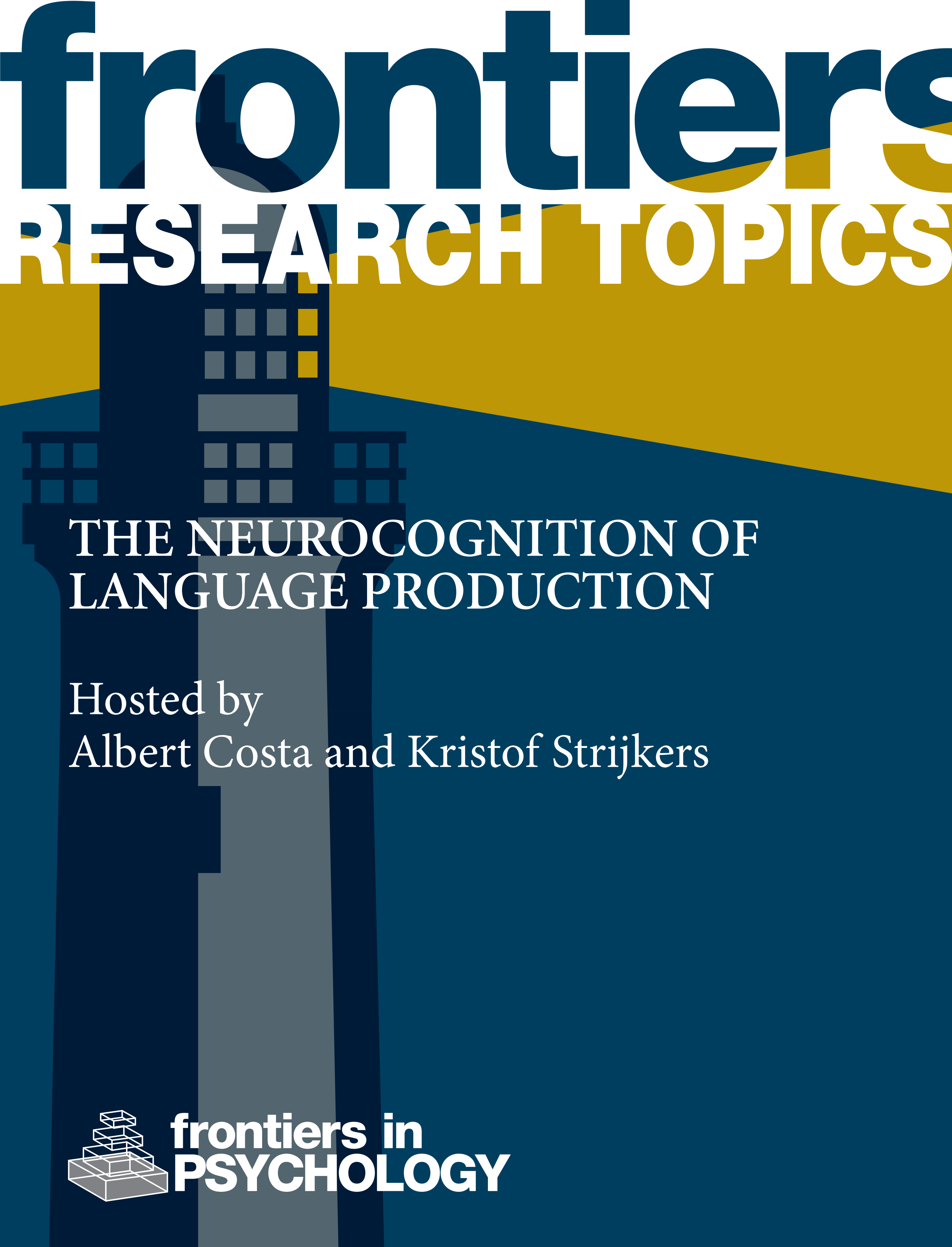
frontiers
RESEARCH TOPICS



THE NEUROCOGNITION OF LANGUAGE PRODUCTION

Hosted by
Albert Costa and Kristof Strijkers

frontiers in
PSYCHOLOGY

FRONTIERS COPYRIGHT STATEMENT

Cover image provided by Ibbl sarl, Lausanne CH

ISSN 1664-8714
ISBN 978-2-88919-033-1
DOI 10.3389/978-2-88919-033-1

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.

All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

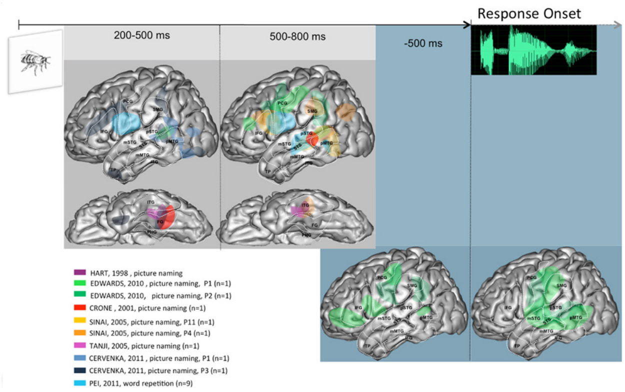
WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

THE NEUROCOGNITION OF LANGUAGE PRODUCTION

Hosted By
Albert Costa, University Pompeu Fabra, Spain
Kristof Strijkers, University of Barcelona, Spain



Over the last decades neuroscience has passed beyond mere "phrenology" and "erpology" and has become an important tool for investigating the spatial, temporal and functional brain dynamics underlying human behavior. In this Special Issue we would like to give a broad overview of recent significant contributions that neuroscientific research has provided to one of the most practiced psychomotor skills unique to humans, namely language production. Physiological studies in language production have not been as extensive compared to many other areas of human cognition and have just currently begun to generate important evidence for uncovering the cognitive mechanisms behind our ability to produce fast and efficient speech. Nonetheless, these findings have already demonstrated their scientific value and interest in neuroscientific approaches for studying language production is increasing exponentially. Therefore, we believe that a topic specially dedicated to neurocognitive advances in language production is not just in its place, but even necessary. Rather than focusing on a specific topic, the idea is to cover many of the important aspects involved in producing speech (semantics, word retrieval, syntax, phonology, motor preparation and control) gathered from various paradigms (e.g., object naming, word naming, etc.) and various populations (monolinguals, bilinguals, patients). The goal is to provide readers with a comprehensive overview of the general questions being addressed in neuroscientific studies on language production, where the research stands, how these findings are of importance for understanding and constraining cognitive models and which future directions have to be taken. To this end we will invite experts in the field who have made significant contributions in the last several years to confer an important topic of language production, critically discuss neuroscientific findings on this topic, relate it to the actual behavior and cognitive models and, importantly, though novel questions which can be derived from their results and facilitate future research in the field. Hereby we hope this Research Topic will be a source of reference both for experts as novices who wish to explore the various mental operations involved in language production from a neurocognitive point of view.

Image credit: From the article "Intra-cranial recordings of brain activity during language production" by Llorens, Trebuchon, Liegeois-Chauval&
Credit: Llorens et al. (2011)

Table of Contents

Review Articles

Original Research Articles

frontiers in
PSYCHOLOGY

EDITORIAL
published: 01 May 2012
doi: 10.3389/fpsyg.2012.00198

The Neurocognition of language production: introduction to the special topic

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

[1] Departament de Psicología Básica, Universitat de Barcelona, Barcelona, Spain
[2] Departament de Tecnología, Institució Catalana de Recerca i Estudis Avançats, Universitat Pompeu Fabra, Barcelona, Spain
*Correspondence: kristof.strijkers@gmail.com

The boom of neuroscientific techniques has opened new ways to study the neural and cognitive processes sustaining human behavior. Combining the traditional behavioral measures with neurophysiological measures does not only provide information about the neurobiological basis of language processing, but also helps to test crucial theoretical hypotheses about the cognitive processes that allow individuals to use language. Hence, it is not surprising that many researchers started to study language processes such as comprehension and visual word recognition with these techniques, leading to an impressive amount of novel observations and significant advances. However, one aspect of language processing has been somewhat neglected during this development, namely the active behavior of speech production. Beyond the several reasons behind this absence of studies exploring the neural basis of language production, namely the theoretical and above all methodological complexity inherent to this psychomotor skill, this state of affairs is changing rapidly.

This is especially so thanks to technical advances and demonstrations that brain activity associated to cognitive processes can be reliably recorded in overt naming tasks with neuroimaging (e.g., Damasio et al., 1996), magnetoencephalographical (e.g., Salmelin et al., 1994; Levelt et al., 1998), and electrophysiological techniques (e.g., Eulitz et al., 2000; Christoffels et al., 2007; Strijkers et al., 2010). The goal of this Special Topic is to provide the reader with a general notion of how these techniques can be used to study the cognition of language production from a plural perspective. The Special Topic comprises both review articles providing current overviews of overt naming studies employing neurophysiological techniques and of methodological aspects of such studies, and original research articles addressing questions of various sub-domains related to speech production and further demonstrating how neurophysiological techniques can be applied to address complex cognitive questions.

Ganushchak et al. (2011) provide a concise review of language production studies employing ERPs. This review is specifically centered on the methodological issues of recording EEG in naming tasks and provides insights to the most relevant components that have been found so far, their possible significance and how well they relate to other ERP deflections observed in the literature. Indefrey (2011) review focuses on both the temporal and spatial correlates of picture naming and links this information to a well-known psychological model, hereby providing a critical update of an influential spatio-temporal meta-analysis on speech production (Indefrey and Levelt, 2004). Llorens et al. (2011) contribute with a critical review on yet another, less frequent but very powerful, technique for studying language production, namely intra-cranial recordings. The authors focus on the advantages and disadvantages of using this technique and, based on the available evidence, they provide a characterization of the neural events occurring in the language network during speech. Finally, Purcell et al. (2011) offer the first quantitative meta-analysis of neuroimaging studies that have examined the neuronal substrates involved in the central and peripheral processes of written word production.

With respect to the original research articles, the Special Topic includes contributions concerning a wide range of speech production theory, illustrating the strong potential of neurophysiological investigations to address relevant cognitive questions for many areas in the field. Wu and Thierry (2011) explored phonological differences in first versus second language production by combining ERPs with a bilingual speech production task, providing novel temporal insights about the progression of language control in bilingual speech production. Also relying on an overt naming paradigm and ERPs, Strijkers et al. (2011) demonstrate how this research strategy can shed light on the role of higher-order intentional and goal-directed processing in accessing the lower-level lexical network during object naming. Price et al. (2011) investigated the interaction between speech production and comprehension through fMRI with the goal of identifying brain activation related to the internal model of speech production after vocalization. Finally, two fMRI studies concerning sentence production form part of the Special Topic. The first one (Shapiro et al., 2012) explored the nature of morphological inflections in sentence production and, in particular, whether neuronal specificity for grammatical operations could be identified. The second one (Tremblay and Small, 2011) examined the hemodynamic correlates involved in the selection of motor responses during sentence production and addressed the question of whether or not motor response selection is different for the production of sentences compared to isolated words.

We believe that the combination of review articles providing critical overviews of the available techniques and the results obtained from them so far with original research articles employing these techniques to study the cognition of language production perfectly satisfies the goals we set out to achieve with the current Special Topic: (1) Offer current and comprehensive insights of the neurophysiological advances in the field both for novices and experts; (2) Remove any lingering skepticisms toward the use of temporally and spatially sensitive measures to study language production; (3) Illustrate with various techniques and for various areas how spatio-temporal knowledge on language production can be exploited to target cognitive questions from a novel point of view. And although the "neurocognition" of language production is still in its infancy, with many open questions and unexplored territories, this is also what makes this such a vivid and exciting field, which will certainly grow exponentially in the years to come.

REFERENCES

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

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., and Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature* 380, 499–505.

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

Ganushchak, L. Y., Christoffels, I. K., and Schiller, N. O. (2011). The use of electroencephalography in language production research: a review. *Front. Psychol.* 2:208. doi: 10.3389/fpsyg.2011.00208

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

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

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.

Llorens, A., Trébuchon, A., Liégeois-Chauvel, C., and Alario, F. (2011). Intra-cranial recordings of brain activity during language production. *Front. Psychol.* 2:375. doi: 10.3389/fpsyg.2011.00375

Price, C. J., Crinion, J. T., and MacSweeney, M. (2011). A generative model of speech production in Broca's and Wernicke's areas. *Front. Psychol.* 2:237. doi: 10.3389/fpsyg.2011.00237

Purcell, J. J., Turkeltaub, P. E., Eden, G. F., and Rapp, B. (2011). Examining the central and peripheral processes of written word production through meta-analysis. *Front. Psychol.*2:239. doi: 10.3389/fpsyg.2011.00239

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

Shapiro, K. A., Moo, L. R., and Caramazza, A. (2012). Neural specificity for grammatical operations is revealed by content-independent fMR adaptation. *Front. Psychol.* 3:26. doi: 10.3389/fpsyg.2012.00026

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

Strijkers, K., Yum, Y. N., Grainger, J., and Holcomb, P. J. (2011). Early goal-directed top-down influences in the production of speech. *Front. Psychol.* 2:371. doi: 10.3389/fpsyg.2011.00371

Tremblay, P., and Small, S. L. (2011). Motor response selection in overt sentence production: a functional MRI study. *Front. Psychol.* 2:253. doi: 10.3389/fpsyg.2011.00253

Wu, Y. J., and Thierry, G. (2011). Event-related brain potential investigation of preparation for speech production in late bilinguals. *Front. Psychol.* 2:114. doi: 10.3389/fpsyg.2011.00114

Received: 28 March 2012; accepted: 28 March 2012; published online: 01 May 2012.

*Citation: Strijkers K and Costa A (2012) The Neurocognition of language production: introduction to the special topic. Front. Psychology **3**:198. doi: 10.3389/fpsyg.2012.00198*

This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Neuroscience.

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 01 September 2011
doi: 10.3389/fpsyg.2011.00208

The use of electroencephalography in language production research: a review

Lesya Y. Ganushchak [1,2] **, **Ingrid K. Christoffels** [1] and **Niels O. Schiller** [1,3]*

[1] *Psychology, Leiden Institute for Brain and Cognition, Leiden, Netherlands*
[2] *Language and Cognition, Max Planck Institute of Psycholinguistics, Nijmegen, Netherlands*
[3] *The Netherlands Institute for Advanced Study in the Humanities and Social Sciences, Wassenaar, Netherlands*

Edited by:
Kristof Strijkers, University of Barcelona, Spain

Reviewed by:
Evelina Fedorenko, Massachusetts Institute of Technology, USA
Els Severens, Ghent University, Belgium

****Correspondence:***
Lesya Y. Ganushchak, Max Planck Institute for Psycholinguistics, Postbus 310, NL 6500 AH, Nijmegen, Netherlands.
e-mail: lganushchak@gmail.com

Speech production long avoided electrophysiological experiments due to the suspicion that potential artifacts caused by muscle activity of overt speech may lead to a bad signal-to-noise ratio in the measurements. Therefore, researchers have sought to assess speech production by using indirect speech production tasks, such as tacit or implicit naming, delayed naming, or meta-linguistic tasks, such as phoneme-monitoring. Covert speech may, however, involve different processes than overt speech production. Recently, overt speech has been investigated using electroencephalography (EEG). As the number of papers published is rising steadily, this clearly indicates the increasing interest and demand for overt speech research within the field of cognitive neuroscience of language. Our main goal here is to review all currently available results of overt speech production involving EEG measurements, such as picture naming, Stroop naming, and reading aloud. We conclude that overt speech production can be successfully studied using electrophysiological measures, for instance, event-related brain potentials (ERPs). We will discuss possible relevant components in the ERP waveform of speech production and aim to address the issue of how to interpret the results of ERP research using overt speech, and whether the ERP components in language production are comparable to results from other fields.

Keywords: overt speech, ERP, review, speech production

THE USE OF ELECTROENCEPHALOGRAPHY IN LANGUAGE PRODUCTION RESEARCH: A REVIEW

Talking is a daily routine in our lives. However, to date there are only few language production studies, in particular on sentence processing, using event-related potential (ERP) measures. This is due to the fact that, for instance, lip, head, and eye movements accompany overt speech (e.g., Grözinger et al., 1975; Brooker and Donald, 1980; Wohlert, 1993). It was feared that such muscle activation would distort the electroencephalography (EEG) signal and therefore make it impossible to investigate language production using EEG. To avoid this problem, language production research focused on meta-linguistic tasks (e.g., phoneme-monitoring), covert naming, and delayed naming (e.g., Van Turennout et al., 1997; Schmitt et al., 2000, 2001; Abdel Rahman et al., 2003). These tasks are successful in avoiding potential speech movement related artifacts, however, they are not without disadvantages. For instance, in case of covert naming, one cannot be sure whether participants follow task instructions. Moreover, the need of actual production of speech may be important to earlier processing and qualitatively influence the speech production process. For instance, intracranial recordings and an fMRI study, showed a different pattern of brain activity for covert versus overt naming (Christoffels et al., 2007b; Pei et al., 2011). In case of button-presses, it is unlikely that only language processes contribute to the response. For instance, in the case of error processing, it cannot be completely excluded that some of the observed errors were due to action slips (e.g., responded with the wrong hand) and were not verbal errors *per se* (e.g., responding "yes" to a phoneme/n/in *lamp*).

The recent increase in published papers measuring overt speech responses using EEG clearly indicates that there is an interest and a great demand for research in language production combining both overt speech responses and EEG recordings. In this paper, we will give an overview of all presently published studies that used tasks requiring immediate overt responses (e.g., picture naming).

The paper is organized as follows: first, we review studies that focused on stimulus-locked analyses, i.e., locked to the time from stimulus onset until a response was given. Within these studies, a division is made between studies investigating native language production, followed by bilingual language production. Second, we will review studies that investigated response-locked ERPs, i.e., processes occurring shortly before or after an overt response was given.

STIMULUS-LOCKED STUDIES

NATIVE LANGUAGE PRODUCTION

To our knowledge, the first published work that combined overt speech with EEG recordings was conducted by Duncan-Johnson and Kopell (1981) and closely replicated much later by Liotti et al. (2000). In both of these studies, a Stroop task was used, where participants were instructed to overtly name the color a word was printed in while ignoring the word itself. However, these earlier studies are limited by sample size (i.e., 12 and 8, respectively) and by number of analyzed electrodes (e.g., only three midline electrodes). Recently, the interest in combining language production with EEG has been revived. The majority of these recent studies investigate the time course of word selection during language production. Most of what we know about the time course of stages of spoken word production comes from chronometric experiments (e.g., voice-key onset latencies; Levelt et al., 1999) and meta-analytic temporal estimates (Indefrey and Levelt, 2004). The high temporal resolution of EEG can provide more information about time course of the spoken word production when combined with tasks that require overt speech production.

According to the Levelt et al. (1999), production of a spoken word consists of lexical selection, lemma retrieval, morphological and phonological code retrieval, and finally articulation. Most of the recent ERP studies focused on the *lexical access* aspect of word production (Hirschfeld et al., 2008; Costa et al., 2009; Dell'Acqua et al., 2010; Strijkers et al., 2010; Aristei et al., 2011).

In a picture–word interference (PWI) paradigm, Hirschfeld et al. (2008) combined each picture with four different distractors: a non-linguistic distractor (e.g., row of Xs), an unrelated distractor word (e.g., flower – DOG), and two types of semantic distractors: words that reflected surface features of a target (e.g., fur – DOG) and words that belong to the same semantic category as a target (e.g., cat – DOG). At a 120–220-ms post-stimulus time interval, the feature related condition resulted in a more negative deflection of the ERP waveform than the unrelated condition. This effect was interpreted as facilitating early stages of visual object processing. During the same time interval, there was a significant difference between all linguistic distractors and the non-linguistic ones. This effect was explained as a result of general conflict-monitoring processes, which are stronger for words than a row of Xs, since only the words have to be suppressed before naming a target picture. However, the 120–220-ms time window approximately corresponds to the time window of 150–250 ms estimated for lexical selection (Indefrey and Levelt, 2004). Thus, it is possible that the observed difference between linguistic and non-linguistic distracters was driven by lexical access, since that is what distinguishes word distractors from a row of Xs. This explanation is in line with the findings of more recent studies (Costa et al., 2009; Sahin et al., 2009; Dell'Acqua et al., 2010; Strijkers et al., 2010; Aristei et al., 2011).

For instance, Aristei et al. (2011) combined PWI with a blocking paradigm [i.e., naming pictures in a semantic context (e.g., cat, dog, horse) and in an unrelated context (e.g., cat, table, flute)]. Aristei et al. (2011) report similar timing for distractor and blocking effects (200 and 250 ms post-stimulus presentation, respectively), possibly suggesting that both effects have similar underlying mechanisms and occur within the time frame of lexical access (Indefrey and Levelt, 2004). In another recent study, Costa et al. (2009) used a so-called cumulative semantic interference paradigm. In this paradigm, participants were asked to name pictures presented in intermixed semantic categories (e.g., turtle, hammer, tree, crocodile, bus, axe, snake, etc.). The typical finding for this paradigm is that naming latencies of a given picture depend on the ordinal position of the picture and on how many items from the same category preceded the pictures (Howard et al., 2006; Costa et al., 2009). Costa et al. (2009) showed that pictures elicited a typical P1/N1/P2 ERP complex in all conditions. In addition, Costa et al. (2009) demonstrated a modulation of the P2, N2, and P300 components. In the N400 window, there was a significant effect of ordinal position; however, it did not correspond to a cumulative pattern seen in other components. Furthermore, similar to Aristei et al. (2011), they showed that lexical access occurred around 200 ms after the onset of the picture. This finding is in line with their previous picture naming study, in which Strijkers et al. (2010) showed that the P2 was sensitive to the lexical frequency of the items, with low-frequency items eliciting more positive amplitudes than high-frequency items.

Further evidence for the time course of lexical access comes from an anomic patient study. Anomic patients have difficulties in word production that could arise at different levels of word production: semantic, lexical, or phonological. Laganaro et al. (2009) recorded ERPs while anomic patients overtly named a series of pictures. They found that patients with lexical-semantic impairment exhibited ERP abnormalities starting at 110 ms after the picture onset. Interestingly, it has also been shown that during object naming, in-depth semantic knowledge about an object causes variation in EEG response 120 ms after object presentation (Abdel Rahman and Sommer, 2008).

Next to lexical access, the time course of *morphological encoding* in overt language production was investigated (Koester and Schiller, 2008). Koester and Schiller (2008) used a long lag-priming paradigm. Participants were presented with words and pictures, and were instructed to read aloud the words and to name the pictures aloud. The words were compounds that were morphologically related to a picture name (e.g., *jaszak* "coat pocket" – JAS "coat") or form-related monomorphemic words (e.g., *jasmijn* "jasmine" – JAS "coat"). The N400 amplitudes, starting 350 ms after the picture onset, were reduced for morphologically related compounds but not for form-related words. This corresponds to the language comprehension literature, where there is evidence that N400 amplitudes are sensitive to morphological processing (e.g., McKinnon et al., 2003). Further evidence comes from a study using intracranial recordings within Broca's area. Sahin et al. (2009) cued participants to inflect nouns (singular/plural) and verbs (past/present). The signal was modulated by the demand of inflection at 320 ms after the target word onset. The neuronal changes were independent of word class. The timing of this effect is also in accordance with meta-analytic temporal estimates of morphological encoding (Indefrey and Levelt, 2004).

Eulitz et al. (2000) mapped the time course of *phonological encoding* during overt picture naming and forming nominal phrases (e.g., using the name and the color of the picture). Eulitz et al. (2000) compared overt production with passive viewing of the same pictures and found ERP markers of phonological encoding between 275 and 400 ms after picture onset. This effect was more pronounced in middle and posterior temporal regions in the left than the right hemisphere, possibly suggesting the involvement of Wernicke's area during phonological encoding. In a PWI paradigm, an effect of phonological distractors occurred in a similar time frame, at about 300 ms after picture onset (Dell'Acqua et al., 2010). Laganaro et al. (2009) showed that anomic patients who had impaired phonological encoding demonstrated normal electro-cortical activity (i.e., similar to healthy control participants) before 300 ms, but abnormal patterns between 300 and 450 ms. This timing was also corroborated by intracranial recordings that showed sensitivity to phonological processes at about 450 ms after the target word onset (Sahin et al., 2009). This time window corresponds to the estimated time course of the phonological encoding (Indefrey and Levelt, 2004).

The papers discussed above have focused on single word production. However, in our everyday communication more complex utterances are produced. To our knowledge, there is only one published paper that investigated conceptual planning in complex utterances in overt language production (Habets et al., 2008). More specifically, Habets et al. (2008) addressed the so-called linearization problem, i.e., the ordering of the event in a sentence (e.g., "before X did A, Y did B" or "after Y did B, X did A"). Participants saw a sequence of two pictures. Each picture consisted of an object that has a strong association with a particular action (e.g., book and reading). Participants were instructed to describe the sequence of two actions associated with the object in chronological/reverse order. A color cue indicated a to-be produced order. ERPs for the "after" condition were more negative than for the "before" condition. This difference emerged between 180 and 230 ms after the vocalization cue, and had a fronto-central distribution. The timing of this effect corresponds closely with comprehension studies investigating temporal order of events in sentences (Münte et al., 1998) and is associated with the engagement of working memory processes in understanding more non-chronological sentences. From 300 ms onward, a parietal distribution was observed. This effect reflects the conceptualization complexity of "before" sentences (Habets et al., 2008).

BILINGUAL LANGUAGE PRODUCTION

To investigate lexical access during production of words in a second language, researchers focused on cognate words (Christoffels et al., 2007a; Verhoef et al., 2009; Strijkers et al., 2010). Cognates are words that are phonologically similar in different languages (e.g., the German – Dutch pair: Apfel – appel). Cognates are typically named faster than non-cognates (e.g., Costa et al., 2000, 2005; Christoffels et al., 2003, 2006). Christoffels et al. (2007a) found more negative amplitudes for cognates compared to non-cognates at about 300 ms after the picture onset, which corresponds with the phonological encoding of words. Strijkers et al. (2010) found a somewhat earlier effect of cognates starting around 200 ms after picture onset, with cognates having more negative amplitudes than non-cognates. The pattern was remarkably similar during both first and second language naming. Note, however, that Figure 5 of Christoffels et al. (2007a) shows a difference between cognates and non-cognates already at around 170 ms after the picture onset. Verhoef et al. (2009) also manipulated cognate status of picture names, however, they do not report any main effect of cognates. Therefore, it is impossible to say whether and when the effects were present.

Next to cognate effects, Christoffels et al. (2007a) and Verhoef et al. (2009) investigated the role of cognitive control and inhibition during language switching. To investigate this issue, a switching paradigm was used, where participants on a given cue were required to name a picture in their first (L1) or second language (L2). Christoffels et al. (2007a) found that naming in L1 was slower and the ERPs were modulated between 275 and 375 ms (time window of N2) compared to naming pictures in L2. Verhoef et al. (2009) manipulated the time between cue and picture onset (i.e., long versus short stimulus onset intervals). They found that preparation time manipulated the degree to which inhibitory control biased language competition as indexed by the N2. Chauncey et al. (2009) also found modulation of the N2 amplitudes. Participants were instructed to overtly name pictures in their L1 (English) and their L2 (French). Pictures were preceded by a word prime, presented for 70 ms. Primes were either the (English or French) name of the to-be named picture or were unrelated to the picture. The language of the prime word affected ERP at about 200 ms after picture onset, but only when pictures were named in L2 and not in L1. The authors argued that the L1 prime interfered with suppression of the L1 lexical activation, which is needed for L2 but not L1 production, thereby creating a conflict reflected in the N2 amplitudes (Chauncey et al., 2009).

There were also first steps taken to investigate processes involved in translation from one language to another. Christoffels et al. (2009) asked participants to translate interlingual homographs: i.e., words that shared orthographic form but had a different meaning in two languages (e.g., "room" refers to CREAM in Dutch) and control words. Participants had to translate targets from and to their first and second language. The authors showed that the brain starts to distinguish between translation directions as early as 200 ms. The results of the study are in line with the idea that language information in the input, a "language cue," rather than an output lexicon, helps to reduce competition between languages when selecting the proper target response (Kroll et al., 2010).

CONCLUSION

The studies discussed above demonstrate that the combination of EEG recording and language production can be successfully employed. The studies provide converging evidence about the time course of word production on both native and second languages. Specifically, the brain engages in lexical selection around 200 ms after picture onset (e.g., Hirschfeld et al., 2008; Costa et al., 2009; Strijkers et al., 2010; Aristei et al., 2011), phonological encoding between 275 and 400 ms (Eulitz et al., 2000), and morphological processes starting around 350 ms after the picture onset (Koester and Schiller, 2008). The ERP research indicates that this time course is in accordance with the estimated timings reported by Indefrey and Levelt (2004). It also demonstrates that EEG recording may be a very sensitive tool to investigate temporal and qualitative differences between first and second language production. However, most of the paradigms used in speech production research require not only production of an utterance, but also comprehension (e.g., reading distractors) and a domain-general processes (e.g., suppressing distractor activation). Potentially more "pure" production paradigm could be a verbal fluency task, where participants required to name members of a given category within given time. However, even within production tasks it is difficult to manipulate different stages, e.g., lexical, morphological, phonological, and speech planning, independently of each other. Thus, ERPs could reflect multiple components associated with various comprehension, production, and domain-general processes. Future studies are needed to disentangle these various aspects during speech production.

RESPONSE-LOCKED

During speech production, we continuously monitor what we say and what we are about to say. In investigating the working of the speech production monitor, researchers have focused on error monitoring. An electrophysiological measure related to error processing is the error-related negativity (ERN; Falkenstein et al., 1991; Gehring et al., 1993), a component of the ERP that has a fronto-central scalp distribution and peaks about 80 ms after an overt incorrect response (Bernstein et al., 1995; Scheffers et al., 1996; Holroyd and Yeung, 2003). The ERN originates in the anterior cingulate cortex (ACC) and/or the supplementary motor area (SMA; e.g., Dehaene et al., 1994; Debener et al., 2005). Recently, studies demonstrated an ERN after errors in meta-linguistics tasks (e.g., Ganushchak and Schiller, 2006, 2008a, 2009; Sebastián-Gallés et al., 2006) and in tasks that require an overt response (e.g., Masaki et al., 2001; Möller et al., 2007; Ganushchak and Schiller, 2008b; Riés et al., 2011). We will review the later studies below.

Masaki et al. (2001) were the first to investigate whether an ERN occurs following speech errors in the Stroop color–word task. Participants were instructed to overtly name the color of each stimulus. Masaki et al. (2001) found an ERN-like component after speech errors, e.g., when participants named the wrong color. Masaki et al. (2001) used loud pink noise to suppress a so-called vocalization-related cortical potential (VRCP). The VRCP is related to movement related potential preceding vocalization and an auditory-evoked potential that follows vocalization (Gunji et al., 2000). The VRCP has a similar time course as the ERN but is independent from the correctness of the response. However, using a masking procedure might not be ideal to study verbal self-monitoring. Speakers use their output as feedback to monitor their own speech (e.g., Levelt et al., 1999). Removing such feedback might interfere with the normal working of the monitoring process (e.g., Christoffels et al., 2007b; Christoffels et al., 2011).

In a more recent study on verbal self-monitoring, no masking procedure was used. Möller et al. (2007) used a so-called SLIP paradigm to induce errors. In this task, participants have to read inductor word pairs such as "ball doze," "bash door," and "bean deck," which are followed by a target word pair such as "darn bore" (see Motley et al., 1982). The reversal of initial phonemes in the target pair compared to the inductor pairs may lead to onset exchange errors such as "barn door." Möller et al. (2007) asked their participants to covertly read the inductor word pairs and vocalize the target word pair preceding a response cue. They found an enlarged negativity on error trials, preceding, and following the response cue. The first negativity reflects conflict at a phonological/phonetic encoding stage. The second negativity indexes conflict at articulatory motor stage. Interestingly, Severens et al. (2011), found a similar negativity following the response cue in the *absence* of error on taboo-eliciting trials (e.g., katten nut → natte k*t; cats sense → wet c*t) compared to neutral trials. The authors concluded that taboo errors were elicited and corrected internally prior to articulation, and suggested that the negativity reflects resolution of conflict rather than detection of conflict.

