim, S., Eickhoff, S. B., and Amunts, K. (2009a). Different roles of cytoarchitectonic BA 44 and BA 45 in phonological and semantic verbal fluency as revealed by dynamic causal modelling. *Neuroimage* 48, 616–624.

Heim, S., Eickhoff, S. B., Friederici, A. D., and Amunts, K. (2009b). Left cytoarchitectonic area 44 supports selection in the mental lexicon during language production. *Brain Struct. Funct.* 213, 441–456.

Henry, M. L., and Gorno-Tempini, M. L. (2010). The logopenic variant of primary progressive aphasia. *Curr. Opin. Neurol.* 23, 633–637.

Hickok, G., and Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends Cogn. Sci. (Regul. Ed.)* 4, 131–138.

Hickok, G., and Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99.

Hickok, G., and Poeppel, D. (2007). Opinion – the cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.

Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., Ishizu, K., Yonekura, Y., Nagahama, Y., Fukuyama, H., and Konishi, J. (1997). Cortical processing mechanism for vocalization with auditory verbal feedback. *Neuroreport* 8, 2379–2382.

Hirschfeld, G., Jansma, B., Bölte, J., and Zwitserlood, P. (2008). Interference and facilitation in overt speech production investigated with event-related potentials. *Neuroreport* 19, 1227–1230.

Hocking, J., McMahon, K. L., and de Zubicaray, G. I. (2009). Semantic context and visual feature effects in object naming: an fMRI study using arterial spin labeling. *J. Cogn. Neurosci.* 21, 1571–1583.

Hocking, J., and Price, C. J. (2009). Dissociating verbal and nonverbal audiovisual object processing. *Brain Lang.* 108, 89–96.

Huang, J., Carr, T. H., and Cao, Y. (2001). Comparing cortical activations for silent and overt speech using event-related fMRI. *Hum. Brain Mapp.* 15, 39–53.

Huettig, F., and Hartsuiker, R. J. (2008). When you name the pizza you look at the coin and the bread: eye movements reveal semantic activation during word production. *Mem. Cognit.* 36, 341–360.

Hultén, A., Vihla, M., Laine, M., and Salmelin, R. (2009). Accessing newly learned names and meanings in the native language. *Hum. Brain Mapp.* 30, 976–989.

Indefrey, P. (2007). "Brain imaging studies of language production," in *Oxford Handbook of Psycholinguistics*, ed. G. Gaskell (Oxford: Oxford University Press), 547–564.

Indefrey, P., and Levelt, W. J. M. (2000). "The neural correlates of language production," in *The New Cognitive Neurosciences*, ed. M. S. Gazzaniga (Cambridge, MA: MIT Press), 845–865.

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

Koester, D., and Schiller, N. O. (2008). Morphological priming in overt language production:

electrophysiological evidence from Dutch. *Neuroimage* 42, 1622–1630.

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

Laganaro, M., Morand, S., Schwitter, V., Zimmermann, C., Camen, C., and Schnider, A. (2009b). Electrophysiological correlates of different anomic patterns in comparison with normal word production. *Cortex* 45, 697–707.

Levelt, W. J. M. (1989). *Speaking: From Intention To Articulation.* Cambridge, MA: 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, 1–38.

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.

Maess, B., Friederici, A. D., Damian, M., Meyer, A. S., and Levelt, W. J. M. (2002). Semantic category interference in overt picture naming: sharpening current density localization by PCA. *J. Cogn. Neurosci.* 14, 455–462.

Malow, B. A., Blaxton, T. A., Sato, S., Bookheimer, S. Y., Kufta, C. V., Figlozzi, C. M., and Theodore, W. H. (1996). Cortical stimulation elicits regional distinctions in auditory and visual naming. *Epilepsia* 37, 245–252.

McGuire, P. K., Silbersweig, D. A., and Frith, C. D. (1996). Functional neuroanatomy of verbal self-monitoring. *Brain* 119, 101–111.

Möller, J., Jansma, B. M., Rodriguez-Fornells, A., and Münte, T. F. (2007). What the brain does before the tongue slips. *Cereb. Cortex* 17, 1173–1178.

Morgan, J. L., Van Elswijk, G., and Meyer, A. S. (2008). Extrafoveal processing of objects in a naming task: evidence from word probe experiments. *Psychon. Bull. Rev.* 15, 561–565.

Murphy, K., Corfield, D. R., Guz, A., Fink, G. R., Wise, R. J., Harrison, J., and Adams, L. (1997). Cerebral areas associated with motor control of speech in humans. *J. Appl. Physiol.* 83, 1438–1447.

Ojemann, G. A. (1983). Brain organization for language from the

perspective of electrical-stimulation mapping. *Behav. Brain Sci.* 6, 189–206.

Ojemann, G. A., Ojemann, J., Lettich, E., and Berger, M. (1989). Cortical language localization in left, dominant hemisphere – an electrical-stimulation mapping investigation in 117 patients. *J. Neurosurg.* 71, 316–326.

Papoutsi, M., de Zwart, J. A., Jansma, J. M., Pickering, M. J., Bednar, J. A., and Horwitz, B. (2009). From phonemes to articulatory codes: an fmri study of the role of Broca's area in speech production. *Cereb. Cortex* 19, 2156–2165.

Peeva, M. G., Guenther, F. H., Tourville, J. A., Nieto-Castanon, A., Anton, J. L., Nazarian, B., and Alario, F. X. (2010). Distinct representations of phonemes, syllables, and suprasyllabic sequences in the speech production network. *Neuroimage* 50, 626–638.