Ganushchak and Schiller (2008b) employed a semantic blocking picture naming task to study error monitoring in speech production. In addition to semantic context, participants' motivation was manipulated. In the high-motivation condition, participants were told that they would be financially punished for speech errors. In the low-motivation condition, neither financial punishment nor reward was administered. The authors obtained an ERN on error trials. The amplitude of the ERN was modulated by semantic context, with larger amplitudes for semantic blocks than unrelated blocks, indicating that semantic relatedness resulted in higher conflict between potential verbal responses. Furthermore, the ERN was larger and peaked later in the high-motivation condition compared to low-motivation condition, indicating higher monitoring activity.

Another component that is associated with error processing is the error positivity (Pe), which is thought to reflect a more thorough evaluation of the error response (Falkenstein et al., 1991). The Pe has a centro-parietal distribution and peaks about 300 ms after the overt error. Contrary to the ERN, the Pe is specific to overt and detected errors (for a review see Overbeek et al., 2005). The Pe after the overt vocal responses is inconsistently reported in the literature. For instance, Masaki et al. (2001) report a Pe after the incorrect trials. However, Riés et al. (2011) showed a Pe following errors that required manual response, but not after overt speech errors. It is possible that during overt speech production some of the errors are left undetected and therefore no Pe is elicited (for discussion on this issue see Riés et al., 2011). More research is needed to determine whether the Pe can be reliably observed following overt vocal responses and what the possible underlying mechanisms are.

The studies reviewed above suggest that verbal monitoring might be a special case of general performance monitoring rather than a completely different process. If so, the ERN should also be observed on correct trials. However, in the studies described above, no ERN was reported on correct trials. In contrast, in non-verbal tasks, the ERN was shown at both correct and incorrect trials (e.g., Vidal et al., 2000, 2003; Bartholow et al., 2005). The ERN-like amplitude on correct trials is smaller than on incorrect trials. During overt speech tasks, this negativity could have been masked by motor artifacts and therefore remained undetected on correct trials (Riés et al., 2011). To analyze overt picture naming data, Riés et al. (2011) used a blind source separation algorithm on the basis of canonical correlation analysis (BSS-CCA; De Clercq et al., 2006). This method reliably reduces the EMG artifacts induced by articulation (see De Vos et al., 2010). This analysis method allowed Riés et al. (2011) to reliably observe the ERN on both correct and incorrect trials, supporting the hypothesis that verbal monitoring involved in speech production is part of the general-purpose mechanism. This electrophysiological evidence is supported by imaging studies, showing the ACC and SMA activation during overt naming (e.g., Christoffels et al., 2007b). Interestingly, McArdle et al. (2009) showed that the Bereitschaftspotential (BP), an electrophysiological index of voluntary movement, was modulated by linguistic processes such as lexical access independently from articulation. This suggests that the premotor system plays a role in lexical access and provides further evidence of a functional interaction between cortical motor and language networks (McArdle et al., 2009).

Taken together, the studies reviewed in this section suggest that the ERN obtained in overt speech production task is comparable to the ERN found in action monitoring studies and can be used as an electrophysiological marker in psycholinguistic research. More generally, the reliable investigation of language processes using overt responses in combination with EEG recordings is possible even in response-locked analyses.

METHODOLOGICAL RECOMMENDATIONS AND CONCLUSION

The above-reviewed studies show that artifact-free brain responses can be measured up to at least 400 ms post-stimulus presentation (e.g., Eulitz et al., 2000; Christoffels et al., 2007a; Aristei et al., 2011). In a stimulus-locked analysis, care needs to be taken to exclude trials that are contaminated by the earliest responses. A recent study, however, using Independent Component Analysis, showed that the early ERP components might not necessarily be artifact-free (Porcaro et al., 2010). Thus, the results should be interpreted with caution and potentially different methods should be used, e.g., Independent Component Analyses, to remove movement related artifacts. For the response-locked analysis, researchers interested in the ERN could use the standard procedures also used in the action monitoring studies. However, this is true only for error trials. The ERN on correct trials, is significantly smaller than the one on error trials and is more likely to be masked by motion artifacts (which are larger in overt speech compared to button-presses) and also largely affected by severe filtering (up to 12 Hz), which is commonly done in the ERN analysis on error trials. Researchers interested in the later processes, such as self-monitoring and response evaluation on correct rather than error trials should preferably use different methods of analysis to remove motion-related artifacts (e.g., BSS-CCA, De Vos et al., 2010; Riés et al., 2011).

In terms of design, a simple and important consideration is to make sure that conditions are comparable in terms of overt output. It is known that the morphology of the speech artifacts in the ERPs varies systematically with the phonetic properties of the utterance. Therefore, it is advisable to compare conditions in which identical words are produced (Aristei et al., 2011) or – when this is impossible – care needs to be taken to match the to-be produced words not only on usual measures, such as frequency of occurrence, but also on their phonetic properties.

The ERP studies reviewed here demonstrate that classical ERP components, among others P2, N400, and ERN, can be observed in the paradigms that require an overt speech response. Thus, this review suggests that combining ERP with overt articulation is not only possible but necessary to provide more insights into the language production processes, allowing investigation of the temporal flow and scalp distributions of well-established behavioral effects (e.g., semantic interference) as well as investigation of various stages of word and sentence production.

REFERENCES

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

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

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

Bartholow, B. D., Pearson, M. A., Dickter, C. L., Sher, K. J., Fabiani, M., and Gratton, G. (2005). Strategic control and medial frontal negativity: beyond errors and response conflict. *Psychophysiology* 42, 33–42.

Bernstein, P. S., Scheffers, M. K., and Coles, M. G. H. (1995). "Where did I go wrong?" A psychophysiological analysis of error detection. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 1312–1322.

Brooker, B. H., and Donald, M. W. (1980). Contribution of speech musculature to apparent EEG asymmetries prior to vocalization. *Brain Lang.* 9, 226–245.

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

Christoffels, I. K., de Groot, A. M. B., and Kroll, J. F. (2006). Memory and language skills in simultaneous interpreting: expertise and language proficiency. *J. Mem. Lang.* 54, 324–345.

Christoffels, I. K., de Groot, A. M. B., and Waldorp, L. J. (2003). Basic skills in a complex task: a graphical model relating memory and lexical retrieval to simultaneous interpreting. *Biling. (Camb. Engl.)* 6, 201–211.

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.

Christoffels, I. K., Ganushchak, L. Y., and Koester, D. (2009). "When ROOM means cream and room: translation, homographs, ERPs and over speech," in *Proceedings of the European Society for Cognitive Psychology*, Krakow.

Christoffels, I. K., van de Ven, V., Waldorp, L. J., Formisano, E., and Schiller, N. O. (2011). The sensory consequences of speaking: parametric neural cancellation during speech in auditory cortex. *PLoS One* 6, e18307. doi: 10.1371/journal.pone.0018307 [Public Library of Science One].

Costa, A., Caramazza, A., and Sebastián-Gallés, N. (2000). The cognate facilitation effect: implications for models of lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1283–1296.

Costa, A., Santesteban, M., and Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain Lang.* 94, 94–103.

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.

De Clercq, W., Vergult, A., Vanrumste, B., Van Paesschen, W., and Van Huffel, S. (2006). Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. *IEEE Trans. Biomed. Eng.* 53, 2583–2587.

De Vos, M., Ries, S., Vanderperren, K., Vanrumste, B., Alario, F.-X., Van Huffel, S., and Burle, B. (2010). Removal of muscle artifacts from EEG recordings of spoken language production. *Neuroinformatics* 8, 135–150.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., Von Cramon, Y., and Engel, A. K. (2005). Trial-by-trial coupling of concurrent EEG and fMRI identifies the dynamics of performance monitoring. *J. Neurosci.* 25, 11730–11737.

Dehaene, S., Posner, M. I., and Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychol. Sci.* 5, 3–23.

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

Duncan-Johnson, C. C., and Kopell, B. S. (1981). The Stroop effect: brain potentials localize the source of interference. *Science* 214, 938–940.

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

Falkenstein, M., Hohnsbein, J., Hoorman, J., and Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalogr. Clin. Neurophysiol.* 78, 447–455.

Ganushchak, L. Y., and Schiller, N. O. (2006). Effects of time pressure on verbal self-monitoring. *Brain Res.* 1125, 104–115.

Ganushchak, L. Y., and Schiller, N. O. (2008a). Effects of auditory distractors on verbal self-monitoring. *J. Cogn. Neurosci.* 20, 927–940.

Ganushchak, L. Y., and Schiller, N. O. (2008b). Motivation and semantic context affect brain error-monitoring activity: an event-related brain potentials study. *Neuroimage* 39, 385–405.

Ganushchak, L. Y., and Schiller, N. O. (2009). Speaking one's second language under time pressure: an ERP study on verbal self-monitoring in German-Dutch bilinguals. *Psychophysiology* 46, 410–419.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., and Donchin, E. (1993). A neural system for error detection and compensation. *Psychol. Sci.* 4, 385–390.

Grözinger, B., Kornhuber, H. H., and Kriebel, J. (1975). Methodological problems in the investigation of cerebral potentials preceding speech: determining the onset and suppressing artefacts caused by speech. *Neuropsychologia* 13, 263–270.

Gunji, A., Hoshiyama, M., and Kakigi, R. (2000). Identification of auditory evoked potential of one's own voice. *Clin. Neurophysiol.* 111, 214–219.

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

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

Holroyd, C. B., and Yeung, N. (2003). Alcohol and error processing. *Trends Neurosci.* 26, 402–404.

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

Indefrey, P., 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.

Kroll, J. F., van Hell, J. G., Tokowicz, N., and Green, D. W. (2010). The revised hierarchical model: a critical review and assessment. *Biling. (Camb. Engl.)* 13, 373–381.

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

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

Liotti, M., Woldorff, M. G., Perez, R. III, and Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia* 38, 701–711.

Masaki, H., Tanaka, H., Takasawa, N., and Yamazaki, K. (2001). Error related brain potentials elicited by vocal errors. *Neuroreport* 12, 1851–1855.

McArdle, J. J., Mari, Z., Pursley, R. H., Schulz, G. M., and Braun, A. R. (2009). Electrophysiological evidence of functional integration between the anguage and motor systems in the brain: a study of the speech Bereitschaftspotential. *Clin. Neurophysiol.* 120, 245–284.

McKinnon, R., Allen, M., and Osterhout, L. (2003). Morphological decomposition involving non-productive morphemes: ERP evidence. *Neuroreport* 14, 883–886.

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

Motley, M. T., Camden, C. T., and Baars, B. J. (1982). Covert formulation and editing of anomalies in speech production: evidence from experimentally elicited slips of the tongue. *J. Verbal Learn. Verbal Behav.* 21, 578–594.

Münte, T. F., Schiltz, K., and Kutas, M. (1998). When temporal terms belie conceptual order. *Nature* 395, 71–73.

Overbeek, T. J. M., Nieuwenhuis, S., and Ridderinkhof, K. R. (2005). Dissociable components of error processing: on the functional significance of the Pe vis-à-vis the ERN/Ne. *J. Psychophysiol.* 19, 319–329.

Pei, X., Leuthardt, E. C., Gaona, C. M., Brunner, P., Wolpaw, J. R., and Schalk, G. (2011). Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. *Neuroimage* 54, 2960–2972.

Porcaro, C., Medaglia, M. T., Meyer, A. S., and Krott, A. (2010). "Artifact-free electrophysiological responses during overt speech production," in *Proceedings of Human Brain Mapping*, Barcelona, 106.

Riés, S., Janssen, N., Dufau, S., Alario, F.-X., and Burle, B. (2011). General purpose monitoring during speech production. *J. Cogn. Neurosci.* 23, 1419–1436.

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.

Scheffers, M. K., Coles, M. G. H., Bernstein, P. S., Gehring, W. J., and Donchin, E. (1996). Event-related brain potential and error-related processing: an analysis of incorrect responses to go and no-go stimuli. *Psychophysiology* 33, 42–53.

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.

Sebastián-Gallés, N., Rodríguez-Fornells, A., De Diego-Balaquer, R., and Díaz, B. (2006). First- and second-language phonological representation in the mental lexicon. *J. Cogn. Neurosci.* 18, 1277–1291.

Severens, E., Janssens, I., Kühn, S., Brass, M., and Hartsuiker, R. J. (2011). When the brain tames the tongue: covert editing of inappropriate language. *Psychophysiology* 48, 1252–1257.

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, 913–928.

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.

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

Vidal, F., Burle, B., Bonnet, M., Grapperon, J., and Hasbroucq, T. (2003). Error negativity on correct trials: a reexamination of available data. *Biol. Psychol.* 64, 265–282.

Vidal, F., Hasbroucq, T., Grapperon, J., and Bonnet, M. (2000). Is the "error negativity" specific to errors? *Biol. Psychol.* 51, 109–128.

Wohlert, A. B. (1993). Event-related brain potentials preceding speech and nonspeech oral movements of varying complexity. *J. Speech Hear. Res.* 36, 897–905.

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

Received: 14 March 2011; paper pending published: 09 June 2011; accepted: 13 August 2011; published online: 01 September 2011.

Citation: Ganushchak LY, Christoffels IK and Schiller NO (2011) The use of electroencephalography in language production research: a review. Front. Psychology ***2****:208. doi: 10.3389/fpsyg.2011.00208*

This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 12 October 2011
doi: 10.3389/fpsyg.2011.00255

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

Edited by:
Albert Costa, University Pompeu Fabra, Spain

Reviewed by:
William W. Graves, Medical College of Wisconsin, USA
Stephen Wilson, University of Arizona, USA

***Correspondence:**
Peter Indefrey, Institut für Sprache und Information, Heinrich Heine Universität Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany.
e-mail: indefrey@phil.uni-duesseldorf.de

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 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.

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.*

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				**Salmelin**	**Levelt**	**Maess**	**Sörös**	**Vihla**	**Hultén**	**Schuhmann**	**Acheson**	**Median**
Year				**1994**	**1998**	**2002**	**2003**	**2006**	**2009**	**2009**	**In press**	
Method				**MEG**	**MEG**	**MEG**	**MEG**	**MEG**	**MEG**	**TMS**	**TMS**	
						PCA		**peak**	**peak**			
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.

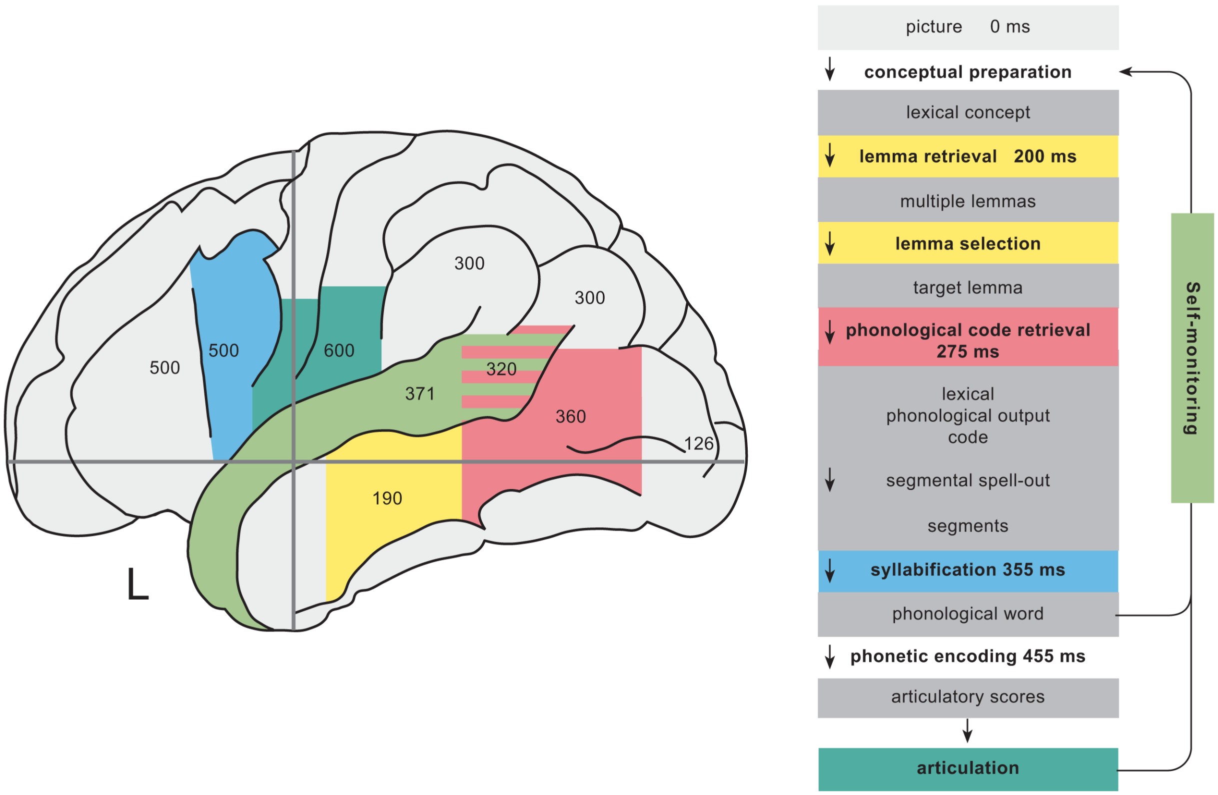


FIGURE 1 | Left column: schematic representation of the activation time course of brain areas involved in word production. Identical colors indicate relationships between regions and functional processing components (right column). The numbers within regions indicate median peak activation time estimates (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., Kümmerer, 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 fMRl 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 supra-syllabic 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.

Received: 15 March 2011; accepted: 13 September 2011; published online: 12 October 2011.

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

This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 27 December 2011
doi: 10.3389/fpsyg.2011.00375

Intra-cranial recordings of brain activity during language production

Anaïs Llorens[1,2], Agnès Trébuchon[1], Catherine Liégeois-Chauvel[1] and F.-Xavier Alario[2]*

[1] *INSERM UMR 751, Aix-Marseille Université, Marseille, France*
[2] *Laboratoire de Psychologie Cognitive, CNRS, Aix-Marseille Université, Marseille, France*

Edited by:
Albert Costa, University Pompeu Fabra, Spain

Reviewed by:
Olaf Hauk, MRC Cognition and Brain Sciences Unit, UK
Lyndsey Nickels, Macquarie University, Australia

Correspondence:
F.-Xavier Alario, Laboratoire de Psychologie Cognitive, CNRS, Aix-Marseille Université, 3 place Victor Hugo (Bâtiment 9, Case D), 13331 Marseille Cedex 3, France.
e-mail: francois-xavier.alario@univ-provence.fr

Recent findings in the neurophysiology of language production have provided a detailed description of the brain network underlying this behavior, as well as some indications about the timing of operations. Despite their invaluable utility, these data generally suffer from limitations either in terms of temporal resolution, or in terms of spatial localization. In addition, studying the neural basis of speech is complicated by the presence of articulation artifacts such as electro-myographic activity that interferes with the neural signal. These difficulties are virtually absent in a powerful albeit much less frequent methodology, namely the recording of intra-cranial brain activity (intra-cranial electroencephalography). Such recordings are only possible under very specific clinical circumstances requiring functional mapping before brain surgery, most notably in patients that suffer from pharmaco-resistant epilepsy. Here we review the research conducted with this methodology in the field of language production, with explicit consideration of its advantages and drawbacks. The available evidence is shown to be diverse, both in terms of the tasks and the cognitive processes tested and in terms of the brain localizations being studied. Still, the review provides valuable information for characterizing the dynamics of the neural events occurring in the language production network. Following modality specific activities (in auditory or visual cortices), there is a convergence of activity in superior temporal sulcus, which is a plausible neural correlate of phonological encoding processes. Later, between 500 and 800 ms, inferior frontal gyrus (around Broca's area) is involved. Peri-rolandic areas are recruited in the two modalities relatively early (200–500 ms window), suggesting a very early involvement of (pre-) motor processes. We discuss how some of these findings may be at odds with conclusions drawn from available meta-analysis of language production studies.

Keywords: electrocorticography, intra-cranial recording, lexical access, phonological encoding, articulation, speech, gamma band activity

INTRODUCTION

Speech is a basic skill that is used quite effortlessly in many daily life activities. Despite this apparent simplicity, speech is subtended by a complex set of cognitive processes and a wide network of brain structures, engaged, and interacting in time within tens of milliseconds. Cognitive models of speech production generally include distinct levels of processing. These allow the retrieval and use of semantic information (i.e., the message to be conveyed), linguistic information (e.g., lexical and phonological representations), as well as pre-motor and motor commands (for articulating).

The network of brain areas engaged during speech production is relatively well described, based on evidence from neuropsychological populations (e.g., speakers suffering from various kinds of aphasia following a stroke) and from functional brain imaging experiments. For example, DeLeon et al. (2007) investigated the linguistic performance and neural integrity of patients within 24 h of acute ischemic stroke. They showed that a deficit in semantic processing (conceptual identification), is associated with dysfunction in anterior temporal brain areas (Brodmann areas BA 21–22–38), while a lexical dysfunction is associated with posterior temporal regions (BA 37–39). Indefrey and Levelt (2004) conducted a meta-analysis of brain activity studies of diverse types, including functional magnetic resonance imaging (fMRI), positron emission tomography (PET), electroencephalography (EEG), magneto-encephalography (MEG), and transcranial magnetic stimulation (TMS). The meta-analysis shows that the selection of a lexical item involves activation of the mid part of left middle temporal gyrus; accessing a word's phonological code is linked to activation in Wernicke's area; and post-phonological encoding (syllabification and metrical encoding) is linked to activation in left inferior frontal regions.

The techniques mentioned above provide complementary insights into neural processing. For instance, fMRI is productive in identifying areas that are central in different language tasks (Price, 2010) but has an indirect, temporally smeared, and poorly understood relationship to neural processing. In contrast, surface EEG and MEG are directly and instantaneously generated by synaptic and active currents in pyramidal apical dendrites. For this reason, they provide valuable information about the timing of neural events that can then be linked to cognitive operations.

For example, the retrieval of lexical linguistic information mentioned above appears to be engaged around 200 ms post-stimulus in the classic picture naming task (Salmelin et al., 1994; Maess et al., 2002; Costa et al., 2009; see also Indefrey and Levelt, 2004; Salmelin, 2007). However, inferring and localizing the cortical sources of these process is a technically complex problem, based on mathematical algorithms constrained by *a priori* hypothesis. The other constraint faced by surface EEG and MEG is that they are both highly sensitive to electro-myographic artifacts. These are electrical signals generated by articulatory muscles well before the onset of speech (Goncharova et al., 2003). They have a large amplitude compared to neural signals and thus interfere with the signal of interest. The presence of massive EMG limits the time windows and frequency bands of activity that can be analyzed fruitfully (although see McMenamin et al., 2009, 2010; De Vos et al., 2010).

Such caveats do not apply, however, when the neurophysiological signal is recorded intra-cranially (intracranial EEG, or iEEG), rather than on the surface of the scalp. iEEG can provide precise spatial resolution and physiological interpretations that are not possible with surface EEG or MEG. The signal is barely contaminated with electro-myographic artifacts because these do not propagate to intra-cranial electrodes. However, these investigations are limited to quite specific circumstances, such as some forms of epilepsy, where intra-cranial electrodes have to be implanted for clinical purposes.

Exceptional as they may be, intra-cranial neurophysiological recordings can contribute to our understanding of the neurocognition of language production, if the appropriate interpretative precautions are taken (see next section on Methodological Considerations). To illustrate this view, our focus here will be on revealing the dynamics of the neural events occurring in the language production network. The inclusion criterion we used when selecting the articles to be reviewed from publication databases was that they reported studies in which brain activity was recorded intra-cranially while participants were engaged in tasks requiring overt or covert language production. Below, we begin with a brief methodological primer, followed by a discussion of the empirical studies.

METHODOLOGICAL CONSIDERATIONS

Patients with pharmacologically resistant epilepsy may be candidates for neurosurgical procedures during which epileptogenic zones are resected. For this procedure to be acceptable, it is of critical importance not only to identify cortical regions that produce seizure onsets but also regions with eloquent functional roles (e.g., motor, language) that *should not* be resected. Intra-cranial measures provide invaluable information for these decisions.

To achieve this, a classic method involves delivering mild intracranial electrical stimulations in different brain structures, and estimating their impact on simple cognitive tasks. Such electrical stimulation procedures are carried out routinely as part of the standard presurgical assessment (Chauvel et al., 1993). This technique can be conducted intraoperatively or extraoperatively. Its major limitation is the triggering of seizures (Hamberger, 2007). For additional pre-surgical evaluation, two main methodological approaches can be used, that we detail below.

The first method, iEEG, involves multi-contact depth electrodes that are implanted inside the brain. These enable measuring coherent activity of local neuronal populations in the vicinity of the recording sites. The electrodes can be stereotactically implanted. This means that they are placed in specific brain structures by reference to a standard atlas (Bancaud et al., 1965; Chauvel et al., 1996). This is the method used in two of the articles discussed in this review (Basirat et al., 2008; Mainy et al., 2008). Alternatively, the electrodes may be placed strictly on the basis of the patients' MRI and its macroscopic structures (e.g., Heschl gyrus, or Broca's area), and later referenced to a standard atlas. This is the method used by Sahin et al. (2009), also discussed in this review. The electrodes, generally between 5 and 15 of them, remain for durations between 1 and 3 weeks. The presence of these electrodes allows recordings over unaffected brain tissue during periods of normal activity in the patient's room (i.e., not during the surgical procedure).

Electrocorticography (ECoG) is the alternative major technique, which was used in all the other studies we reviewed. In this case, subdural grids consisting of 2D arrays of 64-channels 8 × 8 electrodes are positioned directly on the lateral surface of the brain. The location of these electrode-grids with respect to underlying cortical gyral and sulcal anatomy is determined by coregistration of pre-implantation volumetric brain MRI with post-implantation volumetric brain CT. While subdural ECoG grids provide widespread cortical coverage and cortical maps of gyral activity, the iEEG electrodes record activity from both sulci and gyri and go beneath the cortical surface to deep cortical structures.

As is the case with surface EEG, cortical functional mapping can be based on various kinds of data-processing (for review, see Jacobs and Kahana, 2010). First, the signal as it unfolds in time can be averaged across trials, with the reference time being either the onset of stimulus (i.e., stimulus-locked average) or the onset of the overt response (i.e., response-locked average). This averaging yields cortical event-related potentials (ERP), which are an electrical signal generated by neuronal networks in response to a behaviorally significant event. Depending on the location and size of the electrodes (including the reference electrodes), ERPs may integrate neural activity over a range of spatial scales: surface EEG integrates activity over centimeters whereas intracranial ECoG integrates activity on a submillimeter to millimeter scale. Intracranial ERPs are generally referred to as a local field potentials (LFP) in reference to their highly localized origin (Bressler, 2002). The second type of analysis that can be conducted requires decomposing the signal into its frequency components, and then averaging these components across trials. Depending on the types of neural activity that occur in specific language tasks, cortical networks may display different states of synchrony causing cortical signal to oscillate at different frequency bands, referred to as delta band (0–4 Hz), theta band (5–8 Hz), alpha band (9–12 Hz), and gamma band (typically between 40 and 100 Hz; Donner and Siegel, 2011). Analyzing the change in the power spectra within each of these frequency bands provides information on the functional process that generates them. Such analysis is ideally suited for linking neuronal activity to language functions (notably language production), because language related cortical activity is prominently reflected in sustained activities that are not phase locked to external events (i.e., it will be less apparent in ERPs than in oscillatory activity). This is because many linguistic processes result from intrinsic network interactions within the brain (summarizing: from top-down modulations) rather than from an external drive (summarizing: from bottom-up activity). In this context, the most commonly used spectral profile in language cartography is the variation in power in gamma band, especially in high gamma bands ($\gamma_{\text{high}} = 80$–100 Hz, or sometimes 70–160 Hz) which is known to be a robust correlate of local neural activation (Chang et al., 2010). γ_{high} has been shown to be useful for detecting regional processing differences across language tasks, as well as across a variety of cognitive functions (e.g., Crone, 2000; Crone and Hao, 2002; Towle et al., 2008). Recent findings further suggest that cognitive tasks or processing levels (including in tasks requiring the production of words) induce the variations beyond the γ band (i.e., in lower frequency bands; Gaona et al., 2011).

The combination of epilepsy surgery and cognitive neurophysiology has provided a unique window into brain–behavior interactions over the past 60 years. Recent general reviews are provided by Engel et al. (2005) and by Jacobs and Kahana (2010). This being said, there are critical limitations that have to be kept in mind when deriving inferences and generalizations to the healthy population.

First of all, it is critical that recordings are obtained from normal healthy brain tissue, as distant as possible from the epileptogenic zone. In addition, recordings should be obtained at significant temporal distance from the occurrence of seizures, to avoid acute effects linked to seizure activity.

In the spatial domain, iEEG data have a spatial resolution defined in millimeters yet the electrode implantation scheme only provides a restricted sampling of selected cerebral structures. A complete 3D coverage of the brain with a spatial resolution of 3.5 mm has been estimated to require about 10,000 recording sites (Halgren et al., 1998), while the number of sites that are typically recorded with iEEG is approximately 100. Moreover, iEEG data analysis is complicated by the fact that patient populations are often small, and that recording sites are highly variable across patients. A comprehensive view of the large-scale networks involved in various cognitive tasks would therefore require combining data from multiple subjects with both overlapping and complementary electrode positions (a constraint that applies similarly to ECoG grid techniques).

Turning to the time domain, data acquired with this methodology have a temporal resolution in milliseconds. To be interpretable, however, the data need to be aggregated across trials. Thus the time resolution of the phenomenon that are described are rather in the order of tens or hundreds of milliseconds. Finally, a frequent practical limitation concerns the time-frame of participant availability for the cognitive tasks. Because this is often rather limited, the amount of data collected may be small. For this reason, it should always be kept in mind that the pattern of significant gamma band modulation is likely to be underestimated. Had there been more testing time, more electrodes may have shown significant effects.

Overall, however, when these limitations are dealt with carefully, the conclusions can be reasonably generalized beyond the population of epileptic patients. Many iEEG studies have provided spatio-temporal information about a wide range of cognitive processes (e.g., auditory perception, language, memory). These have been shown to be consistent with data from healthy participants, and have even provided the first threads of evidence later corroborated in healthy populations (Liégeois-Chauvel et al., 1989, 1994; Halgren et al., 2006; Axmacher et al., 2008; McDonald et al., 2010). For instance, Liégeois-Chauvel et al. (1999) demonstrated that enhanced sensitivity to temporal acoustic characteristics of sound in left auditory cortex underlaid the left hemispheric dominance for language. Such conclusion has been corroborated by observations from healthy participants (Trébuchon-Da Fonseca et al., 2005). Interestingly, the epileptic population under consideration includes patients suffering from mild to severe linguistic impairments, as well as patients with no apparent linguistic deficit (Mayeux et al., 1980; Hamberger and Seidel, 2003; Hamberger, 2007; Trébuchon-Da Fonseca et al., 2009). For these reasons, we argue that the iEEG studies can provide detailed spatio-temporal information about the dynamics of language production, as discussed below.

INTRACRANIAL ACTIVITY DURING LANGUAGE PRODUCTION

Our review of empirical studies is primarily organized according to the modalities used to elicit language production responses: auditory or visual stimuli. This is motivated by theoretical and clinical considerations. Theoretically, this distinction is thought to provide the most appropriate classification for capturing the spatio-temporal dynamics at stake. Major differences in the speed at which the input is perceived and decoded across modalities may result in major differences in the brain dynamics underlying language production. Furthermore, the networks involved in auditory and visual language production tasks are significantly different. From the clinical perspective, electrode, and subdural grid location are guided by surgical considerations only, and turn out to be highly variable from one participant to another. This has a strong influence on the kind of tasks that a participant may be asked to perform (e.g., auditory but not visual tasks for a patient implanted in superior temporal gyrus).

In addition to the above, we also included studies in which task instructions did not explicitly require that language was produced, but in which the pattern of neural activity indicated that this was most likely to be the case. These are reviewed in the Section on "Other Experimental Tasks."

LANGUAGE PRODUCTION ELICITED BY AUDITORY INPUTS

Among patients suffering from pharmaco-resistant epilepsy, temporal regions are commonly involved, and the posterior part of the superior temporal gyrus is often explored in order to know its possible involvement in seizures and/or to determine the posterior border of cortical excision. These explorations can be conducted by asking participants to repeat linguistic materials they hear (e.g., syllables, words, or sentences: Creutzfeldt et al., 1989; Crone et al., 2001; Crone and Hao, 2002; Towle et al., 2008; Fukuda et al., 2010; Pei et al., 2011), or to engage in a deeper processing of the auditory stimulus (e.g., in word association, definition, and verb generation tasks: Edwards et al., 2010; Thampratankul et al., 2010). These differences in tasks induce differences in the processes engaged to trigger the response, and in the corresponding brain activities. Most of the studies of this kind have shown reliable activity changes in γ_{high} time-frequency spectra. Note that these changes are very focal, and that event-related responses were clearly observed only in approximately one-fifth of the electrodes across the different studies. Here, and elsewhere, this should not be taken to indicate exclusive focal activity on these sites, given that the amount of testing and data collected is often small (see Previous Section and General Discussion). Such evidence should rather be thought of as providing a partial window on the activity of the underlying network.

Broadly speaking, perception, and overt repetition of linguistic materials involve a network comprising posterior superior temporal gyrus (pSTG), and inferior Rolandic gyri. The activation (i.e., increased γ activity) is first seen in pSTG, with a peak around 150 ms post-stimulus. The inferior rolandic gyri are activated somewhat later, around 200 ms before overt vocal response. In the course of response articulation, generally 100 ms after response onset, a second period of activity is seen in pSTG. We discuss these two sites in turn.

Activity in pSTG corresponds to auditory processing of the verbal stimulus, be it externally delivered (first cortical activity) or actually produced by the speaker (second cortical activity). For example, Fukuda et al. (2010) report a series of 15 patients in which STG is sequentially activated with a peak of γ oscillation 500 ms prior the onset of articulation and another one 100 ms after speech onset. This second gamma (γ) activity, linked to overt repetition, was smaller than the first. In between, around the onset of syllable articulation, there was little modulation of gamma oscillations. Crone et al. (2001; see also Crone and Hao, 2002) also observed this sequence of two peaks of activity during oral word repetition at a comparable location. This study reported the single case of a right-handed woman, implanted with ECoG in left temporal, left peri-sylvian, and left basal temporal occipital areas. This patient was a bimodal bilingual (English and ASL) and was tested with oral and signed responses (for further evidence on signed responses see Knapp et al., 2005). In the case of signed responses, only the pre-response peak was present, which is consistent with the idea that it reflects auditory or phonological processing of the verbal stimulus. This pattern has also been observed in another series of 12 patients by Towle et al. (2008). These authors reported that activation in the γhigh band associated with word perception included pSTG and lateral parietal regions (i.e., Wernicke's and surrounding areas). A posterior shift in the distribution of gamma activity is reported when the patients heard the word compared to when they spoke the word. An additional response in Broca's area was also observed in this study, which started 800 ms before the voice onset time. However, this latter activity could be related to the conditions with which repetition was elicited. Only words that had been heard in a previous block were to be repeated. In this case, the frontal response could thus be tied to decision and discrimination processes rather than linguistic processing *per se.*

The pre-response peak repeatedly observed in pSTG allows for a more detailed interpretation. In Fukuda et al. (2010), the peak has different latencies in stimulus- vs. response-locked averages (i.e., if the analysis focuses on the onset of the stimulus or on the onset of the response). Stimulus-locked, the peak is present shortly after stimulus (peak ~260 ms). Response-locked, the peak starts around 500 ms pre-articulation, which is considerably later given that average response time was about 1000 ms. This suggests that this response not only reflects processing *evoked* by external auditory inputs but also longer lasting *induced* preparatory activity (Alain et al., 2007), for example phonological encoding processes. In this respect, the data from the single case reported by Crone et al. (2001) are somewhat different. pSTG activity is only present time-locked to stimulus for 500 ms, followed by a post-response activity at 1500 ms. In contrast, a more anterior recording site showed only pre-response (not post-response) activity, suggesting a role of the later site only in phonological encoding and preparation processes. We come back to this anterior–posterior contrast in Section "General Discussion," after we have presented the evidence from the other modalities in the following sections.

No significant difference in the peak amplitude, onset latency, or peak latency of gamma activity has been reported between the left and right STG (Fukuda et al., 2010). In contrast, the later modulation of gamma oscillation recorded from inferior Rolandic sites displayed a left-hemisphere advantage. The peak of activation was earlier in left compared to right hemisphere. For simple syllables, this activity starts 200 ms before articulation onset and peaks 130 ms after it. Furthermore, a subset of inferior rolandic sites showed phoneme-specific patterns of gamma-augmentation, mostly located on the left side (for a thorough investigation of cortical signal classification to discriminate linguistic materials, i.e., words, see Kellis et al., 2010). Overall, then, these data suggest that primary sensorimotor area on the left side may have a predominant role in movement execution for phoneme articulation, in agreement with Chang et al. (2010; for comparison, Brooker and Donald, 1980, provide a critical discussion of lateralization effects observed at similar timings in surface recordings).

In some cases, participants are asked to engage in deeper processing of the stimulus to construct the response (e.g., word association task, verb generation task, response to definition; Brown et al., 2008; Edwards et al., 2010; Thampratankul et al., 2010). Quite expectedly, this leads to a broader pattern of activity. In the verb generation task, auditory stimulation is followed by gamma activation shifts from pSTG (100 ms post-stimulus) to the posterior part of middle temporal gyrus (pMTG), the parietal operculum, the temporo-parietal junction (300 ms), and the precentral gyrus (superior portion of ventral pre-motor cortex svPM: 400 ms post-stimulus). The middle frontal gyrus and the left inferior frontal gyrus (IFG) are activated later, namely 700 ms after the stimulus onset and around 300 ms prior to verbal responses (Brown et al., 2008; Edwards et al., 2010; Thampratankul et al., 2010).

The propagation of pSTG activation to MTG has been primarily linked to semantic association processes (retrieving the color of a fruit or the answer to a question; Brown et al., 2008; Thampratankul et al., 2010). The more posterior portion, proceeding from the planum temporale and terminating in the mid-to-posterior superior temporal sulcus (STS) connected with the temporo-parietal junction has been linked to verbal phonological working memory and word production (Edwards et al., 2010). The activity seen later in the left supramarginal gyrus (SMG) does not have a well defined function in this data set. Note however that functional imaging data associates fMRI bold responses in this area with phonological processes (Démonet et al., 2005; Vigneau et al., 2006). Finally, the medial pre-frontal cortices (including the supplementary motor area, the pre-supplementary motor area, and the cingulate gyrus) have been associated to voluntary control over the initiation of vocal utterances (Brown et al., 2008; see also Alario et al., 2006, for fMRI evidence).

One issue that runs across all of these studies is the attenuation of gamma responses in STG in speaking vs. listening conditions, which is also seen in non-human primates (Müller-Preuss and Ploog, 1981). In humans, Creutzfeldt et al. (1989) recorded single units while participants repeated sentences. STG neurons showed reduced responsiveness to self-produced speech compared to repetition and naming. This observation was also made by Fukuda et al. (2010), Towle et al. (2008), or Crone et al. (2001). Flinker et al. (2010) report a more specific investigation of this issue. With a phoneme repetition task, they showed that auditory cortex is not homogeneously suppressed but rather that there are fine grained spatial patterns of suppression. They conclude that every time we produce speech, auditory cortex responds with a specific pattern of suppressed and non-suppressed activity. This reduced responsiveness could be due to corollary discharges from motor speech commands preparing cortex for self-generated speech (Creutzfeldt et al., 1989, Towle et al., 2008). One difficulty when investigating this issue is to control for the volume of auditory stimuli in speaking vs. listening conditions. Yet this would be important because this intensity variable is known to affect the magnitude of the brain response (Liégeois-Chauvel et al., 1989).

LANGUAGE PRODUCTION ELICITED BY VISUAL INPUTS

The kinds of visual inputs most generally used include pictures of common objects to be named, overtly or covertly, and words to be recognized and read (Hart et al., 1998; Crone et al., 2001; Crone and Hao, 2002; Tanji et al., 2005; Usui et al., 2009; Edwards et al., 2010; Cervenka et al., 2011; Pei et al., 2011; Wu et al., 2011). In general, the visual modality engages the left baso-temporal region, often referred to as baso-temporal language area (BTLA), which includes fusiform gyrus (FG) and ITG. Activation starts in BTLA around 200 ms after picture presentation. This activity precedes pSTG activity, occurring around 200–600 ms for reading, and 400–750 ms for picture naming, which is much later than the comparable pSTG activity in auditory tasks described in the previous section (∼100 ms post-stimulus; Crone et al., 2001; Edwards et al., 2010). The BTLA presumably plays a crucial role in lexico-semantic processing and picture recognition (Crone et al., 2001; Edwards et al., 2010).

A detailed chronometric analysis of the involvement of BTLA is provided by Hart et al. (1998), on the basis of a pair of electrodes in a single patient. A functional response is recorded around 250–300 ms after visual presentation, and lasts between 450 and 750 ms, depending on the subjective familiarity of the object. Direct electrical cortical stimulation (ECS) delivered shortly after stimulus presentation caused a variety of language deficits (word-finding difficulties, empty speech, paraphasias, and speech arrest), the outcome of which was modulated by the timing of the stimulation. In particular, semantic disruptions were no longer present when the stimulation occurred 750 ms (or later) after the picture. Putting together the activation data with the so-called "time-slicing" cortical stimulation procedure provides lower and upper estimates for the engagement of BTLA in this task: in short, between 250 and 750 ms. These data were comparable in the two languages spoken by the patient.

Further information about the functional role(s) of left BTLA comes from the case of a Japanese speaker implanted with subdural electrodes on the basal temporal cortices bilaterally (Tanji et al., 2005), very close to the locations involved in Hart et al. (1998). The patient was asked to name pictures aloud, and to read silently Japanese words and pseudo-words, the latter being written in the syllabic script Kana or in the logographic and morphographic script Kanji. There were clear responses in the γ_{high} band in the three tasks. The pictures elicited bilateral responses on various recording sites, with a weak anterior–posterior distinction in the left response to animals and tools (animals leading to more intense signal on a more anterior site, and vice-versa). The written materials elicited left lateralized responses only, irrespective of script and lexicality. Pair-wise comparisons of the amount of activation between these conditions showed the following pattern: Kanji pseudo-words > Kanji words > Kana words ∼ Kana pseudo-words. On the basis evoked potentials recorded at similar locations on two patients, Usui et al. (2009) suggested there may be a distinction between anterior and posterior responses to the two scripts in this brain region.

This global pattern does not lend itself to a simple interpretation, however. Increased activity for Kanji pseudo-words compared to Kanji words could reflect the involvement of BTLA in semantic processing, for example if participants effortfully attempted to reach a semantic interpretation of the pseudo-words (Tanji et al., 2005, p. 3291 bottom). The distinction between animals and tools in picture naming would be consistent with this view. However, the similar contrast between words and pseudo-words presented in the Kana script did not produce the same gamma activation difference. It is possible that, compared to logographic Kanji, syllabic Kana promotes phonological processing at the expense of semantic processing, and thus the semantic effect is blurred with this script. This would also explain why there was increased activity for Kanji compared to Kana script at this location. The authors conclude that the overall pattern reflects a role of BTLA at the lexical level, as a convergence zone midway between word form and word meaning. While this interpretation is consistent with the evidence, it is formulated in broad terms and a more detailed account may require further studies (see Wu et al., 2011, for some recent further ECoG evidence, and Usui et al., 2009, for more extensive data on electrical stimulation).

Most of the articles discussed above report activation in IFG without focusing on it. Sahin et al. (2009), however, focused on Broca's area and neighboring regions, in an iEEG study involving three patients implanted in this region. The task they used involved processing a visual stimulus (e.g., word) according to a grammatical rule of English (e.g., transforming a singular noun "horse" to plural "horses," or transforming a verb in present tense "watch" to past tense "watched"). Across all three patients, evoked potentials were recorded at ∼200, ∼320, and ∼450 ms post-target onset. The first 200 ms peak was modulated by lexical manipulations (frequency of use). The second 320 ms peak was modulated by the nature of the task (grammatical manipulations, see examples above). The third 450 ms peak was modulated by articulatory requirements (length of the response in syllables). These results indicate that distinct linguistic processes can be distinguished at a high temporal resolution within the same locus. They also suggest that Broca's area is not dedicated to a single kind of linguistic representation but comprises adjacent but distinct circuits which implement different levels of processes. Although the authors favor a staged model in which these different processes are performed sequentially, these data by themselves do not argue against more integrated processing (e.g., in the form of cascading: Goldrick et al., 2009).

OTHER EXPERIMENTAL TASKS (THAT MAY ENGAGE LANGUAGE PRODUCTION PROCESSES)

This section includes the studies in which participants were not explicitly asked to produce language, but in which language production processes can nonetheless be suspected to have been engaged. For example, patients may be asked to press buttons (rather than speak) on the basis of the linguistic materials they are currently processing (Basirat et al., 2008; Mainy et al., 2008; Chang et al., 2010). Note that, while this studies may provide valuable information about language processing in general, the inferences ones draws with respect to language production processes themselves should be made cautiously. This is because previous comparison between neural responses in overt and covert conditions, especially in fMRI, have shown notable differences. As would be expected, all reports agree that overt responses lead to greater involvement of motor cortices than covert production. A broader brain network is engaged in overt than in covert conditions, including mesial temporal lobe as well as sub-cortical structures (Rosen et al., 2000; Shuster and Lemieux, 2005; Forn et al., 2008; Kielar et al., 2011). Importantly, it has been argued that overt and covert networks show distinct patterns of activation (Borowsky et al., 2005; Shuster and Lemieux, 2005). Barch et al. (1999) concluded that covert conditions cannot be used as simple substitutes for overt verbal responses. One demonstration of such differences between activity linked to overt and covert responding in iEEG comes from Pei et al. (2011) reviewed above. These authors briefly compared these two conditions. They observed that a major difference was in the modulation of pSTG (Wernicke's area) after acoustic processing, while other activations were increased in the covert modality (BA22, mid-STG, BA41/42, and the temporo-parietal junction).

In the first study reviewed in this section, Mainy et al. (2008) tested 10 patients on a hierarchy of judgments on visually presented words and pseudo-words, regarding either the meaning (living vs. non-living word categorization task), the visual properties (analysis of consonant strings) or the phonological make up (rhyme decision task on pseudo-words)[1]. Notably in the latter case, access to phonological information may approximate the processes engaged during language production, for example during overt word reading. Once again, the measure of interest that was analyzed in most detail is the change in spectral power within the γ_{high} band. The results for the early stages of processing are consistent with those reviewed above concerning the visual modality. A rather abrupt onset of activity occurred around 200 ms post-stimulus in associative visual areas, irrespective of the task. With about 100 ms delay, STG showed increased activity in the semantic and phonological tasks, not in the visual property judgment task. This specific response is presumably due to the fact that, while all the stimuli were visual, the materials were pronounceable only in the semantic and the phonological tasks. Relatively similar responses were recorded in the more anterior middle temporal gyrus on some patients. Finally, around 400 ms, a response peaked in IFG, being larger in the phonological than in the semantic task (occasionally the semantic response was larger in more anterior sites). Note that the semantic task involved words and the phonological task involved pseudo-words. The authors link the IFG and the STS gamma band responses to the phonological retrieval processes and/or inner speech production. The respective roles (and their precise interaction) of these two regions stand as an important issue to be clarified.

Chang et al. (2010) also asked patients to perform phonologically based decisions. In this case, participants heard sequences of syllables and had to press a button when they heard a pre-specified target. A secondary control task required patients to repeatedly produce a syllable (/pa/) or a vowel (/a/). In the perception task, an early activity (<120 ms post stimuli) rises sharply in dorsal STG, both to target and non-target items. Quickly after, activity in ventral STG becomes larger for targets. Around 120 ms later, high gamma activity is measured in superior ventral pre-motor cortex (svPM), again larger for target items. The articulation task elicits responses at these same two locations in reverse order: svPM activity starts rising before vocal onset, and STG is activated after vocal onset, presumably as a result of auditory feedback (see Chang et al., 2010, for details on the response decision processes recorded, notably, in pre-frontal cortex). This suggests that the svPM observed early on during speech perception is closely linked to speech-motor activities. The authors discuss how motor cortex may actively participate in sublexical speech perception, perhaps as pSTG accesses the articulatory network to compare externally driven auditory representations with internal motor representations. Additionally, it is noteworthy that STS regions were *not* active prior to articulation, as they were in various studies reviewed above. This could be due to the repetitive nature of the task, whereby the response is not encoded anew on every trial but rather stored ready elsewhere. The production data reported are too scarce to clarify this point.

Finally, Basirat et al. (2008) report a study whose methodology may be fruitful to investigate the monitoring processes that accompany language production (Postma and Oomen, 2005). The original motivation was to investigate so-called multistable perception, i.e., perceptual changes occurring while listening to a briefly cycling stretch of speech. Two patients with iEEG electrodes implanted in frontal, superior temporal, and parietal areas were asked to listen to sequences of repeated syllables. Two experimental conditions were contrasted. In the first one, the two syllables alternated regularly (e.g., /pata. . ./), and patients were asked to press a button whenever they perceived a change in the repeated utterance. In the second condition, the alternation was random (e.g., /. . .papapa. . .tatata. . .papapa. . ./), and patients were asked to detect transitions between/pa/and/ta/. In the first condition the button presses are elicited by endogenous perceptual changes,

---

[1] The different tasks were performed independently from one another and did not involve the same materials.

while in the second one they are elicited by exogenous changes in the signal. Contrasting these two conditions right before a transition was detected revealed significant gamma band activation in the left inferior frontal and supramarginal gyri, but not in temporal sites. This activity could be attributed to the endogenous emergence of the varying speech forms. The authors note the involvement of phonological comparison and decision making in perceptual transitions. These are indeed two standard components of speech monitoring accounts (Christoffels et al., 2007; Möller et al., 2007; Riès et al., 2011).

GENERAL DISCUSSION

This review encompasses articles in which intracranial recordings of patients involved in language production tasks are reported. The combination of the observations made in the different studies provides a patchy yet informative view of the spatio-temporal brain dynamics involved in language production. The review also provides some clues regarding the relative merits of different intracranial indicators of cognitive processing, and allows a number of considerations regarding the relationship between these measures and the gold-standard of brain-function mapping (i.e., brain stimulation). Below we discuss these points in turn. We then finish with some considerations about the amount of available evidence, and avenues for future research using this methodology.

OVERVIEW OF THE SPATIO-TEMPORAL DYNAMICS UNCOVERED

The view that emerges from the ECoG studies we have reviewed is summarized on **Figure 1**, for speech elicited auditorily, and on **Figure 2**, for speech elicited visually. These figures, as well as the discussion below, focus on the left hemisphere.

In the auditory tasks, brain activity starts by being rather focused on the superior temporal gyrus and sulcus (middle and posterior parts), between 200 and 500 ms post-stimulus. At this point of time, a rather consistent activity is also seen in the inferior part of frontal peri-rolandic areas. Later, between 500 and 800 ms post-stimulus, a somewhat broader activation network involves STG and STS, as was the case previously, but also SMG, IFG, and a larger part of peri-rolandic areas. The relative involvement of the different parts of this network seems to be modulated by task demands (e.g., verb generation vs. word repetition). Time-locked to response onset, an overall stable network is observed both before and after speech onset. As in the previous epochs, this network is mostly focused around left peri-sylvian areas, with a larger area of activity in inferior frontal areas before response (notably in the verb generation task).

In the visual tasks, consistent activity is detected in the basal temporal region. This activity starts in its most posterior part (200–500 ms post-stimulus), and is seen later (500–800 ms) in the middle part. A much more anterior locus of activity (temporal pole) is seen unexpectedly early (200 ms post-stimulus), but only in one patient (Cervenka et al., 2011). Concomitantly with this baso-temporal progression, consistent activities have been reported in the posterior part of MTG and the middle and posterior parts of STS, but not in STG, which was clearly seen in the auditory tasks. The similarity between visual and auditory tasks, however, is clearly apparent within the 500–800 ms time window,

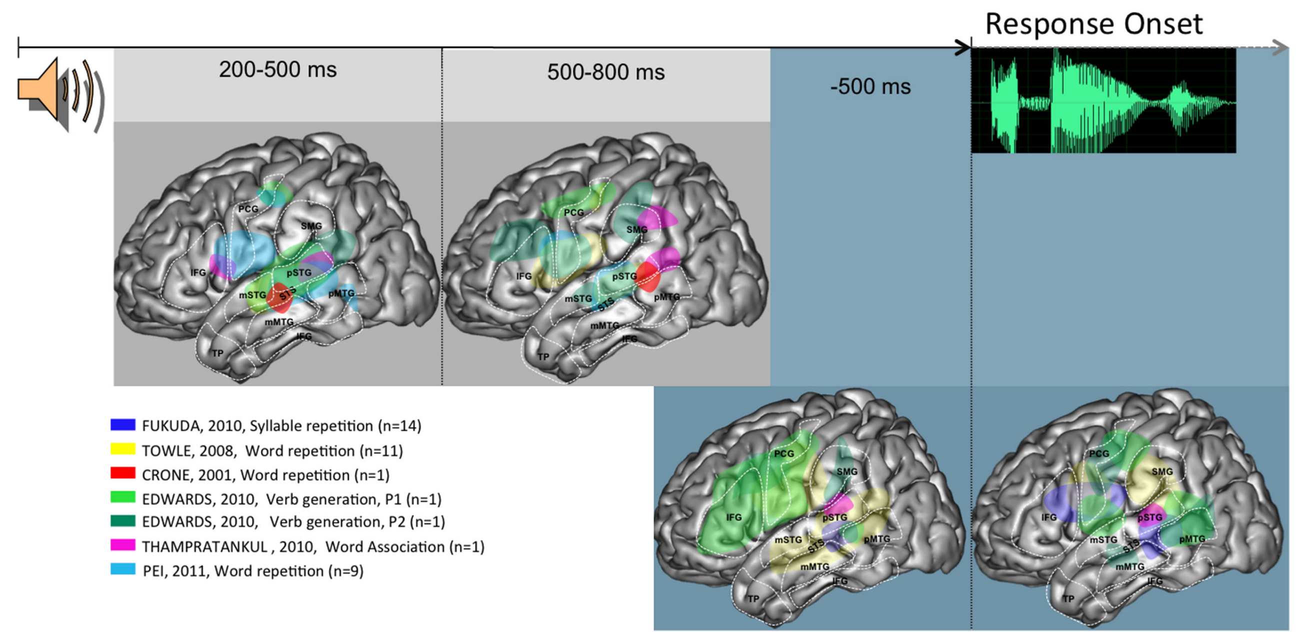


FIGURE 1 | Summary of activities observed across patients and studies when language production is triggered auditorily, projected on a standardized left hemisphere. The two figures on the gray background represent activities time-locked to stimulus presentation, at the indicated timings. The two figures on the blue background represent activities time-locked to response onset. Further details are provided in the Section "Overview of the Spatio-Temporal Dynamics Uncovered" in the General Discussion. Abbreviations used: IFG, inferior frontal gyrus; STS, superior temporal sulcus; SMG, supra marginal gyrus; pSTG, posterior part of the superior temporal gyrus; mSTG, middle part of the superior temporal gyrus; pMTG, posterior part of the middle temporal gyrus; mMTG, middle part of the middle temporal gyrus; PCG, pre central gyrus; TP, temporal pole.

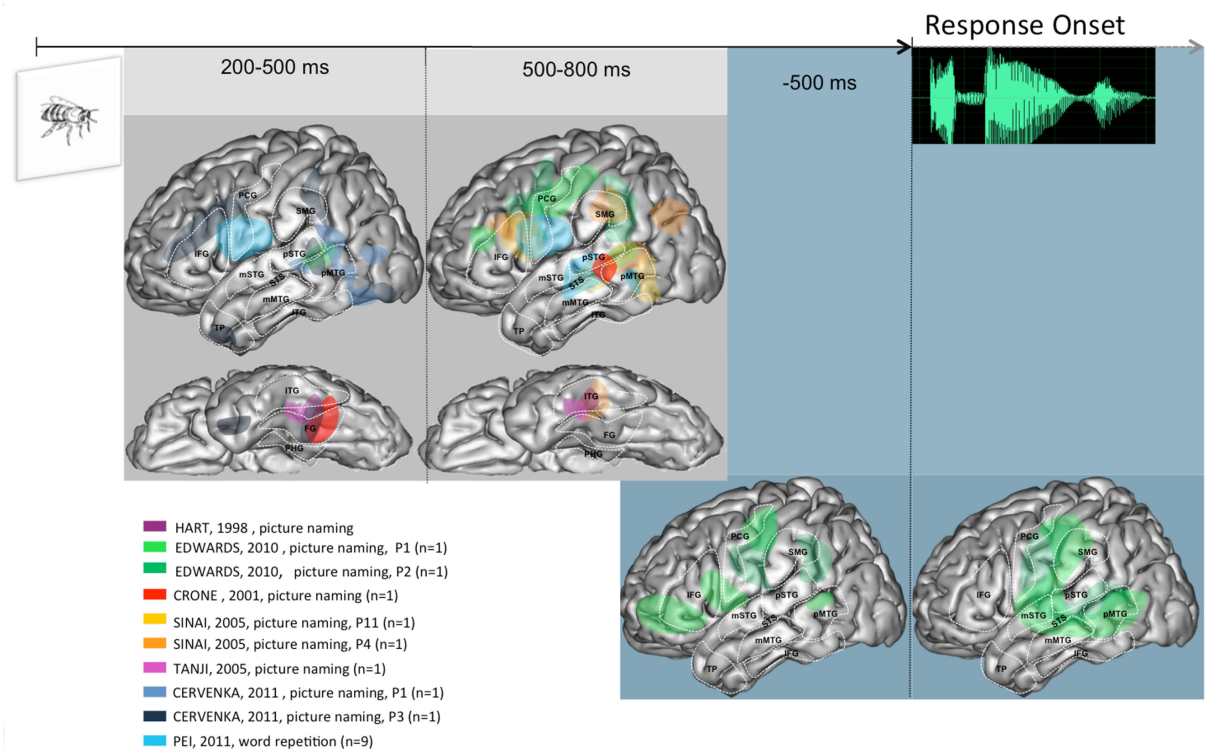


FIGURE 2 | Summary of activities observed across patients and studies when language production is triggered visually, projected on a standardized left hemisphere. The two figures on the gray background represent activities time-locked to stimulus presentation, at the indicated timings; lateral and basal views shown. The two figures on the blue background represent activities time-locked to response onset. Further details are provided in the Section "Overview of the Spatio-Temporal Dynamics Uncovered" in the General Discussion. Abbreviations used: IFG, inferior frontal gyrus; STS, superior temporal sulcus; SMG, supra marginal gyrus; pSTG, posterior part of the superior temporal gyrus; mSTG, middle part of the superior temporal gyrus; pMTG, posterior part of the middle temporal gyrus; mMTG, middle part of the middle temporal gyrus; PCG, pre central gyrus; TP, temporal pole; FG, fusiform gyrus; ITG, inferior temporal gyrus.

in inferior frontal and peri-rolandic areas, as well as in SMG. Finally, time-locked to response onset, very few data are available. The pattern seen in the two patients reported by Edwards et al. (2010) is in keeping with the peri-sylvian networked discussed for auditory tasks.

Overall, these data allow a number of general conclusions. There is a clear early modality effect, whereby auditory tasks recruit STG while visual tasks recruit basal and lateral temporal areas. The possible convergence between these modality specific activities may be in STS, which is a plausible neural correlate of phonological encoding processes (Edwards et al., 2010). These could also involve SMG, whose activity is also seen across modalities, but somewhat later (500–800 ms window). Peri-rolandic areas are recruited in the two modalities relatively early (200–500 ms window). This suggests a very early involvement of (pre-) motor processes, which is consistent with the hypothesis of a dorsal stream in verbal processing. Around and time-locked to response onset the broad peri-sylvian network is not easily characterized in specific cognitive terms, given the reviewed evidence.

What is also clear from this review is the great heterogeneity in the data sets available, across patients and tasks. Our discussion above therefore had to consider both very general phenomena that seem to be reproducible across patients or studies, and more specific hypothesis that have only been discussed or tested in specific studies, or with specific patients. We come back to this issue, in more general terms, in the section below on the limits of this methodology.

SOME LIMITATIONS OF THE REVIEWED EVIDENCE

The number of articles published in this thread of research is remarkably limited. Despite our use of a rather broad scope, in accordance with the topic of this special issue, we could only find about 25 articles in which patients produced language in one way or another while their brain activity was recorded intra-cranially. The number of individuals sampled in each article is also relatively low. Only three of the articles report evidence from more than ten patients (Sinai et al., 2005; Towle et al., 2008; Fukuda et al., 2010). The remaining studies report evidence from four or fewer patients, and six of them focus on single cases (Hart et al., 1998; Crone et al., 2001; Crone and Hao, 2002; Tanji et al., 2005; Thampratankul et al., 2010). The single-case approach is undoubtedly appropriate in this context, in light of the inter-individual variability visible in the few studies reporting more than one patient. However, generalizations from these data to normal function should only be made when the reliability of a given phenomenon has been examined across individuals. Somewhat paradoxically, this relatively limited sample of data was obtained using a great variety of tasks and theoretical approaches, presumably because of clinical motivations underling a great share of the tasks used. While this provides a wide sampling of evidence, it also complicates the comparisons across studies whenever a fine grained definition of cognitive processes is to be used. It remains to be seen whether future research will converge on some specific theoretical questions and experimental paradigms.

COMPARISON WITH OTHER METHODOLOGICAL APPROACHES

The networks revealed by the meta-analysis of Indefrey and Levelt (2004) on diverse brain imaging techniques, and by Price (2010) on fMRI only, already discussed in the Section "Introduction," are largely consistent with one another. The main difference between these two reviews is that Price (2010) reports in more detail areas sub-tending input processes, and that only Price (2010) reports the involvement of medial frontal areas (e.g., pre-SMA) in volition, selection, and execution (see also Alario et al., 2006). The language production network is, by and large, consistent with the activation localizations reported on **Figures 1** and **2**. Likewise, the timing of language production operations that emerges from the reviews by Indefrey and Levelt (2004) and by Ganushchak et al. (2011) is consistent with what is reported here.

Given this context, it becomes interesting to compare in some more detail link between spatial and timing localization of cognitive events, in other words the spatio-temporal dynamics uncovered by these different meta-analysis and reviews. Here, some notable inconsistencies seem to emerge.

Regarding activity in the temporal gyrus, our review does not reveal early middle temporal activity (150–255 ms post-stimulus according to Indefrey and Levelt, 2004) but only posterior temporal gyrus activity. This activity is mostly present between 200 and 800 ms in the visual tasks, which is quite comparable with the 200–400 ms estimate of Indefrey and Levelt (2004). By contrast, this activity is almost absent (and if anything occurs earlier) in the auditory tasks. Thus, the data reviewed here do not support (but neither clearly contradict) the anterior to posterior propagation along the middle temporal gyrus associated to lexical to phonological pathway by Indefrey and Levelt (2004). Note that recent MEG evidence suggests a "reverse" posterior to anterior propagation in picture naming (Liljeström et al., 2009; see also Edwards et al., 2010, for further discussion). Additionally, the most anterior part of STG associated to monitoring and auditory object processing in Indefrey and Levelt's (2004) meta-analysis is not consistently activated in the studies we reviewed. Note that this region was commonly recorded and that monitoring was presumably engaged in the tasks that were used.

Another relevant point of inconsistency concerns the role of different frontal areas. The intracranial studies we reviewed consistently report very fast responses in peri-rolandic areas (around 200 ms post-stimulus), also seen in MEG during picture naming (Liljeström et al., 2009, Figure 3). This is much earlier than the articulatory timing specified by Indefrey and Levelt (2004). While the interpretation of this fast response remains to be settled, its existence seems to go against a very sequential view of the word production process (Goldrick et al., 2009). Finally, IFG is also activated during language production, but, while Price (2010) highlights its role in early stage which is selection process, as we did when considering the data from Towle et al. (2008), Indefrey and Levelt (2004) focus on syllabification process only for this region.