Perfetti, C.A. (1999). "Comprehending written language: a blueprint of the reader," in *The Neurocognition of Language*, eds C. M. Brown and P. Hagoort (Oxford: Oxford University Press), 167–208.

Riecker, A., Ackermann, H., Wildgruber, D., Meyer, J., Dogil, G., Haider, H., and Grodd, W. (2000). Articulatory/phonetic sequencing at the level of the anterior perisylvian cortex: a functional magnetic resonance imaging (fMRI) study. *Brain Lang.* 75, 259–276.

Riecker, A., Mathiak, K., Wildgruber, D., Erb, M., Hertrich, I., Grodd, W., and Ackermann, H. (2005). fMRI reveals two distinct cerebral networks subserving speech motor control. *Neurology* 64, 700–706.

Rodriguez-Fornells, A., Schmitt, B. M., Kutas, M., and Münte, 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.

Sahin, N. T., Pinker, S., Cash, S. 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.

Saur, D., Kreher, B. W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M. S., Umarova, R., Musso, M., Glauche,

V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., and Weiller, C. (2008). Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18035–18040.

Schäffler, L., Lüders, H. O., Dinner, D. S., Lesser, R. P., and Chelune, G. J. (1993). Comprehension deficits elicited by electrical-stimulation of Broca area. *Brain* 116, 695–715.

Schiller, N. O. (2006). Lexical stress encoding in single word production estimated by event-related brain potentials. *Brain Res.* 1112, 201–212.

Schiller, N. O., Bles, M., and Jansma, B. M. (2003). Tracking the time course of phonological encoding in speech production: an event-related brain potential study. *Cogn. Brain Res.* 17, 819–831.

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.

Schuhmann, T., Schiller, N. O., Goebel, R., and Sack, A. T. (2009). The temporal characteristics of functional activation in Broca's area during overt picture naming. *Cortex* 45, 1111–1116.

Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Faseyitan, O., Brecher, A., Dell, G. S., and Coslett, H. B. (2009). Anterior temporal involvement in semantic word retrieval: voxel-based lesion-symptom mapping evidence from aphasia. *Brain* 132, 3411–3427.

Shafto, M. A., Burke, D. M., Stamatakis, E. A., Tam, P. P., and Tyler, L. K. (2007). On the tip-of-the-tongue: neural correlates of increased word-finding failures in normal aging. *J. Cogn. Neurosci.* 19, 2060–2070.

Shergill, S. S., Brammer, M. J., Fukuda, R., Bullmore, E., Amaro, E. Jr., Murray, R. M., and McGuire, P. K. (2002). Modulation of activity in temporal cortex during generation of inner speech. *Hum. Brain Mapp.* 16, 219–227.

Shuster, L. I., and Lemieux, S. K. (2005). An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain Lang.* 93, 20–31.

Sörös, 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.

Stemberger, J. P. (1985). "An interactive activation model of language production," in *Progress in the Psychology of Language*, ed. A. W. Ellis (London: Erlbaum), 143–186.

Stewart, L., Meyer, B. U., Frith, U., and Rothwell, J. (2001). Left posterior BA37 is involved in object recognition: a TMS study. *Neuropsychologia* 39, 1–6.

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.

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

Talairach, J., and Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain.* Stuttgart: Georg Thieme Verlag.

Tourville, J. A., Reilly, K. J., and Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage* 39, 1429–1443.

Trebuchon-Da Fonseca, A., Guedj, E., Alario, F. X., Laguitton, V., Mundler, O., Chauvel, P., and Liegeois-Chauvel, C. (2009). Brain regions underlying word finding difficulties in temporal lobe epilepsy. *Brain* 132, 2772–2784.

Turkeltaub, P. E., Eden, G. F., Jones, K. M., and Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage* 16, 765–780.

van de Ven, V., Esposito, F., and Christoffels, I. K. (2009). Neural network of speech monitoring overlaps with overt speech production and comprehension networks: a sequential spatial and temporal ICA study. *Neuroimage* 47, 1982–1991.

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.

Vihla, M., Laine, M., and Salmelin, R. (2006). Cortical dynamics of visual/semantic vs. phonological analysis in picture confrontation. *Neuroimage* 33, 732–738.

Wheeldon, L. R., and Levelt, W. J. M. (1995). Monitoring the time-course of phonological encoding. *J. Mem. Lang.* 34, 311–334.

Wilson, S. M., Henry, M. L., Besbris, M., Ogar, J. M., Dronkers, N. F., Jarrold, W., Miller, B. L., and Gorno-Tempini, M. L. (2010). Connected speech production in three variants of primary progressive aphasia. *Brain* 133, 2069–2088.

Wilson, S. M., Isenberg, A. L., and Hickok, G. (2009). Neural correlates of word production stages delineated by parametric modulation of psycholinguistic variables. *Hum. Brain Mapp.* 30, 3596–3608.

Wise, R. J. S., Greene, J., Büchel, C., and Scott, S. K. (1999). Brain regions involved in articulation. *Lancet* 353, 1057–1061.

Yagishita, S., Watanabe, T., Asari, T., Ito, H., Kato, M., Ikehira, H., Kanno, I., Suhara, T., and Kikyo, H. (2008). Role of left superior temporal gyrus during name recall process: an event-related fMRI study. *Neuroimage* 41, 1142–1153.

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

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

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.