This brief comparison of results across methodologies indicates that it is a real challenge to arrive at an integrated view of the dynamics of brain activity during language production that is consistent with data from the different available techniques (see also Jerbi et al., 2009; Liljeström et al., 2009). In this context, and with its inherent strengths and limitations, iEEG can provide a powerful method for testing specific explicit hypothesis derived from the meta-analysis of surface EEG and fMRI data and thus to provide valuable details about the spatio-temporal dynamics of language production.

COMPARISON OF INTRACRANIAL ACTIVATION MEASURES TO MAPPING WITH ELECTRICAL CORTICAL STIMULATION

The primary motivation for using invasive brain-activity recording methods, such as iEEG and ECoG, is to help delineating between dysfunctional and functionally eloquent tissue. Yet ECS still is considered as the gold standard for this purpose, whereby elicited focal activity changes induce language task interruptions. It is of clinical importance to compare these methods because in some respects iEEG recordings have a number of advantages over brain stimulation (e.g., iEEG allows fast parallel recording of multiple sites vs. time-consuming sequential recording of individual sites; stimulation can inadvertently influence distant areas through axonal connections). The comparison is also interesting from a cognitive perspective. The two methods do not always provide exactly the same information about which areas are involved. Every patient in which both methods are tested is bound to show sites with converging patterns (significant ECoG effect and disruption through ECS, or neither) as well as sites with diverging patterns (either an ECoG effect and no ECS disruption, or the opposite; see more on this below). This calls for caution when drawing inferences about the healthy brain (just as anatomo-functional correlations established with neuropsychological evidence and with brain activation data should be combined cautiously). In particular, the specific signal recorded in each study (e.g., the frequency ranges considered within or outside the gamma band) has an influence on the sites that may turn out to be significantly active.

Among the studies reviewed above, some report language production related gamma activity in areas that are largely concordant with those observed in brain stimulation (Towle et al., 2008; Fukuda et al., 2010). In contrast, Brown et al. (2008) or Thampratankul et al. (2010) found that the areas showing significant γ_{high} band augmentation were larger than the eloquent areas suggested by the electrical neuro-stimulation procedure. This discrepancy might in part be due to differences in the age of the population of interest, which was diverse across studies. Brown et al. (2008) point out various studies which report positive correlations between the age of patients and the number of sites where neuro-stimulation produced naming errors in language mapping (Ojemann et al., 2003; Schevon et al., 2007).

Two articles were directly devoted to a comparison between γ_{high} band recorded from ECoG and ECS for mapping the language production function. Sinai et al. (2005) probed ECS with different language production tasks, and focused on picture naming for the ECS–ECoG comparison. The primary goal was to test whether ECoG activity predicts ECS disruption. In their data, sites in which no ECoG activity is recorded are rather unlikely to disrupt naming during ECS. By contrast recording ECoG activity in one site does not provide a reliable indication that the site will show disrupted performance during ECS (i.e., the ECoG recording could well be a "false alarm"). These authors also point out that ECS may sometimes overestimate the cortical territory that is critical to function, as shown by occasional good post-operative performance while ECS disrupting sites had to be resected. For these reasons, the authors suggest that, at least for now, ECoG seems to be suitable to provide a preliminary functional map from all implanted subdural electrodes, and to determine cortical sites of lower priority during ECS mapping (i.e., those that do not show ECoG activity in the tasks of clinical interest). It could also be valuable in those cases or sites where ECS is not feasible (because of clinical seizures, after-discharges, or pain). Following a similar logic, Cervenka et al. (2011) show a highly variable degree of overlap between ECS and ECoG sites across four bilingual speakers. In particular, the two methods provide contrastive results regarding the degree of cortical overlap between first and second language. They conclude, in rather general terms, that ECoG provide a useful complement to ECS, notably with bilingual speakers.

ECoG could thus be, in principle, a useful peri-operative tool. Its use is made difficult, however, because of the requirement to conduct off-line statistical analysis on data collected from multiple trials (vs. the immediate "yes–no" answer stemming from ECS). This constraint may in part be relieved by conceiving a statistical procedure which can be implemented online during the surgical procedure. This is the goal of Roland et al. (2010), which seek to reduce the amount of data needed to obtain meaningful ECoG patterns by using an algorithm designed for real-time event detection. Signal modeling for real-time identification and event detection (SIGFRIED) provides such algorithm, along with a non-expert user oriented interface. These allow detecting online task-related modulations in the ECoG γhigh band while patients perform simple motor and speech tasks during awake craniotomy. Their findings indicate that a subset of areas identified by SIGFRIED correspond to those identified by stimulation mapping, without this identification taking much longer in the former case. This method may provide a realistic way, in peri-operative terms, to conduct preliminary mapping of functional sites prior to detail stimulation mapping of predetermined ECoG eloquent areas. This could also be used to circumscribe brain regions during experimental testing.

EVALUATING GAMMA BAND AS AN INDICATOR OF COGNITIVE LINGUISTIC PROCESSING

The primary measure in all the studies we have reviewed is the modulation of γhigh band activity linked to different processing stages. These studies establish that γ_{high} band (>70 Hz) provides a powerful means of cortical mapping and detection of task-specific activations (Crone et al., 2001, Crone and Hao, 2002; Canolty et al., 2007; Towle et al., 2008). Additionally, LFP (i.e., ERPs measured with iEEG, see Introduction) provide a much clearer view of the time course (e.g., Sahin et al., 2009).

Recent intracranial studies of language production have mainly utilized high gamma power rather than ERPs, presumably because they are more focal and are more direct indications of neural activation. However, frequency changes during cognitive tasks are not limited to gamma variations, and there could be important differences within the gamma band itself. Regarding the first point, Canolty et al. (2007) analyzed complex oscillatory responses and found that *theta* was the frequency that was most shared between electrodes. It seemed to be an important regulator of inter-regional communication during complex behavioral tasks (see also Korzeniewska et al., 2011, for a detailed analysis of functional connectivity). Regarding the second point, Gaona et al. (2011) provide evidence that modulations of different stretches of the γhigh band may show differential sensitivity to linguistic tasks and processing stages. Presumably, a complete picture of the spatio-temporal brain and cognitive dynamics involved in language production will only emerge from a full consideration of this intricate pattern of activities across frequency bands.

AVENUES FOR FUTURE RESEARCH

The evidence we have reviewed begs a number of unanswered questions regarding the spatio-temporal dynamics of the brain areas involved in language production. For example, activity in the visual tasks clearly engages lateral temporal areas, while this has not been described in auditory tasks. Note however the implicit "correlation" between the tasks used and the electrode implantation, whereby patients with more superior electrodes are more likely to be engaged in auditory tasks, while patients with lateral and basal electrodes are likely to be engaged in visual tasks. A more accurate description of the truly modality specific activities would benefit from cross-evidence where these two populations of patients are tested in both modalities. This would allow a more specific interpretation of the activity seen is STG and STS (see Edwards et al., 2010, for some hypothesis).

As another example, the early activity seen in peri-rolandic areas could suggest very early preparation of the response, or early engagement of motor areas in speech decoding, or both. This is an important issue, as it ties with the interaction between the perception and the action streams involved in language processing. Experimental tasks directly designed to clarify this kind of issues are still lacking detailed tests with this population. As a final example, the studies we have reviewed do not report (by lack of available data) any evidence about some brain structures that are known to be important for language production (e.g., temporal pole; Tsapkini et al., 2011). This leaves a number of open questions in our understanding of the brain dynamics in which they may be involved. They are potentially important testing grounds, should they be testable in a given patient.

Many specific aspects of word production remain largely unexplored with intracranial recordings. It is clear however, that combining specific cognitive hypotheses with the temporal and spatial resolution of this technique can provide a powerful tool to uncover the dynamics of language production.

ACKNOWLEDGMENTS

This research has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant agreement n 263575).

REFERENCES

Alain, C., Snyder, J. S., He, Y., and Reinke, K. S. (2007). Changes in auditory cortex parallel rapid perceptual learning. *Cereb. Cortex* 17, 1074–1084.

Alario, F.-X., Chainay, H., Lehericy, S., and Cohen, L. (2006). The role of the supplementary motor area (SMA) in word production. *Brain Res.* 1076, 129–143.

Axmacher, N., Schmitz, D. P., Wagner, T., Elger, C. E., and Fell, J. (2008). Interactions between medial temporal lobe, prefrontal cortex, and inferior temporal regions during visual working memory: a combined intracranial EEG and functional magnetic resonance imaging study. *J. Neurosci.* 28, 7304–7312.

Bancaud, J., Talairach, J., Bonis, A., Schaub, C., Szikla, G., Morel, P., and Bordas-Ferrer, M. (1965). *La stéréoélectroencéphalographie dans l'Epilepsie: Informations Neurophysiopathologiques Apportées par l'Investigation Fonctionnelle Stéréotaxique.* Paris: Masson.

Barch, D. M., Sabb, F. W., Carter, C. S., Braver, T. S., Noll, D. C., and Cohen, J. D. (1999). Overt verbal responding during fMRI scanning: empirical investigations of problems and potential solutions. *Neuroimage* 10, 642–657.

Basirat, A., Sato, M., Schwartz, J. L., Kahane, P., and Lachaux, J. P. (2008). Parieto-frontal gamma band activity during the perceptual emergence of speech forms. *Neuroimage* 42, 404–413.

Borowsky, R., Owen, W. J., Wile, T. L., Friesen, K. K., Martin, J. L., and Sarty, G. E. (2005). Neuroimaging of language processes: fMRI of silent and overt lexical processing and the promise of multiple process imaging in single brain studies. *Can. Assoc. Radiol. J.* 56, 204–213.

Bressler, S. L. (2002). "Event related potentials," in *The Handbook of Brain Theory and Neural Networks*, ed. M. A. Arbib. (Cambridge, MA: MIT Press), 412–415.

Brooker, B. H., and Donald, M. W. (1980). Contribution of the speech musculature to apparent human EEG asymmetries prior to vocalization. *Brain Lang.* 9, 226–245.

Brown, E. C., Rothermel, R., Nishida, M., Juhász, C., Muzik, O., Hoechstetter, K., Sood, S., Chugani, H. T., and Asano, E. (2008). In vivo animation of auditory-language-induced gamma-oscillations in children with intractable focal epilepsy. *Neuroimage* 41, 1120–1131.

Canolty, R. T., Soltani, M., Dalal, S. S., Edwards, E., Dronkers, N. F., Nagarajan, S. S., Kirsch, H. E., Barbaro, N. M., and Knight, R. T. (2007). Spatiotemporal dynamics of word processing in the human brain. *Front. Neurosci.* 1:1. doi:10.3389/neuro.01/1.1.014.2007

Cervenka, M. C., Boatman-Reich, D. F., Ward, J., Franaszczuk, P. J., and Crone, N. E. (2011). Language mapping in bilingual patients: electrocorticography and cortical stimulation during naming. *Front. Hum. Neurosci.* 5:13. doi:10.3389/fnhum.2011.00013

Chang, E. F., Edwards, E., Nagarajan, S. S., Fogelson, N., Dalal, S. S., Canolty, R. T., Kirsch, H. E., Barbaro, N. M., and Knight, R. T. (2010). Cortical spatio-temporal dynamics underlying phonological target detection in humans. *J. Cogn. Neurosci.* 23, 1437–1446.

Chauvel, P., Landre, E., Trottier, S., Vignal, J. P., Biraben, A., Devaux, B., and Bancaud, J. (1993). Electrical stimulation with intracerebral electrodes to evoke seizures. *Adv. Neurol.* 63, 115–121.

Chauvel, P., Vignal, J. P., Biraben, A., Badier, J. M., and Scarabin, J. M. (1996). "Stereoelectroencephalography," in *Focus Localization. Multimethodological Assessment of Focalization-Related Epilepsy*, eds G. Pavlik and H. Stefan (Berlin: Thomas Wiese), 135–163.

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

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.

Creutzfeldt, O., Ojemann, G., and Lettich, E. (1989). Neuronal activity in the human lateral temporal lobe I. Responses to speech. *Exp. Brain Res.* 77, 476–489.

Crone, N. E. (2000). Functional mapping with ECoG spectral analysis. *Adv. Neurol.* 84, 343–351.

Crone, N. E., and Hao, L. (2002). Functional dynamics of spoken and signed word production: a case study using electrocorticographic spectral analysis. *Aphasiology* 16, 903–926.

Crone, N. E., Hao, L., Hart, J. Jr., Boatman, D., Lesser, R. P., Irizarry, R., and Gordon, B. (2001). Electrocorticographic gamma activity during word production in spoken and sign language. *Neurology* 5, 2045–2053.

De Vos, M., Riès, S., Vanderperren, K., Vanrumste, B., Alario, F. X., Van Huffel, S., and Burle, B. (2010). Removal of muscle artifacts from eeg recordings of spoken language production. *Neuroinformatics* 8, 135–150.

DeLeon, J., Gottesman, R. F., Kleinman, J. T., Newhart, M., Davis, C., Heidler-Gary, J., Lee, A., and Hillis, A. E. (2007). Neural regions essential for distinct cognitive processes underlying picture naming. *Brain* 130, 1408–1422.

Démonet, J.-F., Thierry, G., and Cardebat, D. (2005). Renewal of the neurophysiology of language: functional neuroimaging. *Physiol. Rev.* 85, 49–95.

Donner, T. H., and Siegel, M. (2011). A framework for local cortical oscillation patterns. *Trends Cogn. Sci.* 15, 191–199.

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

Engel, A. K., Moll, C. K. E, Fried, I., and Ojemann, G. A. (2005). Invasive recordings from the human brain: clinical insights and beyond. *Nat. Rev. Neurosci.* 6, 35–47.

Flinker, A., Chang, E. F., Kirsch, E. H., Barbaro, N. M., Crone, N. E., and Knight, R. T. (2010). Single-trial speech suppression of auditory cortex activity in humans. *J. Neurosci.* 30, 16643–16650.

Forn, C., Ventura-Campos, N., Belenguer, A., Belloch, V., Parcet, M. A., and Avila, C. (2008). A comparison of brain activation patterns during covert and overt paced auditory serial addition test tasks. *Hum. Brain Mapp.* 29, 644–650.

Fukuda, M., Rothermel, R., Juhász, C., Nishida, M., Sood, S., and Asano, E. (2010). Cortical gamma-oscillations modulated by listening and overt repetition of phonemes. *Neuroimage* 49, 2735–2745.

Ganushchak, L. Y., Christoffels, I. K., and Schiller, N. O. (2011). The use of electroencephalography in language production research: a review. *Front. Psychol.* 2:208. doi:10.3389/fpsyg.2011.00208

Gaona, C. M., Sharma, M., Freudenburg, Z. V., Breshears, J. D., Bundy, D. T., Roland, J., Barbour, D. L., Schalk, G., and Leuthardt, E. C. (2011). Nonuniform high-gamma (60-500 Hz) power changes dissociate cognitive task and anatomy in human cortex. *J. Neurosci.* 31, 2091–2100.

Goldrick, M., Dell, G. S., Kroll, J., and Rapp, B. (2009). Sequential information processing and limited interaction in language production. *Science.* [E-letter].

Goncharova, I. I., McFarland, D. J., Vaughan, T. M., and Wolpaw, J. R. (2003). EMG contamination of EEG: spectral and topographical characteristics. *Clin. Neurophysiol.* 114, 1580–1593.

Halgren, E., Marinkovic, K., and Chauvel, P. (1998). Generators of the late cognitive potentials in auditory and visual oddball task. *Electroencephalogr. Clin. Neurophysiol.* 106, 156–164.

Halgren, E., Wang, C., Schomer, D. L., Knake, S., Marinkovic, K., Wu, J., and Ulberte, I. (2006). Processing stages underlying word recognition in the anteroventral temporal lobe. *Neuroimage* 30, 1401 – 1413.

Hamberger, M. J. (2007). Cortical language mapping in epilepsy: a critical review. *Neuropsychol. Rev.* 17, 477–489.

Hamberger, M. J., and Seidel, W. T. (2003). Auditory and visual naming tests: normative and patient data for accuracy, response time, and tip-of-the-tongue. *J. Int. Neuropsychol. Soc.* 9, 479–489.

Hart, J. JR., Crone, N. E., Lesser, R. L., Sieracki, J., Miglioretti, D. L., Hall, C., Sherman, D., and Gordon, B. (1998). Temporal dynamics of verbal object comprehension. *Proc. Natl. Acad. Sci. U.S.A.* 95, 6498–6503.

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

Jacobs, J., and Kahana, M. J. (2010). Direct brain recordings fuel advances in cognitive electrophysiology. *Trends Cogn. Sci.* 14, 162–171.

Jerbi, K., Ossandon, T., Hamame, C. M., Senova, S., Dalal, S. S., Jung, J., Minotti, L., Bertand, O., Berthoz, A., Kahane, P., and Lachaux, J. P. (2009). Task-related gamma-band dynamics from an intracerebral perspective: review and implications for surface EEG and MEG. *Hum. Brain Mapp.* 30, 1758–1771.

Kellis, S., Miller, K., Thomson, K., Brown, R., House, P., and Greger, B. (2010). Decoding spoken words using local field potentials recorded from the cortical surface. *J. Neural Eng.* 7, 056007.

Kielar, A., Milman, L., Bonakdarpour, B., and Thompson, C. K. (2011). Neural correlates of covert and overt production of tense and agreement morphology: evidence from fMRI. *J. Neurolinguistics* 24, 183–201.

Knapp, H. P., Corina, D. P., Panagiotides, H., Zanos, S., Brinkely, J., and Ojemann, G. (2005). "Frontal lobe representations of American sign language, speech, and reaching: data from cortical stimulation and local field potentials. Program No. 535.7," in *Neuroscience Meeting Planner*, Washington, DC: Society for Neuroscience.

Korzeniewska, A., Franaszczuk, P. J., Crainiceanu, C. M., Kus, R., and Crone, N. E. (2011). Dynamics of large-scale cortical interactions at high gamma frequencies during word production: event related causality (ERC) analysis of human electrocorticography (ECoG). *Neuroimage* 56, 2218–2237.

Liégeois-Chauvel, C., De Graaf, J. B., Laguitton, V., and Chauvel, P. (1999). Specialization of left auditory cortex for speech perception in man is based on temporal coding. *Cereb. Cortex* 9, 484–496.

Liégeois-Chauvel, C., Morin, C., Musolino, A., Bancaud, J., and Chauvel, P. (1989). Evidence for a contribution of the auditory cortex to audiospinal facilitation in man. *Brain* 112, 375–391.

Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., and Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalogr. Clin. Neuro.* 92, 204–214.

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

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.

Mainy, N., Jung, J., Baciu, M., Kahane, P., Schoendorff, B., Minotti, L., Hoffmann, D., Bertrand, O., and Lachaux, J. P. (2008). Cortical dynamics of word recognition. *Hum. Brain Mapp.* 29, 1215–1230.

Mayeux, R., Brandt, J., Rosen, J., and Benson, D. F. (1980). Interictal memory and language impairment in temporal lobe epilepsy. *Neurology* 30, 120–125.

McDonald, C. R., Thesen, T., Carlson, C., Blumberg, M., Girard, H. M., Trongnetrpunya, A., Sherfey, J. S., Devinsky, O., Kuzniecky, R., and Dolye, W. K. (2010). Multimodal imaging of repetition priming: using fMRI, MEG, and intracranial EEG to reveal spatiotemporal profiles of word processing. *Neuroimage* 53, 707–717.

McMenamin, B. W., Shackman, A. J., Maxwell, J. S., Bachhuber, D. R., Koppenhaver, A. M., Greischar, L. L., and Davidson, R. J. (2010). Validation of ICA-based myogenic artifact correction for scalp and source-localized EEG. *Neuroimage* 49, 2416–2432.

McMenamin, B. W., Shackman, A. J., Maxwell, J. S., Greischar, L. L., and Davidson, R. J. (2009). Validation of regression-based myogenic correction techniques for scalp and source-localized EEG. *Psychophysiology* 46, 578–592.

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

Müller-Preuss, P., and Ploog, D. (1981). Inhibition of auditory cortical neurons during phonation. *Brain Res.* 215, 61–76.

Ojemann, S. G., Berger, M. S., Lettich, E., and Ojemann, G. A. (2003). Localization of language function in children: results of electrical stimulation mapping. *J. Neurosurg.* 98, 465–470.

Pei, X., Leuthardt, E. C., Goanal, C. M., Brunner, P., Wolpaw, J. R., and Schalk, G. (2011). Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. *Neuroimage* 54, 2960–2972.

Postma, A., and Oomen, C. E. (2005). "Critical issues in speech monitoring," in *Phonological Encoding and Monitoring in Normal and Pathological Speech*, eds R. Hartsuiker, Y. Bastiaanse, A. Postma, and F. Wijnen (Hove: Psychology Press), 157–166.

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann. N. Y. Acad. Sci.* 119, 62–88.

Riès, S., Janssen, N., Dufau, S., Alario, F. X., and Burle, B. (2011). General-purpose monitoring during speech production. *J. Cogn. Neurosci.* 23, 1419–1436.

Roland, J., Brunner, P., Johnston, J., Schalk, G., and Leuthardt, E. C. (2010). Passive real-time identification of speech and motor cortex during an awake craniotomy. *Epilepsy Behav.* 18, 123–128.

Rosen, H. J., Ojemann, J. G., Ollinger, J. M., and Petersen, S. E. (2000). Comparison of brain activation during word retrieval done silently and aloud using fMRI. *Brain Cogn.* 42, 201–217.

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. (2007). Clinical neurophysiology of language: the MEG approach. *Clin. Neurophysiol.* 118, 237–254.

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

Schevon, C. A., Carlson, C., Zaroff, C. M., Weiner, H. J., Doyle, W. K., Miles, D., Lajoie, J., Kuzniecky, R., Pacia, S., Vazquez, B., Luciano, D., Najjar, S.,and Devinsky, O. (2007). Pediatric language mapping: sensitivity of neurostimulation and Wada testing in epilepsy surgery. *Epilepsia* 48, 539–545.

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.

Sinai, A., Bowers, C. W., Crainiceanu, C. M., Boatman, D., Gordon, B., Lesser, R. P., Lenz, F. A., and Crone, N. E. (2005). Electrocorticographic high gamma activity versus electrical cortical stimulation mapping of naming. *Brain* 128, 1556–1570.

Tanji, K., Suzuki, K., Delorme, A., Shamoto, H., and Nakasato, N. (2005). High-frequency γ-band activity in the basal temporal cortex during picture-naming and lexical-decision tasks. *J. Neurosci.* 25, 3287–3293.

Thampratankul, L., Nagasawa, T., Rothermel, R., Juhasz, C., Sood, S., and Asano, E. (2010). Cortical gamma oscillations modulated by word association tasks: intracranial recording. *Epilepsy Behav.* 18, 116–118.

Towle, V. L., Yonn, H. A., Castelle, M., Edgar, J. C., Biassou, N. M., Frim, D. M., Spire, J. P., and Kohrman, M. H. (2008). ECoG gamma activity during a language task: differentiating expressive and receptive speech areas. *Brain* 131, 2013–2027.

Trébuchon-Da Fonseca, A., Giraud, K., Badier, J. M., Chauvel, P., and Liegeois-Chauvel, C. (2005). Hemispheric lateralization of voice onset time (VOT): comparison between depth and scalp EEG recordings. *Neuroimage* 27, 1–14.

Trébuchon-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.

Tsapkini, K., Frangakis, C. E., and Hillis, A. E. (2011). The function of the left anterior temporal pole: evidence from acute stroke and infarct volume. *Brain* 134, 3094–3105.

Usui, K., Ikeda, A., Nagamine, T., Matsuhashi, M., Kinoshita, M., Mikuni, N., Miyamoto, S., Hashimoto, N., Fukuyama, H., and Shibasaki, H. (2009). Temporal dynamics of Japanese morphogram and syllabogram processing in the left basal temporal area studied by event-related potentials. *J. Clin. Neurophysiol.* 26, 160–166.

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., and Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage* 30, 1414–1432.

Wu, H. C., Nagasawa, T., Brown, E. C., Juhasz, C., Rothermel, R., Hoechstetter, K., Shah, A., Mittal, S., Fuerst, D., Sood, S., and Asano, E. (2011). Gamma-oscillations modulated by picture naming and word reading: intracranial recording in epileptic patients. *Clin. Neurophysiol.* 122, 1929–1942.

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

Received: 08 July 2011; accepted: 28 November 2011; published online: 27 December 2011.

Citation: Llorens A, Trébuchon A, Liégeois-Chauvel C and Alario F-X (2011) Intra-cranial recordings of brain activity during language production. Front. Psychology ***2****:375. doi: 10.3389/fpsyg.2011.00375*

This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 11 October 2011
doi: 10.3389/fpsyg.2011.00239

Examining the central and peripheral processes of written word production through meta-analysis

Jeremy J. Purcell[1], Peter E. Turkeltaub[2], Guinevere F. Eden[1] and Brenda Rapp[3]*

[1] *Department of Pediatrics, Center for the Study of Learning, Georgetown University, Washington, DC, USA*
[2] *Department of Neurology, Georgetown University, Washington, DC, USA*
[3] *Department of Cognitive Science, Johns Hopkins University, Baltimore, MD, USA*

Edited by:
Albert Costa, University Pompeu Fabra, Spain

Reviewed by:
Pelagie M. Beeson, University of Arizona, USA
Steven Z. Rapcsak, University of Arizona, USA

****Correspondence:***
Brenda Rapp, Department of Cognitive Science, Johns Hopkins University, 135 Krieger Hall, 3400 North Charles Street, Baltimore, MD 21218-2685, USA.
e-mail: rapp@cogsci.jhu.edu

Producing written words requires "central" cognitive processes (such as orthographic long-term and working memory) as well as more peripheral processes responsible for generating the motor actions needed for producing written words in a variety of formats (handwriting, typing, etc.). In recent years, various functional neuroimaging studies have examined the neural substrates underlying the central and peripheral processes of written word production. This study provides the first quantitative meta-analysis of these studies by applying activation likelihood estimation (ALE) methods (Turkeltaub et al., 2002). For alphabet languages, we identified 11 studies (with a total of 17 experimental contrasts) that had been designed to isolate central and/or peripheral processes of word spelling (total number of participants = 146). Three ALE meta-analyses were carried out. One involved the complete set of 17 contrasts; two others were applied to subsets of contrasts to distinguish the neural substrates of central from peripheral processes. These analyses identified a network of brain regions reliably associated with the central and peripheral processes of word spelling. Among the many significant results, is the finding that the regions with the greatest correspondence across studies were in the left inferior temporal/fusiform gyri and left inferior frontal gyrus. Furthermore, although the angular gyrus (AG) has traditionally been identified as a key site within the written word production network, none of the meta-analyses found it to be a consistent site of activation, identifying instead a region just superior/medial to the left AG in the left posterior intraparietal sulcus. These meta-analyses and the discussion of results provide a valuable foundation upon which future studies that examine the neural basis of written word production can build.

Keywords: spelling, writing, meta-analysis, fMRI, dysgraphia, fusiform gyrus, intraparietal sulcus, angular

INTRODUCTION

There are a number of reasons why it is important to understand the neural substrates of written language production. Communicating through written language is critically important to professional success and for effective functioning in everyday life (e.g., writing checks, keeping a calendar, taking messages, etc.). In this regard, its importance has only increased with the rise of electronic communication (e-mail, internet, texting, etc.). As a result, deficits of written communication have a very significant impact on the well-being of individuals who suffer from acquired and developmental dysgraphia. Understanding the neural substrates of written language production is important for developing accurate prognoses and effective remediation of these written language impairments. Furthermore, written language processing is an interesting domain from a basic neurobiological perspective. Written language is a relatively recent human invention, appearing approximately 5000 years ago and used by only a limited portion of the human population until very recently. As a result, it is unlikely to have had an impact on the human genome and, accordingly, there is unlikely to be a genetic blueprint for the specific neural circuitry of written language processing. Nonetheless, with instruction, most people learn to comprehend and produce written language with remarkable ease. Thus, written language offers an opportunity to investigate the brain's capacity to develop expertise in skills that are not specifically biologically predetermined.

While in the past two decades there has been a great deal of functional neuroimaging research directed at understanding the brain-basis of written language comprehension (reading), relatively little attention has been directed at investigating written language production (spelling and writing). Recently, however there has been an upswing in the number of functional neuroimaging investigations in this domain. The findings from these studies, along with those from the more traditional clinical literature examining correlations between lesions and deficits, have provided important insights into the neurobiology of written language production. The neuroimaging studies, quite naturally, differ with regard to a number of variables such as experimental and control tasks, neuroimaging modalities, etc. This heterogeneity, as well as the current critical mass of functional neuroimaging studies of spelling, makes this an appropriate moment to attempt to integrate findings across studies. In this paper, we report on our efforts to do so by carrying out a meta-analysis of existing positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) studies of word spelling in alphabetic language involving adult participants.

Producing written words involves a number of interacting cognitive processes that have been described in various models of written language production (Roeltgen and Heilman, 1985; Rapp and Caramazza, 1997; Rapcsak and Beeson, 2002; Hillis and Rapp, 2004). Although these cognitive processes are highly integrated, an important distinction is often made between central and peripheral components (see **Figure 1**). The different patterns of impairment that have been observed in cases of acquired dysgraphia subsequent to brain lesions have constituted the major source of empirical support for the distinctions between central and peripheral processing components as well as for the more fine-grained distinctions described below and depicted in **Figure 1**. In addition, convergent evidence for many of these distinctions has been confirmed by behavioral studies of spelling and writing in neurologically healthy participants. While it is outside the scope of this paper to review these literatures, we refer the interested reader to various reviews (Ellis, 1979; Burt and Fury, 2000; Burt and Tate, 2002; Weingarten, 2005).

Spelling typically begins by hearing words (e.g., taking notes in a lecture, a message over the phone, etc.) or with internally generated word meanings (e.g., writing a letter, a grocery list, etc.). These auditory comprehension and semantic processes and mechanisms are not specific to spelling, yet serve as the basis for the subsequent retrieval or assembly of spellings. Spelling-specific, central processes are usually identified as: orthographic long-term memory (O-LTM; the orthographic lexicon), phoneme–grapheme (PG) conversion, and orthographic working memory (the graphemic buffer). O-LTM is the store of the word spellings that an individual is familiar with. As indicated in **Figure 1**, information in O-LTM may be retrieved on the basis of a word's meaning or, according to some researchers, directly from a representation of the word's sound (Patterson, 1986). In addition to retrieval from O-LTM, word spellings may be assembled from a phonological stimulus via the PG conversion processes that apply learned information regarding the relationships between sounds and letters (or other sub-lexical units) to generate plausible spellings for sound strings. For example, the sound stimulus "wuns" could result in the retrieval of the information O-N-C-E from O-LTM and/or in the assembly of a plausible spelling such as W-U-N-S-E from the PG conversion system. The letter representations assembled or retrieved are assumed to be abstract, lacking format-specific information (such as shape, size, motor plan, etc.). The abstract letter strings are then processed by O-WM, a limited capacity system responsible for maintaining letter identity and order information active so that they can be selected for further processing by peripheral components (Rapp and Kong, 2002; Kan et al., 2006). These central processes interact with one another, with evidence specifically supporting bi-directional interactions between O-WM and O-LTM (McCloskey et al., 2006) and between O-LTM and PG conversion processes (Rapp et al., 2002).

In terms of peripheral processes, it is generally assumed that there are multiple stages involved in going from the abstract letters representations in O-WM to the correct ordering and execution of the effector-specific muscle movements required for expressing

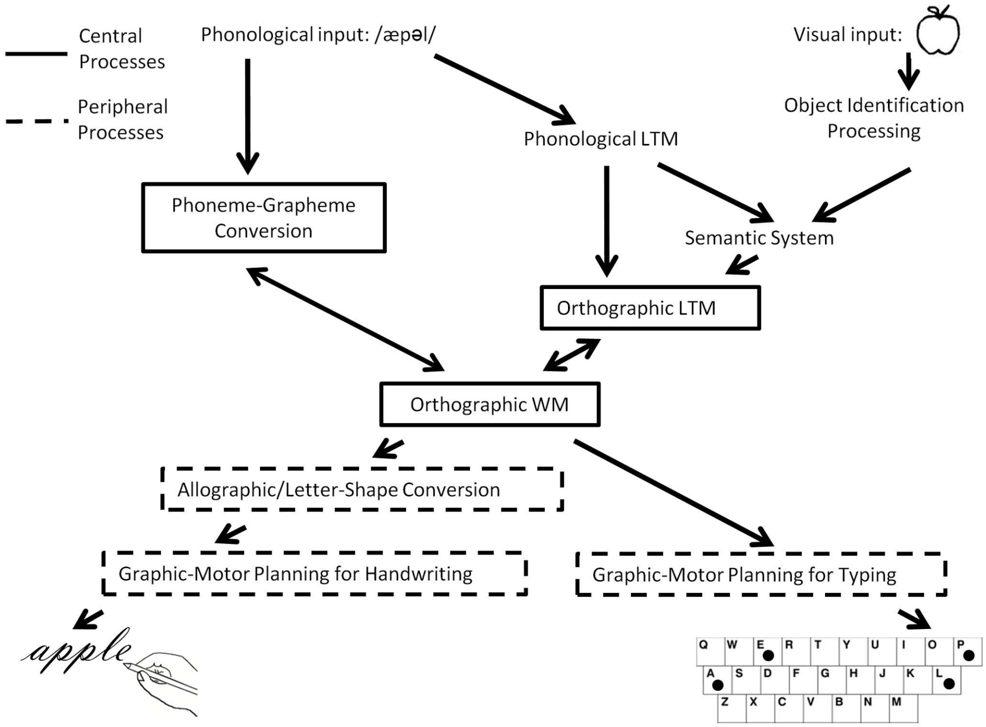


FIGURE 1 | A schematic depiction of the cognitive architecture of the written word production system.

these letters. These peripheral processes generate written language in the major modalities of oral spelling, written spelling, or typing. As of yet there is no strong consensus on the characterization of the multiple peripheral processes, and so we present a general overview of some of the basic components (Ellis, 1979, 1982; Margolin, 1984; Ellis and Young, 1988; Rapp and Caramazza, 1997).

While we will not be concerned with oral spelling in this paper (as there have been no functional neuroimaging studies of oral spelling), it is worth mentioning the peripheral processes that are involved in this spelling "format." On the basis of the information held in O-WM, the names of letters (e.g., /si//ei//ti/) are retrieved in their correct order from the phonological LTM store of words (the phonological lexicon). Subsequently, the phonetic and articulatory planning processes of speech are recruited for the spoken production of the letter names.

In the case of writing, the specific letter forms or allographs must first be selected (upper-case, lower case, print, cursive, etc.). The mechanism responsible for converting the abstract letters held in O-WM to specific letter shapes is often referred to as Allographic or Letter-Shape Conversion. Following Allographic Conversion, motor plans for producing the letter forms are then assembled by Graphic Motor Planning processes, specifying such things as the size and ordering of the strokes. Subsequently, Effector-Specific Motor Programming compiles instructions for the specific limb/s (e.g., right or left hand, foot, etc.) to be used in carrying out the motor actions. The programming and execution of effector-specific actions is subject to multi-modal feedback during execution (haptic, tactile, visual, etc.). In addition, some researchers have posited buffers/WM components that are associated with motor planning and programming components (Ellis and Young, 1988). Furthermore, the various motor planning processes necessarily interact with others responsible for certain spatial aspects of the writing process such as the coordination of the ego-centric and allo-centric reference frames required for producing the specific motor acts in the correct spatial locations.

In contrast to writing, typing requires the generation of a series of hand and finger motor commands in standard keyboard space in order to generate the letters held in O-WM. If we assume an experienced typist, motor plans may be stored and/or pre-compiled not only for individual letters but also for larger graphemic chunks, allowing for rapid production of letter sequences. Motor planning in typing is sensitive to the particular sequencing of letters, with the planning of the timing of hand movements for multiple letters typically occurring in parallel. As a result, the movements required to produce each letter are modulated by the immediately preceding and following context; this is sometimes referred to as co-articulation (West and Sabban, 1982; Salthouse, 1986; Salthouse and Saults, 1987). The motor programs used in typing also require parameterization for such things as speed of typing, size of keyboard, etc. As is the case for writing, the motor planning and execution processes required for typing also interact with processes that ensure the coordination of the spatial frames of reference required to move the hands and fingers in external keyboard space.

Handwriting and typing face a number of similar computational demands and, as a result, possibly share processing components. For example, motor sequences for both typing and handwriting have been shown to be similarly influenced (in terms of timing of strokes and pauses) by the types of grapheme units being produced (digraphs, double letters, syllable, and morpheme boundaries; Weingarten et al., 2004). Some investigators have even posited that handwriting and typing may share buffering or working memory components (Magrassi et al., 2010). For example, recent work has reported similar impairments to handwritten and typed production, but not oral spelling, subsequent to electrical stimulation to a restricted area of the superior parietal lobe (SPL; Magrassi et al., 2010).

Turning to the neural correlates of these central and peripheral components, findings from numerous studies of acquired dysgraphia subsequent to neural injury (typically stroke, but also surgical resection and progressive neurological disease) consistently indicate that the central processes of spelling are subserved by a network of regions including the left fusiform/inferior temporal, middle temporal, superior temporal, inferior frontal, angular, and supramarginal gyri (Beauvois and Derouesne, 1981; Shallice, 1981; Roeltgen and Heilman, 1984; Kawahata et al., 1988; Rapcsak and Beeson, 2004; Philipose et al., 2007; Rapcsak et al., 2009; Tsapkini and Rapp, 2010). In the lesion-based literature, peripheral processes have been most consistently associated with a left frontal–parietal network including, but not limited to, the dorsal premotor cortex and SPL (Exner, 1881; Ritaccio et al., 1992; Tohgi et al., 1995; Lubrano et al., 2004; Roux et al., 2009). In addition, peripheral dysgraphias have been described that are associated with damage to neural regions involved in general motor processing (not specifically in written language production) such as motor cortex, cerebellum, and various sub-cortical structures, e.g., caudate, putamen, and thalamus (Tanridag and Kirshner, 1985; Pramstaller and Marsden, 1996; Denes et al., 2005; Fournier Del Castillo et al., 2011). Finally, it is worth mentioning that spatial dysgraphias have been associated with some forms of spatial neglect (e.g., Caramazza and Hillis, 1990).

Although the lesion literature has provided critical information regarding the necessary substrates for written language production, it is subject to the well-known challenges of large and uncontrolled lesions and concerns regarding functional reorganization. In particular, precise localizations are typically quite difficult. Thus, lesion studies are well-complemented by functional neuroimaging techniques in non-impaired, healthy individuals. This effort has been strengthened in recent years by numerous fMRI or PET studies that have examined all or some component processes of written language production.

In general, functional neuroimaging studies of written production can be grouped into two types; those that make use of tasks that involve only central processing and those that make use of tasks that combine both central and peripheral processing. For instance the "central-only" tasks are those that require subjects to access the spelling of a word, but which do not involve writing (or typing). The "central + peripheral" tasks on the other hand, are those that require subjects to write or type words. In studies using central + peripheral tasks, central and peripheral processing components can sometimes be distinguished from one another depending on the additional experimental conditions employed and the specific contrasts they afford. We discuss these issues is more detail in the Section "Methods."

With regard to central processes, the neuroimaging literature has generally supported the findings of the lesion literature that the left fusiform/inferior temporal gyrus (ITG), inferior frontal gyrus (IFG), and supramarginal gyrus (SMG) are associated with the central processes of spelling. As would be expected, activations in these regions have been reported for tasks involving either only central processing (Booth et al., 2002, 2004; Rapp and Lipka, 2011), or both central and peripheral processing (Beeson et al., 2003; Purcell et al., 2011). Interestingly, the few neuroimaging studies that have examined both reading and spelling within the same study have identified overlapping activations in both the left fusiform/ITG and IFG, suggesting that the these substrates in particular are important for central processes shared by both spelling and reading (Purcell et al., 2011; Rapp and Lipka, 2011). In addition the left SMG has also been associated with central processes (Petrides et al., 1995; Booth et al., 2002; Beeson et al., 2003; Purcell et al., 2011). Specifically, it has primarily been associated with sub-lexical (PG conversion) processing. This association is supported by the observation of SMG activation in studies of Japanese Kana writing, which is thought to rely heavily on PG conversion processing (Katanoda et al., 2001; Sugihara et al., 2006).

One apparent contradiction between lesion and functional neuroimaging studies concerns the left angular gyrus (AG). Although the lesion literature has long suggested it plays a role in central processes of written production, this region has not been consistently identified in the neuroimaging literature. While some studies have reported activation in this region (Booth et al., 2002, 2003a), a number of others have not (Beeson et al., 2003; Purcell et al., 2011). Furthermore, some of the latter studies have reported clusters in the intraparietal sulcus (IPS) region superior and medial to the left AG (Beeson et al., 2003; Purcell et al., 2011). Given the heterogeneity of these findings, the role of the posterior parietal cortex in written language production, and in particular with regard to the central processes, merits more careful scrutiny.

With regard to the peripheral components of written language production, in the neuroimaging literature these have primarily been associated with a left dorsal premotor/superior parietal network, as well as with activity in the cerebellum and sub-cortical structures such as the basal ganglia and thalamus (Menon and Desmond, 2001; Beeson et al., 2003; Sugihara et al., 2006; Roux et al., 2009; Purcell et al., 2011). The left dorsal premotor cortex in particular has been considered relevant to writing-specific processes involving the generation of graphemic motor commands (Menon and Desmond, 2001; Beeson et al., 2003). Although some of the earlier neuroimaging work in this field suggested that the left dorsal premotor region was associated with the generation of handwriting-specific graphic motor commands (Beeson et al., 2003), recent findings indicate that this general region also plays a role in the generation of graphic motor commands of typed production (Purcell et al., 2011).

Although a qualitative narrative summary such as the one presented above is valuable, it does not allow for a precise localization of the shared activations reported across studies. Meta-analytic methods allow us to address these challenges by quantitatively identifying brain locations that are consistently associated with tasks or cognitive functions of interest. Therefore, we applied the activation likelihood estimation (ALE) technique (GingerALE 2.1a3, BrainMap.org) to the study of written word production. The ALE technique is a widely used, validated, automated, quantitative method for a voxel-wise meta-analysis of neuroimaging foci which has been used in a range of cognitive domains such as reading (Turkeltaub et al., 2002), speech perception (Turkeltaub and Coslett, 2011), and object naming (Price et al., 2005).

Briefly, the goal of the ALE approach is to estimate, for each voxel in a normalized brain, the likelihood that it corresponds to the peak of a significant cluster in a task/contrast of interest. The logic underlying the approach is that, although significant activations are reported as discrete *X*, *Y*, *Z* locations, there is uncertainty regarding their precise location. This uncertainty can be modeled as a three-dimensional Gaussian probability density distribution around the activation peaks that have been reported for a study. By combining the probability distributions corresponding to all of the significant activation peaks from all of the contributing studies, and then applying appropriate statistical corrections and thresholds, the ALE algorithm estimates the likelihood that a voxel corresponds to a location of peak activity in the literature. This analysis yields "clusters" of significant activation likelihood estimates that represent the spatial overlap of peak activity among the contributing studies.

In this paper, we report on the results of a series of meta-analyses. In the first, we applied the ALE algorithm to the findings of 11 written language production neuroimaging studies with a combined total of 17 separate contrasts. We then analyzed two subsets of the contrasts separately to identify central-only components of the spelling process and central + peripheral components. Finally, we compared the results of central + peripheral to central-only ALE analyses in order to identify neural substrates that are reliably associated with the peripheral processes of written production. In combination, this set of analyses allowed us to identify the brain regions that are most reliably associated with central and peripheral written language production processes in alphabetic writing.

METHODS

SELECTION OF STUDIES

We searched Pubmed and Googlescholar online databases for studies associated with written language production using keywords "writing," "handwriting," "spelling," "orthographic," "fMRI," "PET," and "neuroimaging" in relevant combinations. Reference lists for appropriate publications were also searched for additional studies that could be included. Direct e-mail communication with some researchers also provided additional data sets for analysis.

We included studies based on the following inclusion criteria: (1) the neuroimaging technique used was fMRI or PET; (2) subjects were neurologically healthy, right-handed adults; (3) experiments required participants to generate orthographic lexical and/or sub-lexical representations; (4) studies involved an alphabetic written language; (5) no visual word or non-word stimuli were presented during the task of interest (in order to allow us to more clearly distinguish spelling from reading processes); (6) results were reported in a stereotactic three-dimensional coordinate system; (7) the field of view covered the whole brain; (8) analyses were not restricted to regions of interest; (9) a random-effects analysis was performed.

SELECTION AND CATEGORIZATION OF CONTRASTS

We assigned the selected studies to either central-only (C) or central and peripheral (C + P) categories. More precisely, we placed the selected contrasts into one of these two categories. This ensured that the meta-analyses would be based not simply on the characteristics of the experimental tasks, but rather on the cognitive processes that were "isolated" on the basis of the contrast performed. The C-contrasts were those that predominantly isolated central processing components, whereas the C + P contrasts were those that included both central and substantive peripheral processes.

An example of a C contrast is the one reported in Rapp and Lipka (2011). In this study, the experimental task required subjects to view a letter, listen to an auditorily presented word and then respond with a button press if the visually presented letter was in the spelling of the auditorily presented word. This task requires recruitment of O-LTM, and possibly PG conversion, as well as O-WM, but does not engage peripheral spelling processes. The contrasting (baseline) task was a perceptual-motor control task that allowed for the visual, auditory, motor, and decision components of the experimental task to be "subtracted out" in the contrast. An example of a C + P contrast comes from Beeson et al. (2003). In this study, the experimental task required subjects to think of and handwrite words from a specific category (e.g., fruits) during a fixed time period. The baseline task was a circle-drawing task that served to exclude the most general peripheral motor aspects of handwriting, while retaining not only the central processes but also many of the writing-specific peripheral components.

Some of the studies that met the inclusion criteria for the meta-analysis included multiple contrasts. For example, the Beeson et al. (2003) study, in addition to a word writing task, included a circle-drawing task as well as an alphabet writing task. For the latter task, subjects had to write as much of the alphabet as possible during a limited time period (comparable to the time period for writing words). While the word writing vs. circle-drawing contrast provided information regarding both central + peripheral processes, the word writing vs. alphabet writing contrast served to more specifically isolate central processes. In this way, some studies contributed multiple sets of contrasts and their corresponding results to the meta-analysis.

ALE METHODS

The *X*, *Y*, *Z* coordinates of every significant peak (or subpeak) for all eligible contrasts constituted the input to the meta-analysis. Coordinates that were reported in Talairach space (Talairach and Tournoux, 1988) were converted to Montreal Neurological Institute (MNI) coordinates (Lancaster et al., 2007). The ALE meta-analysis was implemented using GingerALE 2.1a3 (www.brainmap.org). A new ALE algorithm was employed which consists of three steps resulting in an ALE map that is unbiased by the number of foci or the number of contrasts included from each study (Turkeltaub et al., 2011).

The three steps of the ALE analyses are as follows: (1) localization uncertainty is modeled for each focus of activity as a Gaussian, the width of which is determined from the number of subjects of the study (Eickhoff et al., 2005); (2) taking the union of the study-specific localization probabilities identified for each voxel yields the voxel-wise ALE value; (3) significance is tested using a random-effects method with a null hypothesis that the location of activation in each study is independent of the others (Eickhoff et al., 2009). The ALE maps generated according to steps 1–3 were thresholded with a False Discovery Rate of $q > 0.05$ and clusters size $>100\,\text{mm}^3$. For each of the resulting significant clusters, we additionally considered the number of studies that contributed to each. We did so by tallying the studies with foci within 2 SD of localization uncertainty (see Eickhoff et al., 2009). In order to ensure that the reported results represented coherence across multiple experiments, we eliminated ALE clusters and peaks that were based on fewer than three different contrasts.

Three ALE analyses were carried out: the "All-Contrasts" analysis included all of the eligible contrasts, the "C-only" analysis included all eligible C-contrasts and the "C + P" analysis included all of the eligible C + P contrasts. Additionally, we performed a subtraction analysis to identify locations where the C + P ALE values were significantly greater than C-only ALE values, thus providing a means for identifying regions associated with peripheral processing. This subtraction analysis simply subtracts two ALE maps generated from two different groups of studies. Significance is tested via a permutation method which reassigns the studies randomly into two groups of the same size as the original ones over 5000 iterations. The distribution of ALE values in these random ALE subtraction maps provides a null hypothesis for the significance test.

Visualizations of the results were implemented with MRIcron, using the Colin brain template in MNI space (Holmes et al., 1998). Surface renderings are maximum intensity projections with a search depth of 16 mm. Gyral anatomical labels were assigned based on the automated anatomical labeling (AAL) atlas developed for SPM (Tzourio-Mazoyer et al., 2002). Brodmann areas (BAs) were identified from the template developed for MRIcron. Activation likelihood peaks which were found deep in the gray matter at the sulcus boundaries between two gyri were assigned to the appropriate sulcus name. For instance, if the maximum probability label at a given peak was the SPL and the corresponding cluster was located deep in the gray matter of the IPS (which defines the boundary of the AG and the SPL), then a label of superior parietal lobule/intraparietal sulcus (SPL/IPS) was used in the tables and in the text.

RESULTS

IDENTIFICATION OF ELIGIBLE STUDIES/CONTRASTS

A total of 17 experimental contrasts were identified in 11 different publications which met our inclusion and exclusion criteria for the All-Contrasts meta-analysis and together represented a total of 146 subjects. All but one study was conducted using fMRI and all employed block designs except for one which employed an event-related design (Cohen et al., 2004). Most studies were conducted with English speaking participants, with two of the 11 studies performed with French native speakers. The eligible contrasts with a brief description of the tasks, the number of subjects, the language and imaging modality are listed in **Table 1**.

ANALYSIS 1: ALL-CONTRASTS

As reported in **Table 2** and **Figure 2A**, the ALE analysis identified 16 significant clusters with their corresponding peaks, as well as six additional significant sub-peaks. The peaks correspond to the voxel with the highest ALE value within a cluster of significant ALE values. Sub-peaks are the non-maximal peaks within clusters that have higher ALE values than their surrounding voxels. It is important to emphasize that the results reported in the table and figure do not correspond to the activation maps that are typically reported in functional neuroimaging. Instead, all significant voxels identified by the ALE analysis correspond to locations where, considering the peak coordinates identified by the set of eligible studies, an activation peak is likely to occur.

The ALE locations identified in the first analysis were primarily in the left hemisphere. In the left frontal lobe, the analysis identified four ALE clusters with the peaks located in the left posterior IFG, the superior frontal gyrus/sulcus, the supplementary

Table 1 | Meta-analysis dataset.

Reference	MRI/PET	Language	*N*	Contrast descriptions		Contrast	
				Experimental condition	Control condition(s)	ID*	
Petrides et al. (1995)	PET	English	11	Handwriting: write spoken words	Recall: say aloud previously memorized words	1	C + P
Booth et al. (2002)	MRI	English	13	Spelling: hear three words, determine if third word has same rime spelling as either of the first two (e.g., "hope," "colt," "soap" or "hold," "plant," "cold")	Rhyming: hear three words, determine if third word rhymes with either of first two (e.g., "has," "last," "jazz")	2a	C
					Tones: hear three tones, determine if pitch of the third tone matches either of the first two	2b	C
Booth et al. (2003b)	MRI	English	15	Spelling: same as in Booth et al. (2002)	Tones: same as in Booth et al. (2002)	3	C
Beeson et al. (2003)	MRI	English	12	Handwriting: write words from a given semantic category (e.g., animals)	Drawing: draw continuous circles	4a	C + P
					Handwriting letters: write alphabet	4b	C
					Naming: subvocally name items from a semantic category (e.g., animals)	4c	C + P
Booth et al. (2004)	MRI	English	16	Spelling: same as in Booth et al. (2002)	Tones: same as in Booth et al. (2002)	5	C
Cohen et al. (2004)	MRI	French	17	Orthographic task: determine if the spelling of a spoken word has a descending letter (e.g., "p")	Phoneme task: determine if a spoken word contains a pre-specified phoneme	6	C
Cho et al. (2009)/e-mail	MRI	English	15	Handwriting: write names of pictures	Checkerboard: look at visual checkerboard pattern	7a	C + P
					Copy scribbles: copy images by circular, vertical, or horizontal drawing motions	7b	C + P
					Picture naming: subvocally name pictures	7c	C + P
Roux et al. (2009)/e-mail	MRI	French	12	Handwriting: write spoken words	Repeating: subvocally repeat spoken words	8a	C + P
					Rest	8b	C + P
Purcell et al. (2011)	MRI	English	17	Typing: type spoken words	Motor: hear word "motor" then type pre-practiced sequence, i.e., a;sldkfj on QWERTY keyboard	9	C + P
Rapp and Lipka (2011)	MRI	English	10	Spelling: press button if visually presented letter is in the spelling of a spoken word	Case verification: press button if visually presented letter is upper/lower case; spoken word is ignored	10	C
Rapp and Dufor (2011)	MRI	English	8	Handwriting: write spoken words	Drawing: draw continuous circles	11	C + P

**The contrast ID letters designate tasks associated with each study. The C + P and C label denotes whether the task involved was either a central + peripheral (C + P) or central-only (C) task.*

Table 2 | Results for the all-contrasts written production meta-analysis.

Location (Brodmann area)		Extent (mm^3)	Peak MNI			Peak TAL			ALE value ($\times 10^{-3}$)	Contrasts contributing to peaks
			X	*Y*	*Z*	*X*	*Y*	*Z*		
FRONTAL										
Left	IFG (9)	3456	−46	16	18	−44	12	21	19.3	2a, 3, 4b, 5
	IFG (9)	–	−44	6	28	−42	2	29	15.0	2b, 9, 10, 11
Left	SFG/SFS (6)	3720	−22	−8	54	−22	−14	51	20.2	4b, 8a, 9, 11
	SMA (6)	–	−4	−10	50	−5	−15	48	15.7	1, 4c, 7a, 11
Left	SMA (6)	1544	−4	22	46	−5	15	47	20.3	2a, 4a, 4b, 5, 9
Left	Precentral gyrus (4)	1120	−36	−24	60	−35	−29	55	15.2	1, 7a, 7c, 8a, 8b
	Postcentral gyrus (3)	–	−30	−24	48	−29	−28	44	11.3	4b, 7a, 7c
Right	Insula (13)	1144	42	24	−4	38	21	3	15.2	4a, 4b, 7b, 9
	Insula (13)	–	36	26	−8	32	23	0	13.6	4a, 4b, 5, 7b
PARIETAL										
Left	SPL/IPS (7)	1128	−32	−46	58	−31	−50	51	12.8	4c, 7a, 7c
Left	SPL/IPS (7)	2840	−30	−60	46	−29	−62	39	25.0	2a, 3, 4a, 4b, 9
Left	SMG (40)	704	−52	−32	34	−50	−34	31	14.4	1, 2b, 7c
TEMPORAL										
Left	STG/STS (21)	1512	−60	−12	−2	−57	−12	0	29.5	2a, 3, 5, 9, 11
Right	STG/STS (21)	728	52	−12	−6	47	−13	−2	17.3	2a, 3, 5, 11
OCCIPITOTEMPORAL										
Left	ITG (37)	4024	−50	−60	−16	−47	−56	−16	26.5	2a, 2b, 3, 4a, 4b, 7a, 9, 10
	FG (37)	–	−44	−56	−12	−42	−53	−12	22.9	2b, 4a, 4b, 6, 7a, 7b, 9, 10
	ITG (37)	–	−42	−58	−6	−40	−55	−7	20.3	2b, 4a, 6, 7a, 7b, 11
SUB-CORTICAL										
Left	Thalamus	744	−14	−10	10	−14	−12	12	13.0	7a, 7c, 11
Left	Putamen	576	−24	−4	−2	−23	−5	2	15.6	4a, 7a, 7c
CEREBELLUM										
Right	Cerebellum	552	14	−48	−24	12	−44	−21	16.3	1, 4c, 7a, 7c
Right	Cerebellum	464	30	−68	−30	27	−63	−28	15.4	2a, 5, 7b
Right	Cerebellum	448	8	−72	−18	6	−67	−18	11.8	7b, 9, 11

motor area (SMA), and the fourth straddling the pre and postcentral gyri. In the parietal lobe there were three clusters: two were located in the superior parietal lobule and one in the SMG. In the left temporal lobe there was one cluster in the middle portion of the superior temporal gyrus/sulcus. In the occipitotemporal region there was one large cluster that included peaks in the ITG as well as one in the mid-fusiform gyrus (FG). In addition to these cortical locations in the left hemisphere, sub-cortical ALE clusters were identified in the thalamus and putamen. In the right hemisphere, there was a significant cluster in the insula, the posterior superior temporal gyrus/sulcus as well as three clusters in the right cerebellum.

ANALYSIS 2: CENTRAL (C)-CONTRASTS

For the C-contrasts analysis, a total of seven contrasts were identified which met our inclusion and exclusion criteria (**Table 1**). These contrasts were derived from six different publications and represent a total of 83 subjects.

As reported in **Table 3** and **Figure 2B**, the C-contrasts ALE analysis identified six significant clusters with their corresponding peaks, and one additional sub-peak. All but one of the clusters were in the left hemisphere, and all were cortical. In the left frontal lobe there were two clusters: one large cluster in the posterior IFG and another in the anterior cingulate cortex. In the parietal lobe there was a single cluster in the superior parietal lobule/intraparietal sulcus. In the temporal lobe there was one cluster in the left superior temporal gyrus/superior temporal sulcus (STG/STS). Finally, in the occipitotemporal area there was a large cluster in ITG that extended into the FG. In the right hemisphere, only one ALE cluster was identified in the superior temporal gyrus/sulcus.

ANALYSIS 3: CENTRAL + PERIPHERAL (C + P) CONTRASTS

For the C + P Contrasts analysis, a total of 10 experimental contrasts were identified which met our inclusion and exclusion criteria (**Table 1**). These contrasts were derived from 6 different papers and included a total of 75 subjects.

As reported in **Table 3** and **Figure 2C**, the ALE analysis identified eight significant clusters with their corresponding peaks, and five additional sub-peaks. Clusters were primarily in the left hemisphere, with only three right hemisphere clusters. In the left frontal lobe there was one cluster in the superior frontal gyrus/sulcus

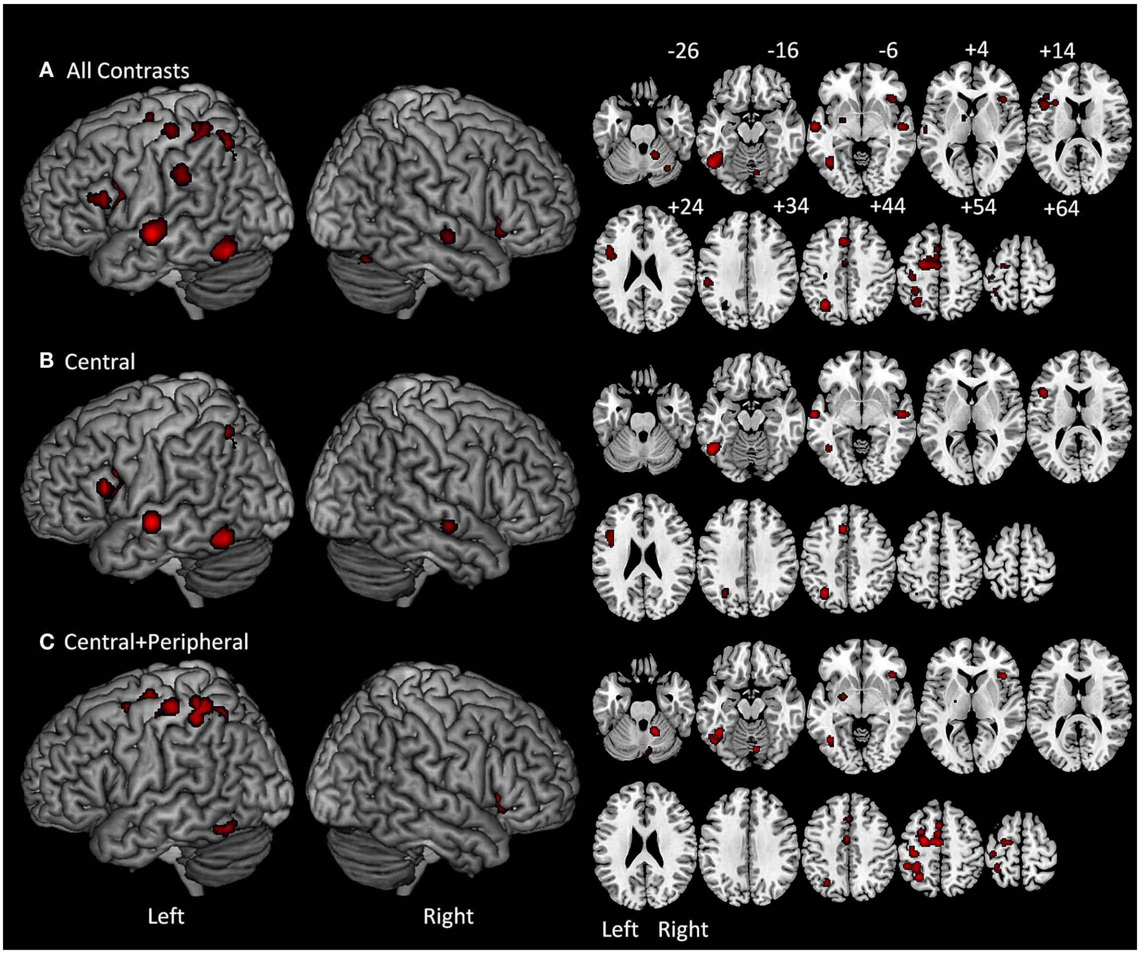


FIGURE 2 | On the left are ALE meta-analysis results projected on a standard rendered template brain and on the right are the corresponding axial slices from z −26 to +64 in 10 mm increments. (A) All-contrasts meta-analysis included all of the tasks referenced in **Table 1**. **(B)** Central-contrasts meta-analysis included only the tasks that involved central, but not peripheral written production processes. **(C)** Central + peripheral contrast meta-analysis included all of the tasks that involved both central and peripheral written production processes.

that extended into the SMA and middle frontal gyrus. Additionally, there was a cluster in the left precentral gyrus that extended posteriorly into the postcentral gyrus. In the parietal lobe there was a large cluster in the superior parietal lobule/intraparietal sulcus. In the occipitotemporal region there was a large cluster within the FG. Sub-cortically, there was a cluster in the left putamen. In the right hemisphere, two clusters were identified in the cerebellum.

ANALYSIS 4: COMPARISON OF C + P VS. C

The subtraction analysis allowed us to identify locations that were significant for the C + P but not the C analysis. The results revealed three significant clusters with their corresponding peaks and three additional sub-peaks. One cluster was in the left precentral gyrus and extended into the superior frontal sulcus. A second cluster was just posterior to the first, but in the postcentral gyrus. The third cluster was in the left anterior IPS extending dorsally into the superior parietal lobule and ventrally into the SMG (BA40). (**Table 4**; **Figure 3**).

DISCUSSION

This study provides the first quantitative meta-analysis of functional neuroimaging studies of written word production. The analysis is based on 11 studies of alphabetic languages (total $N = 146$) consisting of a total of 17 experimental contrasts that had been designed to isolate the Central or Central + Peripheral processes of word spelling. Three ALE analyses were carried out. One involved the complete set of 17 contrasts; two others were applied to subsets of contrasts with the goal of distinguishing neural substrates supporting Central from Peripheral processes. In addition a subtraction analysis was performed in order to directly contrast the two subset analyses to identify areas associated specifically with peripheral processes of written word production.

The ALE analysis that considered the full set of 17 contrasts identified significant activation locations in a distributed set of primarily left hemisphere regions that have been previously associated with spelling such as the IFG, superior frontal gyrus/superior frontal sulcus (SFG/SFS), SPL/IPS,

Table 3 | Results for the central and the central + peripheral ALE meta-analyses.

Location (Brodmann area)		Extent (mm^3)	Peak MNI			Peak TAL			ALE value (×10^{-3})	Contrasts contributing to peaks
			X	*Y*	*Z*	*X*	*Y*	*Z*		
CENTRAL-CONTRASTS										
Frontal										
Left	IFG (9)	2064	−46	16	18	−44	12	21	18.8	2a, 3, 4b, 5
Left	ACC (32)	920	−6	22	44	−7	15	45	17.5	2a, 4b, 5
Parietal										
Left	SPL/IPS (7)	1872	−30	−62	44	−29	−63	37	17.9	2a, 3, 4b
Temporal										
Left	STG/STS (22)	976	−62	−12	−3	−58	−12	−1	20.8	2a, 3, 5
Right	STG/STS (22)	872	52	−12	−6	47	−13	−2	15.7	2a, 3, 5
Occipitotemporal										
Left	ITG (37)	2360	−52	−58	−16	−49	−54	−16	18.6	2a, 2b, 3, 4b, 10
	ITG (37)	–	−46	−56	−12	−44	−53	−12	16.5	2a, 2b, 3, 4b, 6, 10
CENTRAL + PERIPHERAL CONTRASTS										
Frontal										
Left	SFG/SFS (6)	3992	−22	−8	54	−22	−14	51	18.1	8a, 9, 11
	SMA (6)	–	−4	−10	50	−5	−15	48	15.7	1, 4c, 7a, 11
Left	SMA (6)	–	−14	−10	60	−15	−16	57	11.6	4a, 7a, 7c
	Precentral gyrus (3)	1552	−36	−24	60	−35	−29	55	15.2	1, 7a, 7c, 8a, 8b
Right	Insula (13)	960	42	24	−4	38	21	3	12.5	4a, 7b, 9
Parietal										
Left	SPL (7)	3320	−28	−56	54	−28	−59	47	14.7	4a, 7a, 11
	SPL/IPS (7)	–	−30	−66	42	−29	−67	35	10.3	4a, 9, 11
Occipitotemporal										
Left	FG (37)	2128	−40	−52	−12	−38	−49	−12	14.3	4a, 7a, 7b, 9
	FG (37)	–	−42	−56	−12	−40	−53	−12	14.2	4a, 7a, 7b, 9
Sub-cortical										
Left	Putamen	920	−24	−4	−2	−23	−5	2	15.6	4a, 7a, 7c
Cerebellum										
Right	Cerebellum	1264	6	−64	−20	5	−60	−19	13.2	4a, 7a, 7c
Right	Cerebellum	–	8	−72	−18	6	−67	−18	11.8	7b, 9, 11
	Cerebellum	776	14	−48	−24	12	−44	−21	16.3	1, 4c, 7a, 7c

STG/STS, FG, and ITG. In addition to confirming reports from the existing lesion and functional neuroimaging literatures, this analysis provided the most likely specific location of each of these sites. Among the results there are two findings that are particularly noteworthy. We briefly mention them here and discuss them more fully below. First, the analysis revealed that the regions with the highest correspondences across studies were in the left FG/ITG (observed in 11 of the 17 contrasts) and left IFG (observed in 8 of the 17 contrasts). Second, although the lesion literature has traditionally identified the AG as a key site within the written word production network, the meta-analysis did not identify the AG as a consistent activation site. Instead, the analysis identified a region just superior and medial to the AG in the left posterior IPS.

Additional analyses allowed further clarification of the functional roles of the identified regions, in particular whether they were associated with central vs. peripheral written language processes. Areas most consistently involved with central processes were: the left FG/ITG, IFG, and posterior IPS. Peripheral processes, as identified through the subtraction of ALEs, were most robustly associated with foci within the left SFG/SFS, precentral gyrus, postcentral gyrus and SPL. We discuss these results in the sections below within the context of both the neuroimaging and lesion literatures on spelling (**Figure 1**); we also discuss these findings in relation to the reading literature when appropriate. This discussion highlights points of increasing convergence across the various literatures and also identifies controversies and topics to be pursued in future research.

THE CENTRAL PROCESSES OF SPELLING

Fusiform gyrus/inferior temporal gyrus

One of the most salient and robust findings of the All-Contrasts analysis is that the left FG/ITG region (with three separate MNI peaks at −50, −60, −16; −44, −56, −12; and −42, −58, −6) had one of the highest ALE values generated in the analysis, as well as the greatest number (11) of contributing contrasts. Thus, despite the very considerable diversity in experimental tasks and contrasts, this region was consistently identified as contributing to spelling, and more specifically to the central processes in spelling.

In terms of cognitive processes, this region has been most frequently associated with Orthographic LTM. This is supported by a number of lesion studies (Rapp and Caramazza, 1997; Rapcsak and Beeson, 2004; Philipose et al., 2007; Tsapkini and Rapp, 2010) reporting that lesions to this area resulted in the behavioral profile expected from damage to the Orthographic LTM system. This profile includes difficulties in spelling lower frequency irregular words (i.e., words with atypical PG mappings, e.g., yacht), in the face of accurate spelling of regular words and pseudowords. Consistent with the association of this region with Orthographic LTM, recent fMRI studies have found lexical frequency effects in this region, with larger BOLD responses observed for low vs. high frequency words (Rapp and Lipka, 2011; Rapp and Dufor, 2011). Thus, the lesion and neuroimaging spelling literatures are consistent in indicating that this region plays some role in Orthographic LTM.

Table 4 | Central + peripheral (C + P) > central (C) contrasts meta-analysis.

Location (Brodmann area)		Extent (mm^3)	Peak MNI			Peak TAL		
			X	*Y*	*Z*	*X*	*Y*	*Z*
(C + P) > C								
Frontal								
Left	Precentral gyrus (6)	256	−24	−11	64	−24	−17	60
Left	SFG/SFS (6)	–	−26	−4	62	−26	−11	59
Parietal								
Left	Postcentral gyrus (3)	272	−39	−27	63	−38	−32	57
Left	Postcentral gyrus (3)		−34	−27	58	−33	−32	53
Left	SPL/IPS (7)	1744	−36	−40	57	−35	−44	51
Left	SPL/IPS (7)	–	−33	−46	53	−32	−49	47

In addition, from the time of the earliest investigations of written language, the left occipitotemporal cortex has being considered to be critical for reading words (Dejerine, 1892). Since then, numerous functional neuroimaging and lesion studies have confirmed the association of the left mid-FG (bordering on the ITG), with visual word processing in literate individuals (Chialant and Caramazza, 1998; Cohen et al., 2000, 2002; McCandliss et al., 2003; Kronbichler et al., 2004; Gaillard et al., 2006; Sheldon et al., 2008; Dehaene et al., 2010). These findings are also generally consistent with the visual object processing literature which includes the left occipitotemporal cortex as part of the ventral visual system of areas selective to common complex visual objects such as faces or body parts (Kanwisher et al., 1997; Schwarzlose et al., 2005). While the specific role of this region in reading is debated (Price and Devlin, 2003), there is evidence, similar to that found for spelling, that this region is sensitive to the lexical frequency of words (Kronbichler et al., 2004) and it has been designated by some investigators as the storage site of the orthographic word forms used in reading (Glezer et al., 2009).

Although findings from the spelling and reading literatures suggests that both skills share neural substrates in the left occipitotemporal cortex, this claim is most clearly established by studies that examine both spelling and reading within the same individuals. In fact, four recent fMRI studies determined that there was overlapping activation for spelling and reading in the left FG/ITG (Cho et al., 2009; Purcell et al., 2011; Rapp and Lipka, 2011; Rapp and Dufor, 2011). These functional neuroimaging findings are also supported by studies of acute and chronic lesions reporting that lesions or hypoperfusion to the left FG/ITG can impair both spelling and reading in a similar manner (Philipose et al., 2007; Tsapkini and Rapp, 2010).

In sum, the finding of overlap within the FG/ITG region for both reading and spelling in both lesion and neuroimaging literatures, as well as the sensitivity of the region to word frequency, are consistent with some role for this region in the Orthographic LTM system.

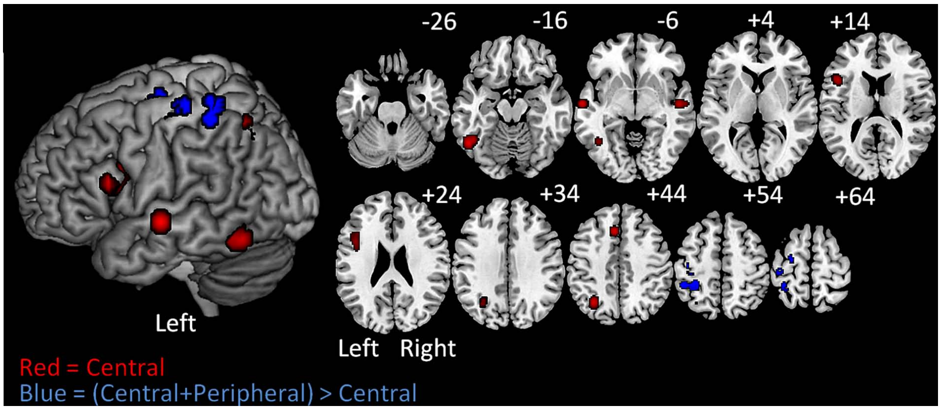


FIGURE 3 | On the left are ALE Meta-analysis results projected on a standard rendered template brain and on the right are the corresponding axial slices from *z* −26 to +64 in 10 mm increments. In red are the clusters corresponding to the central-contrast analysis. These clusters correspond to the central processes of written word production. In blue are the clusters identified in the (central + peripheral) > central-contrast analysis. These clusters correspond to peripheral processes of written production.

Left inferior frontal gyrus

The overall ALE analysis identified a high activation likelihood region within the left IFG (with MNI peaks at: −46, 16, 18 and −44, 6, 28) that was supported by eight contrasts, making this region the second most consistently observed across studies. The peaks of this cluster are in the posterior part of the IFG (opercularis), with the cluster extending into the precentral sulcus. This neuroanatomical region has been referred to by some researchers as the Inferior Frontal Junction, IFJ (Brass and Von Cramon, 2002; Brass et al., 2005; Derrfuss et al., 2005).

Damage to this area of cortex is often associated with written and spoken language deficits which is unsurprising considering this region is located within the dorsal extent of Broca's area. More specifically, with regard to acquired dysgraphia in acute stroke, hypoperfusion to the left IFG has been associated with impairments in accessing Orthographic LTM for spelling (Hillis et al., 2002; Hillis and Rapp, 2004). Also consistent with a role for this region in Orthographic LTM, Rapp and Dufor (2011) found the region to be sensitive to word frequency, showing stronger BOLD response to low vs. high frequency words. However, it should be noted that there are studies of chronic stroke associating damage to this region with deficits affecting PG conversion processes (Henry et al., 2007).

With regard to reading, like the FG/ITG region, the posterior IFG is often found to be active in studies of reading (e.g., Fiez and Petersen, 1998; Price, 2000; Turkeltaub et al., 2002; Joubert et al., 2004; Bolger et al., 2005). Furthermore, three of the neuroimaging studies that have considered both reading and spelling in the same individuals reported overlap between spelling and reading in this region (Purcell et al., 2011; Rapp and Lipka, 2011; Rapp and Dufor, 2011). This latter finding suggests a role for this region in the central processes of spelling.

Specific interpretation of the functional role of the posterior IFG in spelling is complicated because activations in this general region are reported in a large number of studies across quite disparate skill domains. For example, Brass and Von Cramon (2002) suggested that the IFJ is involved in cognitive control (see Rapp and Lipka, 2011 for discussion in the context of spelling). Worth noting is that even within the All-Contrasts meta-analysis, we found heterogeneity in this region with six of the contrasts contributing to this area coming from the Central-only group of studies and only two from the Central + Peripheral group (although the contrasts in this latter group must also draw on central spelling processes). This heterogeneity may reflect the sensitivity of this region to the specific task demands and the particular features of the tasks used in the contrasts. One interpretation, which falls generally within a cognitive control account, is that activation in the left IFG for written language processing is not associated directly with Orthographic LTM, but rather with the coordination of activity in more posterior regions (i.e., in temporal or parietal cortex) that are, themselves, more directly involved in processing the stored representations of written language (Mesulam, 1998; Booth et al., 2002; Bitan et al., 2005). This interpretation suggests that the left IFG/IFJ's participation in orthographic processing may primarily consist of its contribution to the processes of orthographic lexical selection and retrieval, rather than storage. This idea is supported by the findings in spoken word production studies which report that the IFJ/IFG region is associated with lexical selection in speaking (Martin et al., 1994; Thompson-Schill et al., 1997).This proposal is further elaborated by Kan and Thompson-Schill (2004) who suggested that biased competition processes (Desimone and Duncan, 1995) draw on this neuroanatomical region when there is competition for selection in visual, lexical, or conceptual domains. That is, while storage of orthographic forms may reside in FG/ITG regions, lexical selection and the resolution of competition at orthographic, lexical, conceptual and/or phonological levels, in spelling and reading, may rely on the left IFJ.

The angular gyrus controversy

The left angular gyrus (AG) was one of the first regions associated with central processes of spelling and reading (Dejerine, 1892). The region has continued to be associated with literacy and developmental dyslexia (Horwitz et al., 1998; Rumsey et al., 1999; Pugh et al., 2000; Carreiras et al., 2009). The specific function of this region with regard to reading is unclear, however, with some researchers positing that it is part of a larger dorsal parietotemporal region which plays a role in grapheme–phoneme conversion in reading (e.g., Pugh et al., 2000), or that it contributes to lexical processing associated with reading high frequency words (e.g., Joubert et al., 2004). The function of this region with regard to written production is also unclear. Although the left AG has continued to be associated with acquired dysgraphia, there is no clear consensus regarding its functional role in writing. For instance, some studies have linked impairments in irregular word spelling to the left AG suggesting it is involved in Orthographic LTM (Beauvois and Derouesne, 1981; Roeltgen and Heilman, 1984), while others have found it to be associated with deficits to sub-lexical PG conversion spelling processes (Hillis et al., 2002; Sheldon et al., 2008).

Conflicting with the lesion literature findings, the left AG has not been clearly identified in functional neuroimaging studies of spelling. In particular, Beeson et al. (2003) specifically examined the functional activation in a left AG region of interest and found no significant activation for the group analysis, although individual subjects did show some significant effects (Beeson et al., 2003). Interestingly, the left AG has not been consistently identified in the neuroimaging literature of reading either (e.g., Fiez and Petersen, 1998; Turkeltaub et al., 2002; Jobard et al., 2003). Consistent with these previous findings, the meta-analysis did not identify any reliable activation peaks in the AG. This confirms that this region is not reliably identified in functional neuroimaging studies of written word production.

It may be useful to consider these inconsistent reports in light of recent observations that activation in the left AG has been found in neuroimaging studies involving conceptual or semantic processing (Binder et al., 2009; Graves et al., 2009; Seghier et al., 2011). While spelling (as indicated in **Figure 1**) and reading both engage semantic processes, tasks involving spelling or reading which are not designed to focus the participant's attention on the semantic features of the words may not produce sufficiently consistent activation in semantically related areas, such as the AG, to be detected by experimental paradigms/analysis techniques employed. In sum, the inconsistently observed activation of the AG may be due to low and variable activation of semantic processing regions in neuroimaging studies involving spelling or reading tasks that are not focused on semantic processing.

That said, the All-Contrasts analysis did identify a cluster in the posterior IPS which is supported by five experimental contrasts and is located just superior and medial to the left AG, in the left posterior IPS (MNI peak: −30, −60, 46). It is important to point out that the posterior IPS is essentially the neuroanatomical demarcation between the superior parietal lobule and the AG and, therefore, that the IPS cluster that we report could be considered to be at the most superior edge of the AG. Importantly, a similar region was observed in both the C-Contrasts and C + P Contrasts separately suggesting that this region plays a role in central spelling processes.

If we consider cognitive processes that have been associated with the SPL/IPS region specifically, we find that damage to superior parietal cortex, entirely, or largely sparing the AG, has been associated with spelling deficits, specifically ones affecting Orthographic WM (Miceli et al., 2004; Buchwald and Rapp, 2009). In particular, one of the individuals described by Buchwald and Rapp (2009) had a lesion restricted to the parietal lobe anterior to the AG. Furthermore, also consistent with a role for this region in Orthographic WM is the finding reported by Rapp and Dufor (2011) of sensitivity of the BOLD response in this area to word length in written word production.

One way to reconcile the current body of functional neuroimaging and the lesion findings implicating the left AG, is to assume that the IPS region is, in fact, the critical region for spelling and that lesions to the left AG may have also damaged the posterior IPS (e.g., due to the close proximity of these regions) or connections to the IPS. Still another interpretation is that lesions which have previously been associated with the left AG have actually affected a portion of the left inferior temporal lobe or FG that was important for written spelling. This possibility has been suggested in lesion studies examining deficits in spelling (Rapcsak and Beeson, 2004) as well as reading (Epelbaum et al., 2008). One final possibility is that the AG serves to modulate activity in areas related to orthographic processing and that it may not be detectable via classic functional neuroimaging measures utilized in the studies considered in this meta-analysis, but instead can be observed via anatomical or functional connectivity measures. This is supported by recent work which reported gray matter increases in the bilateral AG as well as white matter increases in the splenium of the corpus callosum (which serves to functionally connect the bilateral angular gyri)were associated with learning to read and write as an adult (Carreiras et al., 2009). These varied interpretations indicate that much further research is needed in order to more clearly discern the relevance of the AG to written spelling.

Other perisylvian regions: SMG and STG/STS

The remaining left hemisphere cortical sites that were identified in the All-Contrasts analysis and that were not attributed to peripheral processes are the mid left STG/STS (MNI peak: −60, −12, −2) and the SMG (MNI peak: −52, −32, 34). Lesions to either the left STG/STS or SMG have typically led to impaired written word production by affecting phonological processing thought to be associated with the PG conversion system. This type of impairment results in difficulties in pseudoword spelling with relatively more intact spelling of both regular and irregular familiar words (Henry et al., 2007; also see Philipose et al., 2007 for parallel deficits in spelling and reading pseudowords subsequent to damage to BA40). Furthermore, the left SMG has been directly implicated in functional neuroimaging studies of Japanese Kana writing (a phonetic written language system) which is thought to rely heavily on PG conversion (Katanoda et al., 2001; Sugihara et al., 2006). Analogous studies in an alphabetic language such as English would involve pseudoword writing, but to date no such neuroimaging study has been performed.

With specific regard to the STG/STS sites, it is worth noting that this was the one cortical area in which bilateral activation was observed. Activation in these sites (especially the right hemisphere location) were supported largely by the specific contrasts used by Booth et al. (2002, 2003b, 2004) involving a task that required subjects to compare the rime spelling of three different auditory words. This task, unlike many of the others examined, involved considerable phonological processing and phonological working memory. In this regard it is worth noting that the mid to posterior STG/STS region has been associated in both lesion and neuroimaging studies with phonological processing and phonological deficits (Turkeltaub and Coslett, 2011). Given this, it may not be surprising that this area is involved in PG conversion processes. The similarity and likely close topographic proximity of regions implicated in PG conversion and phonological processes more generally, underscores the importance for future work to be specifically directed at distinguishing spelling-specific from more general phonological processes (e.g., Rapcsak et al., 2009).

THE PERIPHERAL PROCESSES OF SPELLING

A subtraction of the ALE maps of the Central from the Central + Peripheral contrasts was designed to reveal the locations most likely to have peak activations associated with the peripheral processes of spelling: those processes that are involved with the format-specific production of spelling knowledge (see **Figure 1**). It is worth making two points before continuing. First, for almost all of the contrasts we considered, there were control conditions designed to "subtract out" the most peripheral aspects of the motor responses produced in these tasks, with the goal of more clearly identifying spelling-specific peripheral processes. Although these contrast conditions may have been more or less successful in accomplishing this goal, it is for this reason that we will not focus on identifying or discussing general motor processes and substrates. Second, although there are empirical and computational reasons to assume that there are multiple components involved in the peripheral aspects of spelling (as depicted in **Figure 1** and discussed in the Introduction), relatively little work has been carried out to differentiate them or their neural substrates. Thus, the discussion of the results in this section will be far more limited than was the discussion of the central processes of spelling.

The regions identified as being primarily associated with peripheral processes were: the left precentral gyrus and SFG/SFS, the left postcentral gyrus and left SPL/IPS. In addition, there were other motor related regions, such as the left SMA and the right cerebellum, that were significant in the All-Contrasts meta-analysis and were supported by contrasts that involved handwriting or typing output (some of which did not include motor baselines). All of these areas have been implicated in the lesion literature as being important for writing. Some have been linked specifically to writing such as the left SFG/SFS and the SPL/IPS, while others have been associated with motor processes more generally, such as the left SMA, left postcentral gyrus and the right cerebellum (see Gerloff et al. (1997) for discussion regarding the SMA). In subsequent sections we focus our discussion on the left SFG/SFS and the SPL/IPS.

The precentral gyrus and superior frontal gyrus/sulcus

The subtraction analysis designed to identify peripheral writing substrates identified a high ALE region centered in the left precentral gyrus (MNI peak: −24, −11, 64) and extending anteriorly into the SFG/SFS (MNI peak: −26, −4, 62).

In the lesion literature, the conversion of graphemic representations to motor commands has been associated with a region in the left posterior middle frontal gyrus and SFG (BA6) often referred to in the clinical neuroscience literature as Exner's area. For handwriting, this region has primarily been associated with the translation of orthographic representations into letter shapes (i.e., allographic processing; Exner, 1881; Ritaccio et al., 1992; Tohgi et al., 1995; Lubrano et al., 2004). Interestingly a recent fMRI study of Japanese Kana handwriting identified a peak in the left SFS (MNI peak: −28, −9, 54; Sugihara et al., 2006) that is within 7 mm of the SFS peak identified in our Central + Peripheral meta-analysis (MNI peak: −22, −8, 54). These results suggest that this region may be associated with processes common to both alphabetic and syllabic writing systems such as Kana. Additionally, upon further examination of the SFG/SFS peak from the Central + Peripheral analysis from **Table 3**, we find that this peak was associated with two handwriting tasks and one keyboard typing study, thus suggesting that it may play a role in the conversion of graphemic representations to motor commands regardless of whether the word is handwritten or typed. Based on these findings it seems that, although the left SFG/SFS is consistently associated with written motor production, the specific functional role of this region in the conversion of graphemes to motor commands needs to be examined further with regard to whether or not there is specificity for output modality (i.e., handwriting or typing).

Another possibility regarding the function of the precentral/SFG/SFS region is that it plays a critical role in Orthographic WM. This was proposed in a recent fMRI study comparing activations from the spelling of longer as compared to shorter words (matched for writing times and motor output) which reported length dependent activation in both the left SFS and SPL (Rapp and Dufor, 2011). These authors argued that not only was the SFS finding indicative of Orthographic WM functions for this frontal region, but that this claim would explain the findings that the area is recruited for spelling across modalities of output (handwriting and typing), as both output formats should rely on a common Orthographic WM system. Reporting findings consistent with this proposal, Cloutman et al. (2009) carried out a study of individuals in the first 48 h after left hemisphere stroke onset, with one group showing clear signs of Orthographic WM impairment while the other had graphemic buffer sparing. They found that, among other regions, the left premotor cortex was reliably associated with graphemic buffer damage.

The inconsistency that the Rapp and Dufor (2011) proposal creates is that Orthographic WM is considered to be a component of central spelling, yet it was identified by the meta-analysis as being especially associated with peripheral processes. However, it is important to note that Orthographic WM is situated at the interface between central and peripheral processing and may actually interact with peripheral component processes in a manner which may make it difficult to distinguish between central and peripheral processes that are in the same or neighboring regions. Further work will need to be done to better characterize the functional role of the left SFS with regard to its role in Orthographic WM as compared to the more peripheral processes of written production. In addition, if there turns out to be a reliable relationship between Orthographic WM and the posterior SPL/IPS regions (as suggested in an earlier section), the specific roles of the frontal and parietal areas in Orthographic WM will also need to be examined. In that regard it is interesting to note that there have been several proposals of multiple cognitive components of Orthographic WM, much like the multi-component view of other WM systems (see Rapp and Kong, 2002; Kan et al., 2006). Presumably, different components of Orthographic WM would draw on different neural substrates and may explain the sensitivity of both SFS and SPL to the length of words being spelled.

Superior parietal lobule

Aside from the generation of graphemic motor commands, writing requires the generation of the correct sequences of motor commands. The left SPL in particular has been associated with sequence production in written language. For instance, damage to left SPL has be associated with apraxic agraphia, a disorder characterized by deficits in the generation of correct sequences of movements required for handwriting (Alexander et al., 1992; Sakurai et al., 2007). In support of these findings we identified a left anterior SPL/IPS cluster which was shown to be associated with the peripheral but not central processes of written production (MNI peak: −36, −40, 57). This peak is notably anterior to the one identified in the Central-contrasts analysis (MNI peak: −30, −62, 44) which suggests that there may be a distinction between certain peripheral processes that draw on the more anterior SPL and central processes (e.g., Orthographic WM) that make use of the more posterior SPL/IPS region.

The left SPL has also been associated with the learning and generation of complex motor sequences (Haaland et al., 2004) as well as with the integration of the body-schema with allo- and egocentric frames of reference (Neggers et al., 2006). Other work has shown that the left SPL, in particular the region including the IPS, is topographically organized such that the more anterior portion is associated with the coordination of body movements in space (e.g., grasping and pointing) and the more posterior regions are associated with calculation and eye saccades (Simon et al., 2002). Understanding the relationship amongst the cognitive requirements of writing and these other tasks and, in turn, the specific neural substrates that support these cognitive processes will be an important research direction.

SUMMARY AND CONCLUSION

This study provides the first quantitative review and synthesis of findings from functional neuroimaging studies of written language production. The analyses carried out identify a network of left hemisphere frontal, parietal, and temporal sites that are reliably and consistently associated with written word production, providing estimates of the most likely locations of peak activations within these areas. Furthermore, these analyses provide a coherent assignment of these locations to central or peripheral processes of spelling. The General Discussion relates these sites to more specific cognitive functions by bringing together results from the lesion and neuroimaging literatures. This study sets the stage for future research regarding the neural substrates of written word production by identifying the literature's most reliable findings and by highlighting critical questions regarding the neural basis of written language productions that still need to be addressed by future studies.

ACKNOWLEDGMENTS

This work was supported by NIDCD006740 (Brenda Rapp), NICHD R01HD056107 (Guinevere F. Eden), NIDCD (F31DC009545) (Jeremy J. Purcell), NSF SBE 0541953 (Jeremy J. Purcell and Guinevere F. Eden), and the American Academy of Neurology Foundation Clinical Research Training Fellowship (Peter E. Turkeltaub). We would like to thank Drs Pelagie Beeson and Olivier Dufor for providing additional data for these analyses.

REFERENCES

Alexander, M. P., Fischer, R. S., and Friedman, R. (1992). Lesion localization in apractic agraphia. *Arch. Neurol.* 49, 246–251.

Beauvois, M. F., and Derouesne, J. (1981). Lexical or orthographic agraphia. *Brain* 104, 21–49.

Beeson, P., Rapcsak, S., Plante, E., Chargualaf, J., Chung, A., Johnson, S., and Trouard, T. (2003). The neural substrates of writing: a functional magnetic resonance imaging study. *Aphasiology* 17, 647–665.

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.

Bitan, T., Booth, J. R., Choy, J., Burman, D. D., Gitelman, D. R., and Mesulam, M. M. (2005). Shifts of effective connectivity within a language network during rhyming and spelling. *J. Neurosci.* 25, 5397–5403.

Bolger, D. J., Perfetti, C. A., and Schneider, W. (2005). Cross-cultural effect on the brain revisited: universal structures plus writing system variation. *Hum. Brain Mapp.* 25, 92–104.

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., and Mesulam, M. M. (2002). Functional anatomy of intra- and cross-modal lexical tasks. *Neuroimage* 16, 7–22.

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., and Mesulam, M. M. (2003a). Relation between brain activation and lexical performance. *Hum. Brain Mapp.* 19, 155–169.

Booth, J. R., Burman, D. D., Meyer, J. R., Lei, Z., Choy, J., Gitelman, D. R., Parrish, T. B., and Mesulam, M. M. (2003b). Modality-specific and -independent developmental differences in the neural substrate for lexical processing. *J. Neurolinguistics* 16, 383–405.

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., and Mesulam, M. M. (2004). Development of brain mechanisms for processing orthographic and phonologic representations. *J. Cogn. Neurosci.* 16, 1234–1249.

Brass, M., Derrfuss, J., Forstmann, B., and Von Cramon, D. Y. (2005). The role of the inferior frontal junction area in cognitive control. *Trends Cogn. Sci. (Regul. Ed.)* 9, 314–316.

Brass, M., and Von Cramon, D. Y. (2002). The role of the frontal cortex in task preparation. *Cereb. Cortex* 12, 908–914.

Buchwald, A., and Rapp, B. (2009). Distinctions between orthographic long-term memory and working memory. *Cogn. Neuropsychol.* 26, 724–751.

Burt, J. S., and Fury, M. B. (2000). Spelling in adults: the role of reading skills and experience. *Read. Writ.* 13, 1–30.

Burt, J. S., and Tate, H. (2002). Does a reading lexicon provide orthographic representations for spelling? *J. Mem. Lang.* 46, 518–543.

Caramazza, A., and Hillis, A. E. (1990). Spatial representation of words in the brain implied by studies of a unilateral neglect patient. *Nature* 346, 267–269.

Carreiras, M., Seghier, M. L., Baquero, S., Estevez, A., Lozano, A., Devlin, J. T., and Price, C. J. (2009). An anatomical signature for literacy. *Nature* 461, 983–986.

Chialant, D., and Caramazza, A. (1998). Perceptual and lexical factors in a case of letter-by-letter reading. *Cogn. Neuropsychol.* 15, 167–201.

Cho, H., Rapcsak, S. Z., and Beeson, P. M. (2009). Evidence for a common neural substrate of orthographic processing during reading and spelling. Neurobiology of Language Conference, Chicago, IL.

Cloutman, L., Gingis, L., Newhart, M., Davis, C., Heidler-Gary, J., Crinion, J., and Hillis, A. E. (2009). A neural network critical for spelling. *Ann. Neurol.* 66, 249–253.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., and Michel, F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123(Pt 2), 291–307.

Cohen, L., Jobert, A., Le Bihan, D., and Dehaene, S. (2004). Distinct unimodal and multimodal regions for word processing in the left temporal cortex. *Neuroimage* 23, 1256–1270.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., and Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain* 125, 1054–1069.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., and Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science* 330, 1359–1364.

Dejerine, J. (1892). Contribution a l'étude anatomopathologique et clinique des différentes variétés de cécité verbale. *Mem. Soc. Biol.* 4, 61–90.

Denes, G., Signorini, M., and Volpato, C. (2005). Post graphemic impairments of writing: the case of micrographia. *Neurocase* 11, 176–181.

Derrfuss, J., Brass, M., Neumann, J., and Von Cramon, D. Y. (2005). Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. *Hum. Brain Mapp.* 25, 22–34.

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

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., and Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. *Hum. Brain Mapp.* 30, 2907–2926.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., and Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335.

Ellis, A. W. (1979). Slips of the pen. *Visible Lang.* 13, 265–282.

Ellis, A. W. (1982). "Spelling and writing (and reading and speaking)," in *Normality and Pathology in Cognitive Functions*, ed. A. W. Ellis (New York: Academic Press), 113–143.

Ellis, A. W., and Young, A. W. (1988). *Human Cognitive Neuropsychology*. London: Lawrence Erlbaum Associates.

Epelbaum, S., Pinel, P., Gaillard, R., Delmaire, C., Perrin, M., Dupont, S., Dehaene, S., and Cohen, L. (2008). Pure alexia as a disconnection syndrome: new diffusion imaging evidence for an old concept. *Cortex* 44, 962–974.

Exner, S. (1881). *Untersuchungen über die Lokalisation der Functionen in der Grosshirnrinde des Menschen*. Wien: W. Braumuller.

Fiez, J. A., and Petersen, S. E. (1998). Neuroimaging studies of word reading. *Proc. Natl. Acad. Sci. U.S.A.* 95, 914–921.

Fournier Del Castillo, M. C., Maldonado Belmonte, M. J., Ruiz-Falco Rojas, M. L., Lopez Pino, M. A., Bernabeu Verdu, J., and Suarez Rodriguez, J. M. (2011). Cerebellum atrophy and development of a peripheral dysgraphia: a paediatric case. *Cerebellum* 9, 530–536.

Gaillard, R., Naccache, L., Pinel, P., Clemenceau, S., Volle, E., Hasboun, D., Dupont, S., Baulac, M., Dehaene, S., Adam, C., and Cohen, L. (2006). Direct intracranial, FMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron* 50, 191–204.

Gerloff, C., Corwell, B., Chen, R., Hallett, M., and Cohen, L. G. (1997). Stimulation over the human supplementary motor area interferes with the organization of future elements in complex motor sequences. *Brain* 120(Pt 9), 1587–1602.

Glezer, L. S., Jiang, X., and Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the "visual word form area." *Neuron* 62, 199–204.

Graves, W. W., Desai, R., Humphries, C., Seidenberg, M. S., and Binder, J. R. (2009). Neural systems for reading aloud: a multiparametric approach. *Cereb. Cortex* 20, 1799–1815.

Haaland, K. Y., Elsinger, C. L., Mayer, A. R., Durgerian, S., and Rao, S. M. (2004). Motor sequence complexity and performing hand produce differential patterns of hemispheric lateralization. *J. Cogn. Neurosci.* 16, 621–636.

Henry, M. L., Beeson, P. M., Stark, A. J., and Rapcsak, S. Z. (2007). The role of left perisylvian cortical regions in spelling. *Brain Lang.* 100, 44–52.

Hillis, A., Kane, A., Tuffiash, E., Beauchamp, N., Barker, P., Jacobs, M., and Wityk, R. (2002). Neural substrates of the cognitive processes underlying spelling: evidence from MR diffusion and perfusion imaging. *Aphasiology* 16, 425–438.

Hillis, A., and Rapp, B. (2004). "Cognitive and neural substrates of written language: comprehension and production," in *The Cognitive Neurosciences III*, ed. M. S. Gazzaniga (Cambridge, MA: The MIT press), 775–787.

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., and Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. *J. Comput. Assist. Tomogr.* 22, 324–333.

Horwitz, B., Rumsey, J. M., and Donohue, B. C. (1998). Functional connectivity of the angular gyrus in normal reading and dyslexia. *Proc. Natl. Acad. Sci. U.S.A.* 95, 8939–8944.

Jobard, G., Crivello, F., and Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. *Neuroimage* 20, 693–712.

Joubert, S., Beauregard, M., Walter, N., Bourgouin, P., Beaudoin, G., Leroux, J. M., Karama, S., and Lecours, A. R. (2004). Neural correlates of lexical and sublexical processes in reading. *Brain Lang.* 89, 9–20.

Kan, I. P., Biran, I., Thompson-Schill, S. L., and Chatterjee, A. (2006). Letter selection and letter assembly in acquired dysgraphia. *Cogn. Behav. Neurol.* 19, 225–236.

Kan, I. P., and Thompson-Schill, S. L. (2004). Selection from perceptual and conceptual representations. *Cogn. Affect. Behav. Neurosci.* 4, 466–482.

Kanwisher, N., Mcdermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.

Katanoda, K., Yoshikawa, K., and Sugishita, M. (2001). A functional MRI study on the neural substrates for writing. *Hum. Brain Mapp.* 13, 34–42.

Kawahata, N., Nagata, K., and Shishido, F. (1988). Alexia with agraphia due to the left posterior inferior temporal lobe lesion – neuropsychological analysis and its pathogenetic mechanisms. *Brain Lang.* 33, 296–310.

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., and Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: evidence from a parametric fMRI study. *Neuroimage* 21, 946–953.

Lancaster, J. L., Tordesillas-Gutierrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., Mazziotta, J. C., and Fox, P. T. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Hum. Brain Mapp.* 28, 1194–1205.

Lubrano, V., Roux, F. E., and Demonet, J. F. (2004). Writing-specific sites in frontal areas: a cortical stimulation study. *J. Neurosurg.* 101, 787–798.

Magrassi, L., Bongetta, D., Bianchini, S., Berardesca, M., and Arienta, C. (2010). Central and peripheral components of writing critically depend on a defined area of the dominant superior parietal gyrus. *Brain Res.* 1346, 145–154.

Margolin, D. I. (1984). The neuropsychology of writing and spelling: semantic, phonological, motor, and perceptual processes. *Q. J. Exp. Psychol. A.* 36, 459–489.

Martin, N., Dell, G. S., Saffran, E. M., and Schwartz, M. F. (1994). Origins of paraphasias in deep dysphasia: testing the consequences of a decay impairment to an interactive spreading activation model of lexical retrieval. *Brain Lang.* 47, 609–660.

Mesulam, M. M. (1998). From sensation to cognition. *Brain* 121 (Pt 6), 1013–1052.

McCandliss, B. D., Cohen, L., and Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci. (Regul. Ed.)* 7, 293–299.

McCloskey, M., Macaruso, P., and Rapp, B. (2006). Grapheme-to-lexeme feedback in the spelling system: evidence from a dysgraphic patient. *Cogn. Neuropsychol.* 23, 278–307.

Menon, V., and Desmond, J. E. (2001). Left superior parietal cortex involvement in writing: integrating fMRI with lesion evidence. *Brain Res. Cogn. Brain Res.* 12, 337–340.

Miceli, G., Capasso, R., Benvegnu, B., and Caramazza, A. (2004). The categorical distinction of vowel and consonant representations: evidence from dysgraphia. *Neurocase* 10, 109–121.

Neggers, S. F., Van Der Lubbe, R. H., Ramsey, N. F., and Postma, A. (2006). Interactions between ego- and allocentric neuronal representations of space. *Neuroimage* 31, 320–331.

Patterson, K. (1986). Lexical but nonsemantic spelling? *Cogn. Neuropsychol.* 3, 341–367.

Petrides, M., Alivisatos, B., and Evans, A. C. (1995). Functional activation of the human ventrolateral frontal cortex during mnemonic retrieval of verbal information. *Proc. Natl. Acad. Sci. U.S.A.* 92, 5803–5807.

Philipose, L. E., Gottesman, R. F., Newhart, M., Kleinman, J. T., Herskovits, E. H., Pawlak, M. A., Marsh, E. B., Davis, C., Heidler-Gary, J., and Hillis, A. E. (2007). Neural regions essential for reading and spelling of words and pseudowords. *Ann. Neurol.* 62, 481–492.

Pramstaller, P. P., and Marsden, C. D. (1996). The basal ganglia and apraxia. *Brain* 119(Pt 1), 319–340.

Price, C. (2000). The anatomy of language: contributions from functional neuroimaging. *J. Anat.* 197, 335–359.

Price, C. J., and Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage* 19, 473–481.

Price, C. J., Devlin, J. T., Moore, C. J., Morton, C., and Laird, A. R. (2005). Meta-analyses of object naming: effect of baseline. *Hum. Brain Mapp.* 25, 70–82.

Pugh, K. R., Mencl, W. E., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Constable, R. T., Skudlarski, P., Marchione, K. E., Jenner, A. R., Fletcher, J. M., Liberman, A. M., Shankweiler, D. P., Katz, L., Lacadie, C., and Gore, J. C. (2000). The angular gyrus in developmental dyslexia: task-specific differences in functional connectivity within posterior cortex. *Psychol. Sci.* 11, 51–56.

Purcell, J. J., Napoliello, E. M., and Eden, G. F. (2011). A combined fMRI study of typed spelling and reading. *Neuroimage* 55, 750–762.

Rapcsak, S., and Beeson, P. (2002). "Neuroanatomical correlates of spelling and writing," in *The Handbook of Adult Language Disorders*, ed. A. Hillis (New York: Psychology Press), 71–100.

Rapcsak, S. Z., and Beeson, P. M. (2004). The role of left posterior inferior temporal cortex in spelling. *Neurology* 62, 2221–2229.

Rapcsak, S. Z., Beeson, P. M., Henry, M. L., Leyden, A., Kim, E., Rising, K., Andersen, S., and Cho, H. (2009). Phonological dyslexia and dysgraphia: cognitive mechanisms and neural substrates. *Cortex* 45, 575–591.

Rapp, B., and Caramazza, A. (1997). From graphemes to abstract letter shapes: levels of representation in written spelling. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 1130–1152.

Rapp, B., and Dufor, O. (2011). The neurotopography of written word production: an fMRI investigation of the distribution of sensitivity to word length and frequency. *J. Cogn. Neurosci.* PMID: 21812571. [Epub ahead of print].

Rapp, B., Epstein, C., and Tainturier, M. J. (2002). The integration of information across lexical and sublexical processes in spelling. *Cogn. Neuropsychol.* 19, 1–29.

Rapp, B., and Kong, D. (2002). Revealing the component functions of the graphemic buffer. *Brain Lang.* 83, 112–114.

Rapp, B., and Lipka, K. (2011). The literate brain: the relationship between spelling and reading. *J. Cogn. Neurosci.* 23, 1–18.

Ritaccio, A. L., Hickling, E. J., and Ramani, V. (1992). The role of dominant premotor cortex and grapheme to phoneme transformation in reading epilepsy. A neuroanatomic, neurophysiologic, and neuropsychological study. *Arch. Neurol.* 49, 933–939.

Roeltgen, D. P., and Heilman, K. M. (1984). Lexical agraphia. Further support for the two-system hypothesis of linguistic agraphia. *Brain* 107(Pt 3), 811–827.

Roeltgen, D. P., and Heilman, K. M. (1985). Review of agraphia and a proposal for an anatomically-based neuropsychological model of writing. *Appl. Psycholinguist.* 6, 205–229.

Roux, F. E., Dufor, O., Giussani, C., Wamain, Y., Draper, L., Longcamp, M., and Demonet, J. F. (2009). The graphemic/motor frontal area Exner's area revisited. *Ann. Neurol.* 66, 537–545.

Rumsey, J. M., Horwitz, B., Donohue, B. C., Nace, K. L., Maisog, J. M., and Andreason, P. (1999). A functional lesion in developmental dyslexia: left angular gyral blood flow predicts severity. *Brain Lang.* 70, 187–204.

Sakurai, Y., Onuma, Y., Nakazawa, G., Ugawa, Y., Momose, T., Tsuji, S., and Mannen, T. (2007). Parietal dysgraphia: characterization of abnormal writing stroke sequences, character formation and character recall. *Behav. Neurol.* 18, 99–114.

Salthouse, T. A. (1986). Perceptual, cognitive, and motoric aspects of transcription typing. *Psychol. Bull.* 99, 303–319.

Salthouse, T. A., and Saults, J. S. (1987). Multiple spans in transcription typing. *J. Appl. Psychol.* 72, 187–196.

Schwarzlose, R. F., Baker, C. I., and Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *J. Neurosci.* 25, 11055–11059.

Seghier, M. L., Fagan, E., and Price, C. J. (2011). Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. *J. Neurosci.* 30, 16809–16817.

Shallice, T. (1981). Phonological agraphia and the lexical route in writing. *Brain* 104, 413–429.

Sheldon, C. A., Malcolm, G. L., and Barton, J. J. (2008). Alexia with and without agraphia: an assessment of two classical syndromes. *Can. J. Neurol. Sci.* 35, 616–624.

Simon, O., Mangin, J. F., Cohen, L., Le Bihan, D., and Dehaene, S. (2002). Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. *Neuron* 33, 475–487.

Sugihara, G., Kaminaga, T., and Sugishita, M. (2006). Interindividual uniformity and variety of the "writing center": a functional MRI study. *Neuroimage* 32, 1837–1849.

Talairach, J., and Tournoux, P. (1988). *Co-planar stereotaxic-atlas of the human brain. 3-Dimensional Proportional System: Approach to Cerebral Imaging.* New York: Thieme Medical Publishers Inc.

Tanridag, O., and Kirshner, H. S. (1985). Aphasia and agraphia in lesions of the posterior internal capsule and putamen. *Neurology* 35, 1797–1801.

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

Tohgi, H., Saitoh, K., Takahashi, S., Takahashi, H., Utsugisawa, K., Yonezawa, H., Hatano, K., and Sasaki, T. (1995). Agraphia and acalculia after a left prefrontal (F1, F2) infarction. *J. Neurol. Neurosurg. Psychiatr.* 58, 629–632.

Tsapkini, K., and Rapp, B. (2010). The orthography-specific functions of the left fusiform gyrus: evidence of modality and category specificity. *Cortex* 46, 185–205.

Turkeltaub, P. E., and Coslett, H. B. (2011). Localization of sublexical speech perception components. *Brain Lang.* 114, 1–15.

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.

Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., and Fox, P. (2011). Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. *Hum. Brain Mapp.* doi: 10.1002/hbm.21186. [Epub ahead of print].

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., and Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289.

Weingarten, R. (2005). Subsyllabic units in written word production. *Writ. Lang. Lit.* 8, 43–61.

Weingarten, R., Nottbusch, G., and Will, U. (2004). "Morphemes, syllables and graphemes in written word production," in *Multidisciplinary Approaches to Language Production*, ed. T. H. Pechmann (Berlin, NY: Mouton de Gruyter), 529–572.

West, L. J., and Sabban, Y. (1982). Hierarchy of stroking habits at the typewriter. *J. Appl. Psychol.* 67, 370–376.

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

Received: 18 April 2011; accepted: 01 September 2011; published online: 11 October 2011.

Citation: Purcell JJ, Turkeltaub PE, Eden GF and Rapp B (2011) Examining the central and peripheral processes of written word production through meta-analysis. Front. Psychology ***2****:239. doi: 10.3389/fpsyg.2011.00239*

This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 21 February 2012
doi: 10.3389/fpsyg.2012.00038

Factors determining semantic facilitation and interference in the cyclic naming paradigm

Eduardo Navarrete[1]*, Paul Del Prato[2] and Bradford Z. Mahon[2,3]*

[1] Dipartimento di Psicologia dello Sviluppo e della Socializzazione, University of Padova, Padova, Italy
[2] Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY, USA
[3] Department of Neurosurgery, University of Rochester Medical Center, Rochester, NY, USA

Edited by:
Albert Costa, University Pompeu Fabra, Spain

Reviewed by:
David Howard, Newcastle University, UK
Gary Oppenheim, University of Illinois at Urbana-Champaign, USA

***Correspondence:**
Eduardo Navarrete, Dipartimento di Psicologia dello Sviluppo e della Socializzazione, University of Padova, Via Venezia 8, 35131 Padova, Italy.
e-mail: enavarrete2007@gmail.com;
Bradford Z. Mahon, University of Rochester, Meliora Hall, Rochester, NY 14627-0268, USA.
e-mail: mahon@rcbi.rochester.edu

The cyclic naming paradigm, in which participants are slower to name pictures blocked by semantic category than pictures in an unrelated context, offers a window into the dynamics of the mapping between lexical concepts and words. Here we provide evidence for the view that incremental adjustments to the connection weights from semantics to lexical items provides an elegant explanation of a range of observations within the cyclic naming paradigm. Our principal experimental manipulation is to vary the within-category semantic distance among items that must be named together in a block. In the first set of experiments we find that naming latencies are, if anything, faster for within-category semantically close blocks compared to within-category semantically far blocks, for the first presentation of items. This effect can be explained by the fact that there will be more spreading activation, and thus greater priming at the lexical level, for within-category semantically close blocks than within-category semantically far blocks. We test this explanation by inserting intervening filler items (geometric shapes), and show as predicted, that while intervening unrelated trials abolish short-lived semantic priming effects, the long-lag interference effect that is characteristic of this paradigm is unaffected. These data place new constraints on explanations of the cyclic naming effect, and related phenomena, within a model of language production.

Keywords: lexical access, speech production, semantic interference, semantic facilitation, semantic distance, cyclic naming paradigm

INTRODUCTION

Many object naming studies have explored lexical access in speech production by manipulating the semantic context within which speakers retrieve words from their mental lexicon. There are at least two motivations for that approach. The first is that lexical access in speech production is semantically driven. Thus, in the course of naming an object, speakers must access and select the semantic representation corresponding to the target name before initiating lexical access. The second motivation for manipulating the semantic context in which target pictures are named is because most extant theories of speech production model information flow between levels of processing in terms of spreading activation (e.g., Lupker, 1979; Dell, 1986; La Heij, 1988; Roelofs, 1992; Caramazza, 1997; Rapp and Goldrick, 2000; but see Bloem and La Heij, 2003). In the context of other assumptions about the dynamics of spreading activation (e.g., Dell, 1986; La Heij, 1988; Roelofs, 1992) current models assume that the amount of activation that one representation propagates to other linked representations is proportional to its level of activation. Thus, when naming a given picture (e.g., horse), semantically related concepts (e.g., ZEBRA, DOG, etc.) will become activated proportional to the degree to which they are semantically related to the target. In addition, most theoretical models also assume that the propagation between the semantic level and the lexical level follows the same principle. As a consequence, in the course of object naming, multiple lexical representations would be activated (i.e., the target word "horse" along with semantically related items, such as "zebra," and "dog").

According to the principle of spreading activation between semantic and lexical stages of processing, the prediction follows that any semantically related context should facilitate the semantic and lexical processing of target items. In line with this prediction, it is a well-established phenomenon in word recognition paradigms, such as lexical decision, that responses to target words are facilitated by semantically related primes (e.g., McRae and Boisvert, 1998). Such effects find a natural explanation in that the amount of activation that propagates from a prime to the target representation scales with the semantic similarity between the two words – the more similar they are, the more activation spreads to the target, and the faster will be the subsequent response to the target.

Semantic distance, or its inverse – semantic similarity – describes the semantic overlap between two items. Thus, closely related items (e.g., HORSE–ZEBRA) share more semantic information than less close (or far) related items (HORSE–SHARK); this may be modeled either in terms of the amount of overlap of semantic features (e.g., Dell, 1986) or the nature and number of links that connect different concepts (e.g., Roelofs, 1992). According to the principle of spreading activation, as the semantic distance between two concepts decreases, more activation would propagate between them. Thus, one would predict that semantic effects on target processing (e.g., horse) would be greater when a within-category close item (zebra) has just been processed compared to when a within-category far item has just been processed (whale).

The same situation obtains in speech production experiments: when prime and target stimuli are presented on consecutive trials, a facilitation effect can emerge. For instance, Huttenlocher and Kubicek (1983) reported faster naming latencies in object naming when the object named on the preceding trial was semantically related than when it was not related (see also Sperber et al., 1979; Humphreys et al., 1988; Lupker, 1988; Biggs and Marmurek, 1990; Hartsuiker et al., 2005; and for similar results without naming the first object see Flores d'Arcais and Schreuder, 1987; Bajo, 1988). The semantic facilitation effects reported in those studies would be congruent with models of lexical access that assume that the time to select the target word is independent of the level of activation of non-target words; or specifically, the time required to select a word depends only on its own level of activation, and is not affected for instance, by the activation level of the last word that was named (Dell, 1986; Caramazza, 1997; Rapp and Goldrick, 2000, for similar conclusions in bilingual lexical access, see Finkbeiner et al., 2006).

However, semantic interference effects are also observed in production, and in fact, have received far more attention and been given far more theoretical prominence than facilitation effects. For instance, Wheeldon and Monsell (1994) required participants to name pictures and written definitions of objects and observed that naming latencies to picture targets (e.g., "shark") were slower when some trials previously a semantic coordinate word (e.g., "whale") was produced as a response to a written definition, compared to when a non-semantic coordinate was previously produced (e.g., "volcano"). Similar semantic costs are reported if instead of naming written definitions, participants name objects (Vitkovitch et al., 2006). Furthermore, when several objects of the same semantic category are presented to be named, the amount of interference is cumulative, so that the amount of the delay observed for each additional instance of the category that is named depends on the total number of exemplars of the same category that have already been named (Brown, 1981; Howard et al., 2006; Costa et al., 2009; Navarrete et al., 2010). Semantic interference, as observed in those studies, has generally been interpreted to indicate that lexical selection is a competitive process in which the time required to select the target word is affected by the levels of activation of non-target words (Roelofs, 1992; Levelt et al., 1999). It is not clear, however, if the theory of lexical selection by competition would (in and of itself) ever "predict" that semantic interference should be observed when different instances from the same category are separated by intervening trials.

Thus, one may derive the following broad empirical generalization: a semantic relationship between the target on trial *n* and the picture named on trial $n-1$ leads to facilitation, while the same relationship that spans multiple trials leads to interference. Regarding the theoretical implications, semantic facilitation is consistent with models that assume that lexical selection depends on target activation only (that is, selection by activation models) but challenges models of lexical selection by competition; in contrast, semantic interference challenges selection by activation models and is consistent with competitive models.

Why do semantically related contexts lead to facilitation in recognition, facilitation in production with no intervening trials, and interference in production with intervening (unrelated) trials? As noted above, over the last two decades, the field has arrived at the view that interference effects in speech production are theoretically more informative than facilitation effects, in that interference (but not facilitation) effects inform a model of the dynamics of lexical selection (i.e., that it occurs by competition). The rationale in relation to the object naming studies described above is that the lexicalization process over an object (or definition) makes its corresponding lexical unit a stronger competitor when on a subsequent trial a semantically related object has to be named (Wheeldon and Monsell, 1994; Howard et al., 2006).

However, recent work indicates that the theory of lexical selection by competition is not without its problems. In particular, the theory must be reinforced against the range of semantic facilitation effects that have been observed, and which, *prima facie*, are contrary to its central prediction (for data and relevant discussions, see e.g., Roelofs, 1992, 2003; Caramazza and Costa, 2001; Damian et al., 2001; Bloem and La Heij, 2003; Costa et al., 2003; Costa et al., 2005; Bloem et al., 2004; Finkbeiner and Caramazza, 2006; Kuipers et al., 2006; Mahon et al., 2007). Even explanations of long-lag semantic interference effects in picture naming require some additional maneuvering for the theory, such as that activation "lingers" at the lexical level for some time. But the fact that trial-to-trial semantic relationships lead to facilitation and not interference indicates that, on that theory, activation cannot "linger" from the previous trial (see Howard et al., 2006). An alternative account of the inference effect proposes that there are incremental learning effects that are expressed as changes in the connection weights between semantic and lexical level representations. Thus, naming "horse" on a given trial will strengthen the connection between the concept HORSE and the lexical representation "horse" and, at the same time, weaken the connections between the concept and the lexical representations of semantic coordinates of the target (e.g., zebra, dog). On a subsequent trial on which "zebra" must be named, naming latencies will be slower because of the weaker connections (Oppenheim et al., 2010; see also Navarrete et al., 2010)[1]. The account in terms of incremental learning need not assume lexical selection by competition. An important issue, to which we return in the General Discussion, is how the incremental learning proposal of semantic interference, and more broadly the model of lexical selection in which it is embedded, can explain observations of semantic facilitation when there are no intervening filler items.

The main goal of the current research is to empirically characterize the boundaries of semantic interference and facilitation effects observed in the cyclic naming paradigm. In the standard version of this task participants are required to name a series of

---

[1] This is not to imply that there is one concept for the concepts "horse," "zebra," "whale," and so on – rather, on one way of modeling lexical semantics, concepts are decomposed into sets of features that are shared among items from the same category. Thus, strengthening and weakening effects occur because the same set of features is used to access one word, of they many words to which those features are connected. However, it is not unreasonable to imagine how such a mechanism could be implemented within models that do not assume features (e.g., Roelofs, 1992) and so we therefore remain agnostic on the representational structure of lexical semantics.

pictures several times in two semantic contexts. In the homogenous semantic context, objects presented in a block belong to the same superordinate semantic category (e.g., horse, dog, cat). In the heterogeneous semantic context, objects in a block belong to different superordinate semantic categories (e.g., horse, table, lemon). Under those conditions, naming latencies in homogenous blocks are on average slower than in heterogeneous blocks – the so-called semantic blocking, or cyclic naming, effect (Damian et al., 2001; see Kroll and Stewart, 1994 for the first demonstration of the effect). Of particular interest is that in the cyclic naming paradigm, participants are required to name semantically related objects on consecutive trials, that is, without interleaved unrelated items between them. This situation raises the empirical question of whether the effects described above from sequential object naming tasks (semantic interference and semantic facilitation) interact with one another in cyclic naming tasks. The answer is yes. Within the cyclic naming paradigm, for the first presentation of items in a block, there is either no effect at all or there is a semantic facilitation effect characterized by faster object naming latencies in homogeneous blocks than in heterogeneous blocks (e.g., Belke et al., 2005a,b; Damian and Als, 2005; Abdel Rahman and Melinger, 2007). In the current research we explored the pattern of semantic effects in the cyclic naming task by manipulating the semantic distance among the items of the homogeneous blocks, and exploring this manipulation as a function of the factors cycle (i.e., repetition within a block) and lag (i.e., the number of intervening trials between a given trial and the previously named within-category item).

THE CURRENT RESEARCH

In the present set of experiments we set out to test the effect of manipulating within-category semantic distance on the pattern of facilitation and interference effects observed within the cyclic naming paradigm. According to the hypothesis of lexical selection by competition, all else equal, more interference is expected when pictures are named in blocks of within-category close items compared to when the same pictures are named in the context of within-category far items (see a similar prediction in Alario and Moscoso del Prado Martín, 2010). According to selection by activation accounts of lexical selection, all else equal, naming latencies should be faster for within-category close contexts compared to within-category far contexts. However, and of central importance, is that "all else may not be equal," in two ways.

1) First we will seek to understand how the effect of manipulating within-category semantic distance interacts with the critical factors of cycle and lag. As noted above, the cyclic naming effect appears only in later cycles (repetitions) and, more generally, intervening trials are necessary in order for interference effects to emerge. That is, the cost in the cyclic naming paradigm is not created by the immediately preceding trial in the block, but rather by the contextual effect created by the picture(s) processed several trials before. If anything, the immediately preceding picture in a homogenous block will yield facilitation and not interference. Thus, in testing for an effect of within-category semantic distance, we are really testing for what happens when two effects of contrasting polarity are directly pitted against one another – a long-lasting interference effect and a trial-by-trial facilitation effect. We will attempt to tease apart these two effects with correlational analyses, and by interspersing non-critical and unrelated filler trials.
2) The second reason why "all else may not be equal," is that according to the incremental learning account of interference effects (Oppenheim et al., 2010), the degree to which connections are strengthened or weakened is proportional to the strength with which lexical items are activated by semantics. Thus, manipulating semantic distance will affect the degree to which those connection weights are altered. This means, that even on a model that does not assume lexical selection by competition, the prediction can be made that a within-category semantically close context may lead to slower naming latencies than a within-category semantically far context. The difference between this prediction as extrapolated from the incremental learning hypothesis, and as made by lexical selection by competition, is that the incremental learning hypothesis but not the hypothesis of lexical selection by competition predicts the within-category interference effect will be observed only for long lags.

Damian and Als (2005) reported evidence that semantic facilitation in the cyclic naming paradigm is a short-lasting phenomenon while semantic interference is a long-lasting phenomenon. In Experiment 4a of their study, the semantic blocking effect was absent for the first occurrence of the items but emerged thereafter, remaining stable for the remainder of the presentations. In Experiment 4b of their study, the same procedure as in Experiment 4a was used with the only difference that filler items from unrelated categories were dispersed throughout the blocks. According the authors, if the lack of the semantic interference for the first occurrence of the block was due to a mechanism of short-lasting semantic priming, such a mechanism should be attenuated when filler items are presented between the target items. Consistent with this, in Experiment 4b the semantic blocking effect was observed in all presentations of the target items, and most notably, the first presentation (i.e., the first cycle).

We thus had three goals with the current set of experiments. The first goal was to characterize the effect of varying within-category semantic contexts (within-category semantically close vs. within-category semantically far). This manipulation is important because it is a "pure" manipulation of semantic distance; unlike the contrast of homogenous vs. heterogeneous contexts, the manipulation of within-category semantic distance is not confounded with a manipulation of semantic category coordinate status. This allows us to test whether the incremental learning account can explain long-lag semantic interference. Two predictions can be outlined. If the mechanism underlying long-lasting semantic interference is incremental, then the prediction is made that within-category semantically close contexts will yield slower naming overall than within-category semantically far contexts. The reason why is that when the picture named on trial $n-2$ or $n-3$ was very close within-category (e.g., "zebra") to the picture on the current trial (e.g., "horse") then the connection between the semantic and lexical levels will be weaker than when the picture on a previous trial is relatively far within-category (e.g., whale) to the picture on the current trial ("horse"). If incremental learning is not the mechanism underlying long-lasting semantic interference, a model that does not assume lexical selection by competition predicts that within-category semantically close will be faster than within-category semantically far contexts, due to the presence of greater trial-to-trial priming (and no additional cost) created by the high semantic similarity among the pictures in a within-category semantically close block.

The second goal was to study the effect of intervening unrelated trials. Thus, in Experiments 1 and 2 there were no interleaved semantically unrelated filler items between the experimental items, while in Experiments 3a and 3b we included unrelated filler trials. The logic here is the same as in Damian and Als (2005), in that having interleaved unrelated trials should abolish short-lived semantic priming effects. By manipulating whether there are interleaved unrelated trials, we can directly study the emergence of a possible within-category semantic distance effect both "under the influence" of short-lived semantic priming and without short-lived semantic priming.

The third goal was to characterize the emergence of within-category semantic distance effects as a function of the factor cycle (i.e., repetition within a block). One prediction made by the hypothesis that the cyclic naming effect (homogenous vs. heterogeneous) is due to incremental learning, is that (see Goal 1 above) within-category semantically close will be slower than within-category semantic far. To anticipate the principal finding, that is the pattern that is obtained. Another prediction that is made is that the within-category semantic distance effect (close > far) will not be present in the first cycle. That is because trials that occur later in a block will have many long-lag within-category items preceding them within the block, while trials early in the block will have fewer such items. Finally, models of lexical access agree on the assumption that more spreading activation is expected between semantically close related items than semantically far related items. Therefore, if semantic facilitation instead of interference is observed in the first cycle (replicating previous studies, e.g., Abdel Rahman and Melinger, 2007) an empirical question is whether the facilitation effect would depend on semantic distance among the items of the homogenous blocks.

In Experiment 1 semantic distance was manipulated between items. Homogenous blocks were created in such a way as to cover a large range of semantic distance. In Experiments 2 and 3 semantic distance was manipulated within items and the same items were presented in two homogenous blocks: within-category close items and within-category far items. Furthermore, in Experiment 3a we choose pictures such that visual similarity among items in the within-category close condition was minimized so as to be comparable to the within-category far condition, while in Experiment 3b we choose pictures such that visual similarity was allowed to be greater among the within-category semantically close pictures than among the within-category semantically far pictures.

EXPERIMENT 1: MANIPULATING WITHIN-CATEGORY SEMANTIC DISTANCE BETWEEN ITEMS WITH NO INTERLEAVED UNRELATED TRIALS

Eight semantic categories were selected for Experiment 1. Objects were presented in two semantic contexts. In the homogeneous context, objects within the blocks belonged to the same semantic category; in the heterogeneous context, objects came from different categories. Experiment 1 had several goals. First, we aimed to replicate the interaction between semantic context and position within a block, with facilitation or no effect for the first presentation, and then interference emerging over later presentations within the block (see Level I analysis below). Second, we aimed to explore whether the semantic effects (interference and/or facilitation) are modulated by the semantic distance among the items of the homogenous blocks. Semantic distance was determined by semantic similarity ratings among the items that appeared together in blocks. Specifically, after the main naming experiment participants were asked to judge the semantic similarity between all possible pairs of items that had been presented together in the homogeneous and the heterogeneous blocks. Finally, correlation analyses between the magnitude of the semantic context effect and the semantic ratings were carried out on an item-by-item basis and for each occurrence of the items within the blocks (see Level II analysis below).

One could argue that semantically related objects share more visual features than unrelated objects, and that this could affect the overall pattern of naming latencies in heterogeneous and homogenous blocks (e.g., Lotto et al., 1999). Two measures were taken in order to mitigate the influence of visual variables in the results. First, before starting the experiment proper participants were exposed three times to all of the experimental pictures (for a similar procedure see Abdel Rahman and Melinger, 2007). This massive pre-exposure was intended to mitigate the role of object identification variables during the main naming experiment. Second, visual similarity between homogeneous and heterogeneous blocks was determined and analyzed through a computational approach to measuring similarity in visual shape.

METHOD

Participants

Twelve native Italian speakers (students at the University of Trento, Italy) took part in the experiment. Participants in this and subsequent experiments had normal or corrected to normal vision and participated in only one experiment.

Materials

Sixty-four black and white photographs depicting objects from 8 semantic categories were selected. Photographs in this and subsequent experiments were taken from the Internet and sized to fit within a square of 400 × 400 pixels. See Section "Appendix A" for a list of materials.

Design

Each picture presented in two naming conditions. In the Homogeneous condition pictures were paired with three semantic coordinate pictures. In the Heterogeneous condition, pictures were paired with three pictures from different semantic categories. There were a total of 32 blocks with 4 pictures in each block. Half of the blocks belonged to the homogeneous naming condition and the other half to the heterogeneous naming condition. Within each block, each of the four pictures was presented four times in random order with the constraint that the same picture never appeared on consecutive trials. Thus, there were a total of four cycles within each block, corresponding to each of the four occurrences of the pictures. In this design a cycle would not necessarily coincide with four consecutive trials. Following Damian et al.'s (2001) design, the 32 experimental blocks were presented in a ABBA design: half of the participants started with 8 semantically related blocks followed by the 16 unrelated blocks and finished with the remaining 8 semantically related blocks; the other half of the participants started with 8 unrelated blocks, followed by the 16 semantically related blocks, and finished with the remaining unrelated blocks. There was a short pause of several seconds between each block.

Visual similarity among the items within homogeneous and heterogeneous conditions was assessed computationally. Each picture in each block was compared with the other three pictures of the same block. Similarity in visual shape was determined using the algorithm of Belongie et al. (2002), which computes the "cost" that would be required to warp an image into the shape of another image (see Mahon et al., 2007, for precedent on the use of this algorithm). By blocks, averages of visual similarity values for the items within the 16 homogenous blocks (mean $= 0.132 \pm 0.03$) were not different than the averages of visual similarity values for the items within the 16 heterogeneous blocks (mean $= 0.129 \pm 0.03$; $t < 1$).

Procedure

Participants were seated approximately 60 cm from the screen. The experimental session was divided into four parts. First, there was a familiarization phase in which each picture was presented with the corresponding name and participants were instructed to read the name while paying attention to the picture. Second, there was a training phase. The training consisted of 2 blocks containing the 64 pictures presented in a random order. During this phase participants were corrected by the experimenter when necessary. Third, was the experimental phase. In this phase each participant was presented with the 32 blocks and asked to name the pictures as fast and as accurately as possible. The fourth phase was a norming study in which participants were required to judge the semantic similarity between all possible pairs of items that had appeared together within the blocks.

Each trial within the experimental naming phase consisted of the following events. A fixation cross was shown in the center of the screen for 500 ms and was followed by a blank interval of 500 ms. The picture was presented for 500 ms. Response latencies were measured from the onset of the picture. The next trial started 1500 ms after the onset of participants' response or 3000 ms after the offset of the target. Stimulus presentation, response times, and response recording were controlled by the program DMDX (Forster and Forster, 2003).

For the semantic norming study words instead of pictures were used. Participants were presented with two words corresponding to pictures they had seen and were instructed to judge the semantic similarity between them (from 1 = not related at all; to 7 = very related). Each item was paired with all of the items with which it had appeared in a block (homogeneous and heterogeneous). Thus, each target word was presented a total of six times: three with semantically related items and three with semantically unrelated items. There were a total of 384 trials. The order of presentation of the trials and the position of the words (left or right) was random for each participant. The semantic rating study started after 20 practice trials using filler words.

Analyses

Four types of responses were excluded from the analyses of response times: (a) production of clearly erroneous picture names; (b) verbal disfluencies (stuttering, utterance repairs, and production of non-verbal sounds that triggered the voice key); (c) naming latencies less than 250 ms or greater than 1500 ms; and (d) the first trial of each block. A total of 7.9% of the data points were excluded following those criteria.

Two different types of analysis were carried out. In the Level I analysis, two within-subject factors, Semantic Context (two levels: Homogenous and Heterogeneous) and Cycles within blocks (four levels: 1–4), and their interaction were modeled. In the Level II analysis we explored whether the semantic effects were affected by the semantic similarity between the items of the homogeneous blocks. We did that for each cycle and on an item-by-item basis. In order to avoid item-intrinsic properties affecting correlation values (as for instance, lexical frequency, or age of acquisition), the semantic blocking effect was calculated for each specific item by subtracting the latency in the heterogeneous block from the latency in the homogenous block. These differences, on an item-by-item basis, were then correlated with the semantic rating value of each item obtained in the related condition of the norming study (that is, the mean semantic rating obtained by comparing the semantic similarity of one specific item with all the other items of the homogenous block).

RESULTS AND DISCUSSION

Semantic ratings for homogenous and heterogeneous conditions were different (homogenous mean $= 5.58 \pm 1.31$; heterogeneous mean $= 1.62 \pm 0.97$; $t(30) = 26.44$, $p < 0.01$).

Level I analysis

Separate analyses were carried out treating subjects and items as random factors, yielding $F1$ and $F2$ statistics, respectively. For all analyses in this article, degrees of freedom were Greenhouse–Geisser corrected when the assumption of Sphericity was violated.

Mean naming latencies and error rates by condition are reported in **Table 1** (see **Figure 1**). The analysis of naming latencies showed a main effect of Semantic Context [$F1(1, 11) = 10.98$, $p < 0.01$, $\eta^2 = 0.54$; $F2(1, 63) = 29.19$, $p < 0.001$, $\eta^2 = 0.31$], with slower response times for homogeneous blocks than for heterogeneous blocks. The main effect of Cycle was also significant [$F1(3, 33) = 143.82$, $p < 0.01$, $\eta^2 = 0.92$; $F2(1.83, 115.59) = 143.17$, $p < 0.001$, $\eta^2 = 0.69$], with decreasing response times with each additional cycle. The interaction between Semantic Context and Cycle was significant [$F1(3, 33) = 20.26$, $p < 0.001$, $\eta^2 = 0.64$; $F2(2.2, 139.86) = 18.89$, $p < 0.01$, $\eta^2 = 0.23$]. This interaction reflects a polarity shift of the semantic effect between the first cycle and the remaining three. Paired t-tests revealed a semantic facilitation effect in the first cycle [$t1(11) = -2.31$; $p < 0.05$; $t2(63) = -2.91$; $p < 0.01$] and semantic interference in the remaining cycles [Cycle 2: $t1(11) = 4.29$; $p < 0.01$; $t2(63) = 5.71$; $p < 0.01$; Cycle 3: $t1(11) = 5.25$; $p < 0.01$; $t2(63) = 4.75$; $p < 0.01$; Cycle 4: $t1(11) = 3.31$; $p < 0.01$; $t2(63) = 5.58$; $p < 0.01$].

Table 1 | Mean naming latencies (RT), SDs in ms, and percentage of error rates (E) by semantic context and cycle in Experiment 1.

Cycle	Semantic context					
	Homogenous			Heterogeneous		
	RT	SD	E	RT	SD	E
1	640	70	2.7	655	66	2.5
2	579	57	1.7	548	61	1.8
3	573	58	2.6	544	61	1.3
4	569	57	2	542	66	0.9
Mean	587		2.2	567		1.6

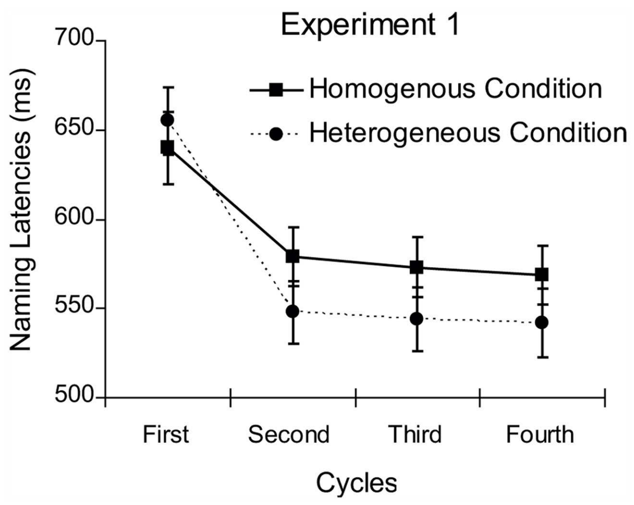


FIGURE 1 | Mean naming latencies by Semantic Context and Cycle for Experiment 1.

In the analysis of error rates the main effect of Semantic Context was not significant [$F1(1, 11) = 1.9, p = 0.19, \eta^2 = 0.14$; $F2(1, 63) = 1.41, p = 0.23, \eta^2 = 0.02$], and the main effect of Cycle was significant only in the item analysis [$F1(2.04, 22.46) = 1.8, p = 0.18, \eta^2 = 0.14$; $F2(2.6, 163.86) = 4.42, p < 0.01, \eta^2 = 0.06$]. The interaction between these two factors was not significant ($Fs < 1$).

Level II analysis

The correlation between the semantic interference effect (RT homogenous–RT heterogeneous) and semantic ratings in the related condition was negative in Cycle 1 ($r = -0.27$, $p < 0.04$, two tailed) and positive in the remaining cycles (Cycle 2: $r = 0.25$, $p < 0.05$; Cycle 3: $r = 0.22$, $p = 0.08$; Cycle 4: $r = 0.39$, $p < 0.01$; all two tailed).

Experiment 1 replicated the previous observation that in the first cycle, semantic facilitation is observed, with the interference effect emerging only for later cycles within the block. Of particular importance, however, is that variation in naming latencies as a function of within-category semantic distance followed the same pattern. For the first cycle, pictures in within-category semantically close contexts were named faster than pictures in within-category semantically far contexts. However, by the second cycle this effect completely reversed, in parallel to the emergence of the semantic blocking effect itself (i.e., related vs. unrelated). The design of Experiment 1 explored within-category semantic distance effects between items, as items appeared only once in a semantic homogenous block, and semantic distance was estimated across different homogenous blocks containing different items. In the next experiments we manipulated within-category semantic distance within items by presenting the same item in two homogenous blocks (one with within-category semantically close items and another with within-category semantically far items).

EXPERIMENT 2: MANIPULATING WITHIN-CATEGORY SEMANTIC DISTANCE WITHIN ITEMS WITH NO INTERLEAVED UNRELATED TRIALS

In this experiment the same pictures were presented in three different semantic contexts: within-category semantically close, within-category semantically far, and unrelated (or heterogeneous) blocks. In order to avoid strategic preparation of the response, item order within cycles was completely randomized and the duration of the fixation point was jittered. The same types of analyses as in Experiment 1 were carried out.

METHOD

Participants

Twelve native English speakers (students at the University of Rochester) took part in the experiment.

Materials

Fifty-four black and white photographs were selected. Items were organized in two groups containing three categories with nine items per category. Group 1 contained the categories animals, tools, and fruits/vegetables, and Group 2 contained the categories animals, tools, and vehicles. Each group was arranged in a 3 × 3 matrix, where each row contained a grouping of three highly semantic similar items (e.g., dog, wolf, fox) and formed the within-category close condition, and each column contained three less similar items (e.g., dog, mouse, lizard) and formed the within-category far condition. The heterogeneous condition was formed by selecting three items from different superordinate categories (e.g., dog, fork, orange). Phonological similarity was reduced or eliminated among all items that appeared together in a block (see Appendix B for all materials). Furthermore, pictures were chosen so as to eliminate, if possible, the tendency for within-category semantically close groupings to be more visually similar than within-category semantically far groupings of items.

DESIGN

Pictures were grouped into blocks that were either within-category close (e.g., dog, wolf, and fox), within-category far (e.g. dog, mouse, and lizard), or heterogeneous (e.g. dog, fork, and orange). Items were repeated three times in each block, and blocks were therefore nine trials long. Item order within block was random. The first occurrence of each item was denoted as cycle 1, the second occurrence as cycle 2, and the last as cycle 3. As item order within the block was completely random, the same item could be presented on two consecutive trials. There were a total of 54 blocks, presented in a different random order for each participant. Participants completed the 486 trials in approximately 35 min, with intermittent break periods between each block.

Procedure

The procedure was the same as in Experiment 1 with two differences: there was no training phase and the trial structure was different. On each trial a fixation cross was presented for either 160, 240, 320, 400, or 480 ms. This jittering was introduced so that participants would not anticipate the timing of picture onset. Following the fixation cross was a blank period of 160 ms followed by the target picture. Targets were presented for 480 ms. The offset of the target picture was followed by a blank screen for 992 ms, during which time participants responded. There were a total of 486 trials. Participants completed a semantic distance rating study following completion of the main experiment. For the rating study, words instead of pictures were used: participants were presented with two words corresponding to pictures they had seen and were instructed to judge the semantic similarity between them (from 1 = not related; to 7 = very related). Each item was paired with all of the items with which it had appeared in a block (within-category semantically close, within-category semantically far, and unrelated). Thus, each target word was presented a total of six times: two with semantically close related items, two with within-category semantically far items, and two with semantically unrelated items. There were a total of 162 trials. The order of presentation of the trials was randomized for each participant.

RESULTS AND DISCUSSION

The mean semantic ratings between all three semantic contexts were different [using the averages for blocks: close vs. far: $t(34) = 9.55, p < 0.01$; close vs. unrelated: $t(34) = 42.88, p < 0.01$: far vs. unrelated: $t(34) = 15.61, p < 0.01$]. The mean ratings were 6.0, 4.3, and 1.8 for close, far, and unrelated contexts, respectively (on a seven point scale).

Analysis level I

The same analyses as in Experiment 1 were performed. A total of 17.6% of the data points were excluded following the same criteria as were used in Experiment 1. In the analysis of naming latencies (see **Table 2**; **Figure 2**), the main effect of Semantic Context was not significant ($Fs < 1$), while the main effect of Cycle was significant, [$F1(2, 22) = 146.41$, $p < 0.01$, $\eta^2 = 0.93$; $F2(1.41, 75.08) = 219.7$, $p < 0.01$, $\eta^2 = 0.8$], with response times decreasing for each cycle. The interaction between Semantic Context and Cycle was significant [$F1(4, 44) = 3.66$, $p < 0.02$, $\eta^2 = 0.25$; $F2(3.12, 165.39) = 2.92$, $p < 0.04$, $\eta^2 = 0.05$]. Paired samples t-tests between the within-category semantically close and the heterogeneous condition showed a semantic facilitation effect in the first cycle, with response times faster in the within-category semantically close condition than in the heterogeneous condition [$t1(11) = -2.5$, $p < 0.03$; $t2(53) = -2.01$, $p < 0.05$]. No significant differences were observed in the second and third cycles [Cycle 2: $t1(11) = 1.7, p = 0.1$; $t2(53) = 1.8, p = 0.08$; Cycle 3: $t1(11) = 1.4$, $p = 0.18$; $t2(53) = 1.53$, $p = 0.13$]. Paired samples t-tests between the within-category semantically far and the heterogeneous condition indicated no significant differences [Cycle 1: $t1(11) = -1.51$, $p = 0.15$; $t2(53) = -1.33$, $p = 0.18$; Cycle 2: $t1(11) = 1.18$, $p = 0.26$; $t2(53) = 1$, $p = 0.32$; Cycle 3: $t1(11) = 1.59$, $p = 0.13$; $t2(53) = 1.79$, $p = 0.07$]. No significant differences were observed between the within-category semantically close and the within-category semantically far conditions ($ts < 1$).

Table 2 | Mean naming latencies (RT), SDs in ms, and percentage of error rates (E) by semantic context and cycle in Experiment 2.

Cycle	Semantic context								
	Close related			Far related			Heterogeneous		
	RT	SD	E	RT	SD	E	RT	SD	E
1	668	48	11.3	676	57	12.7	689	56	9.7
2	580	55	6.7	576	48	6	568	51	5.6
3	581	45	7.4	582	52	6.4	572	50	4.4
Mean	610		8.5	611		8.4	610		6.6

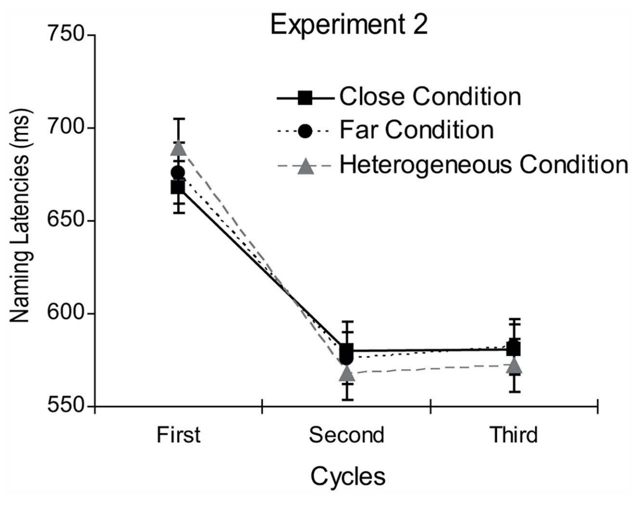


FIGURE 2 | Mean naming latencies by Semantic Context and Cycle for Experiment 2.

There was no semantic interference effect comparing either the within-category semantically close or the within-category semantically far conditions to the unrelated baseline. In order to explore whether there was a semantic interference effect averaging over the two semantically related conditions, cycles 2 and 3 were collapsed, and related was compared to unrelated. A semantic blocking effect was obtained, with slower naming latencies in the semantically related condition than in the unrelated condition [580 and 570 ms, respectively; marginally significant by subjects, $t1(11) = 1.98$; $p = 0.07$; $t2(53) = 2.25$; $p < 0.05$].

In the analysis of error rates the main effect of Semantic Context was significant in the subject analysis [$F1(2, 2) = 4.48$, $p < 0.03$,

$\eta^2 = 0.29$] and marginally significant in the item analysis [$F2(1.79, 95.27) = 2.72$, $p = 0.07$, $\eta^2 = 0.04$]. The main effect of Cycle was not significant [$F1 < 1$; $F2(2, 106) = 1.62$, $p = 0.2$, $\eta^2 = 0.03$]. The interaction between these two factors was not significant ($Fs < 1$).

Level II analysis

The same analyses as in Experiment 1 were performed with the difference that now each item had two data points, one corresponding to the within-category semantically close vs. heterogeneous comparison and one corresponding to the within-category semantically far vs. heterogeneous comparison. Correlations were not significant ($ps > 0.2$).

As in Experiment 1, a semantic facilitation effect in the first cycle (for the within-category semantically close condition only) and an interaction between Semantic Context and Cycle were observed. In Experiment 2 no significant correlations were observed. The semantic blocking effect was observed only when collapsing cycles 2 and 3 together and collapsing the within-category close and far conditions together, and comparing related to unrelated.

EXPERIMENT 3: MANIPULATING WITHIN-CATEGORY SEMANTIC DISTANCE WITHIN ITEMS WITH INTERLEAVED UNRELATED TRIALS

Damian and Als (2005) reported that the interaction between semantic relationship and item repetition disappears when interleaved filler trials are presented in the first cycle of the block. According to those authors, this happens because the short-lasting semantic priming mechanism in the first cycle is attenuated by the presence of filler items. The same experimental blocks as in Experiment 2 were used in Experiment 3 with the difference that unrelated distractor objects (geometrical shapes) were intermixed between the target objects of the blocks. Following, Damian and Als' interpretation, both the semantic facilitation effect in the first cycle of the within-category close condition and the interaction between semantic context and item repetition found in Experiment 2, should disappear in Experiment 3. To ensure that any observed effects could not be explained by a concomitant manipulation of visual variables along with the manipulation of within-category semantic distance, two different versions of Experiment 3 were constructed. In Experiment 3a we used the same photographs as in Experiment 2 (where visual similarity was minimized for within-category close blocks), and in Experiment 3b we selected a new set of photographs of the same items with high visual similarity among the pictures in the within-category close condition. Previous research indicates that higher visual similarity among pictures will slow down response times (Lotto et al., 1999, see also for a specific test in the cyclic paradigm, Belke et al., 2005b). Thus, it is important to be able to rule out, or at least reduce the possibility of, a contribution to a within-category distance effect of a confound of visual similarity with within-category semantic distance.

METHOD

Participants

Twenty-four native English speakers (students at the University of Rochester) took part in the experiment. Twelve participated in Experiment 3a and 12 in Experiment 3b.

Materials

The same experimental items used in Experiment 2 were used here. In order to create the filler condition, three unrelated objects (the geometrical shapes of a circle, a square, and a triangle) were selected and presented intermixed within the experimental objects in the blocks. Each experimental block contained 18 trials (nine objects and nine shape trials). The nine trials of objects and the nine trials of shapes corresponded to three repetitions (cycles) of presentation of the objects and shape images.

Design and procedure

The order of pictures and shapes within a block was random with the constraint that the first six trials contained the three object pictures and the three shapes, the second group of six trials contained the three objects and the three shapes, and the same for the final six trials. The first occurrence of the experimental items is denoted cycle 1, the second cycle 2, and the last cycle 3. The same picture was never presented on adjacent trials. There were a total of 54 blocks, presented in a different random order for each participant. Participants completed the 972 trials in approximately 1 h, with intermittent break periods between each block. The same procedure as in Experiment 2 was used here.

RESULTS AND DISCUSSION

Detailed analyses were performed by object and shape trials separately. We first present the analysis for the object trials. Following the same procedure as in previous experiments, 13.1% of the trials were removed from the analysis.

Level I analysis

In the analysis of naming latencies (see **Table 3**; **Figure 3**) there was a main effect of Semantic Context [$F1(2, 44) = 20.61$, $p < 0.01$, $\eta^2 = 0.48$; $F2(2, 106) = 35.14$, $p < 0.01$, $\eta^2 = 0.39$] and a main effect of Cycle [$F1(1.47, 32.51) = 135.03$, $p < 0.01$, $\eta^2 = 0.86$; $F2(1.45, 77.25) = 214.74$, $p < 0.01$, $\eta^2 = 0.8$]. The main effect of Experiment was significant [$F1(1, 22) = 4.15$, $p = 0.05$, $\eta^2 = 0.15$; $F2(1, 53) = 152{,}4$, $p < 0.01$, $\eta^2 = 0.74$]. The interaction between Semantic Context and Cycle was significant [$F1(4, 88) = 5.48$, $p < 0.01$, $\eta^2 = 0.2$; $F2(2.93, 155.78) = 6.21$, $p < 0.01$, $\eta^2 = 0.1$].

Table 3 | Mean naming latencies (RT), SDs in ms, and percentage of error rates (E) by semantic context and cycle in Experiment 3a and 3b.

Cycle	Close related			Far related			Heterogeneous		
	RT	SD	E	RT	SD	E	RT	SD	E
EXPERIMENT 3A (VISUALLY DISSIMILAR ITEMS)									
1	698	60	15.3	690	53	14.4	682	55	16.4
2	642	47	10.4	630	48	6.8	611	45	8.1
3	644	57	10.3	620	43	7.7	601	41	6.3
Mean	661		12	647		9.6	631		10.3
EXPERIMENT 3B (VISUALLY SIMILAR ITEMS)									
1	757	80	11.9	756	72	10.6	753	79	12.8
2	712	104	9.2	681	94	7	670	105	5.3
3	699	93	7.6	674	110	7.3	661	109	5.9
Mean	723		9.6	704		8.3	695		8

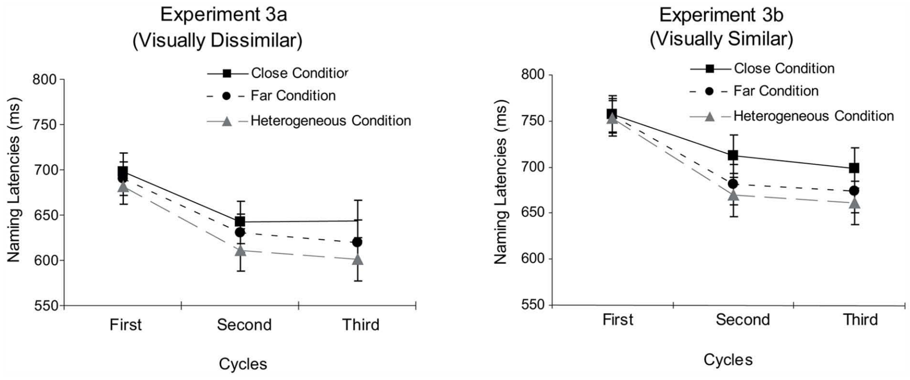


FIGURE 3 | Mean naming latencies by Semantic Context and Cycle for Experiments 3a and 3b.

No other interactions were significant ($ps > 0.16$). In order to provide a direct comparison with Experiment 2, the interaction between Semantic Context and Cycle was analyzed separately for Experiment 3a, confirming that for this Experiment the interaction was significant [$F1(4, 44) = 2.84$, $p < 0.04$, $\eta^2 = 0.2$; $F2(3.08, 163.56) = 3.35$, $p < 0.02$, $\eta^2 = 0.06$].

In order to further explore the interaction between Semantic Context and Cycle we conducted paired t-tests between all three semantic conditions. As there was no interaction between the factor Experiment and the other two factors, data points were collapsed across the factor Experiment. Comparisons between the semantically close and the heterogeneous conditions revealed no effect in the first cycle [$t1(23) = 1.64$, $p = 0.11$; $t2(53) = 1.07$, $p = 0.28$], and semantic interference effects in the second and third cycles [Cycle 2: $t1(23) = 5.33$, $p < 0.01$; $t2(53) = 7.09$, $p < 0.01$; Cycle 3: $t1(23) = 5.95$, $p < 0.01$; $t2(53) = 7.71$, $p < 0.01$]. The same pattern was observed for the within-category semantically far condition: no effect in the first cycle ($ts < 1$), and semantic interference effects in cycles 2 and 3 [Cycle 2: $t1(23) = 2.95$, $p < 0.01$; $t2(53) = 3.23$, $p < 0.01$; Cycle 3: $t1(23) = 4.66$, $p < 0.01$; $t2(53) = 3.12$, $p < 0.01$]. The comparison between the within-category semantically close and far conditions was not significant in the first cycle ($ts < 1$) but significant in the last two cycles [Cycle 2: $t1(23) = 3.89$, $p < 0.01$; $t2(53) = 4.94$, $p < 0.01$; Cycle 3: $t1(23) = 3.9$, $p < 0.01$; $t2(53) = 4.95$, $p < 0.01$], with slower response times in the within-category semantically close condition.

In the analysis of error rates the main effect of Semantic Context was significant [$F1(2, 44) = 3.52$, $p < 0.04$, $\eta^2 = 0.13$; $F2(2, 106) = 5.59$, $p < 0.01$, $\eta^2 = 0.09$]. The main effect of Cycle was also significant [$F1(1.22, 27.01) = 279.27$, $p < 0.01$, $\eta^2 = 0.92$; $F2(1.41, 75.05) = 30.04$, $p < 0.01$, $\eta^2 = 0.36$]. The main effect of Experiment was significant by items only [$F1 < 1$; $F2(1, 53) = 7.62$, $p < 0.01$, $\eta^2 = 0.12$]. The interaction between Semantic Context and Cycle was significant [$F1(4, 88) = 2.87$, $p < 0.03$, $\eta^2 = 0.16$; $F2(4, 212) = 2.54$, $p < 0.05$, $\eta^2 = 0.04$]. The interaction between Cycle and Experiment was significant in the item analysis only [$F1 < 1$; $F2(2, 106) = 4.23$, $p < 0.02$, $\eta^2 = 0.07$]. The interaction between Semantic Context and Experiment was not significant ($Fs > 1$).

Analysis of shape trials

As there were only three filler items (three geometrical shapes), analyses of the data from shape trials were performed only by Subjects. The analysis of naming latencies for shape trials showed no significant effects ($ps > 0.18$) for any of the factors. Mean latencies were 630 ms in the close condition, 632 ms in the far condition and 627 ms in the unrelated condition; and 627, 633, and 629 ms for the first, second, and third cycles, respectively. The analysis of error rates for shape trials also showed no significant effects ($ps > 0.13$).

Level II analysis

The same correlation analyses as in Experiment 2 were performed here. The correlation was not significant in the first Cycle ($p = 0.7$) and significantly positive in Cycles 2 ($r = 0.42$, $p < 0.01$) and Cycle 3 ($r = 0.38$, $p < 0.01$).

Several findings were reported in Experiment 3. First, a semantic blocking effect was reported; the presence of this effect is in contrast to Experiment 2 in which the same materials were used with the difference that there were no interleaved filler trials (geometrical shapes). Second, the magnitude of the semantic blocking effect was independent of the visual similarity of the pictures within the blocks, suggesting that the nature of the effect does not have a perceptual locus, consistent with previous work (e.g., Damian et al., 2001, but see Belke et al., 2005b, for effects of visual similarity in the same paradigm). Third, the semantic facilitation effect reported in the first cycle of the Close condition of Experiment 2 was absent in Experiment 3. The lack of the facilitation effect in the first cycle is congruent with Damian and Als (2005), suggesting that the presence of interleaved unrelated trials countermands the short-lasting facilitation effects. The final observation from Experiment 3 is that the effects were restricted to the object naming trials; the interleaved filler shape naming trials were unaffected by both context and cycle, replicating the pattern obtained by Damian and Als (2005)[2].

---

[2] However, contrary to the formal prediction of those authors, the interaction between the semantic blocking effect and cycle obtained in Experiment 2, without interleaved unrelated trials, was still present in Experiment 3. An explanation of this could be related to the differences between our design and that of Damian and

Analysis of lag

In a further analysis we explored the influence of lag on the pattern of results. Even though this analysis is *post hoc* within the context of Experiment 3, it is of critical theoretical importance, since as discussed above, different theories make different predictions about short-lived facilitation vs. long-lived interference effects. Specifically, models that assume that lexical selection depends on target activation would predict semantic facilitation, while models that assume lexical selection by competition would predict semantic interference (see Introduction). In Experiment 3, cycles were six items long (three objects and three shapes). Thus, the minimum lag between two objects was 0 (i.e., two different objects presented on consecutive trials with no intervening filler trials) and the maximum lag could be 6 (i.e., three shape trials of cycle *n*, followed by the three shape trials of cycle $n + 1$ separating the two different within-category items). As the presentation of the items within a block was pseudo-random, the total number of observations per each lag value varied. Specifically, the number of observations per Lag were: lag $0 = 3961$, lag $1 = 2955$, lag $2 = 1731$, lag $3 = 628$, lag $4 = 123$, lag $5 = 23$, and lag $6 = 16$ observations. In order to perform the analysis of the factor lag with a similar number of observations per cell, we collapsed lag values into three bins (lag 0, lag 1, and lag ≥ 2). Thus, the analysis had three factors (Semantic Context, Cycle, and Lag) with three levels each. Because any given item would not contribute equally to all levels of the Factor Lag for a given subject, item analyses are contraindicated and we therefore performed all analyses treating subjects as random factor.

In the analysis of naming latencies there were main effects of the three factors: Lag [$F1(2, 46) = 8.25$, $p < 0.01$, $\eta^2 = 0.26$], Semantic Context [$F1(2, 46) = 17.16$, $p < 0.01$, $\eta^2 = 0.42$], and Cycle [$F1(1.46, 33.65) = 162.66$, $p < 0.01$, $\eta^2 = 0.87$]. The interaction between Lag and Semantic Condition [$F1(4, 92) = 2.15$, $p = 0.08$, $\eta^2 = 0.08$] was marginally significant. The interaction between Lag and Cycle was not significant [$F1(2.49, 57.3) = 2.13$, $p = 0.11$, $\eta^2 = 0.08$]. Finally, the interaction between Semantic Condition and Cycle was (again) significant [$F1(4, 92) = 3.63$, $p < 0.01$, $\eta^2 = 0.13$; see **Figure 4**].

In the analysis of error rates, the main effect of Lag was not significant [$F1(2, 46) = 1.62$, $p = 0.2$, $\eta^2 = 0.06$], the main effect of Semantic Context was marginally significant [$F1(2, 46) = 3.06$, $p = 0.56$, $\eta^2 = 0.11$], and the main effect of Cycle was significant [$F1(1.17, 27.12) = 20.56$, $p < 0.01$, $\eta^2 = 0.47$]. No interactions were significant (*p*s > 0.13).

Modulation of repetition priming by the identity of an intervening trial

In a final analysis, we studied the situation in which there was one intervening trial between repeats of the same picture. That intervening trial could be either an object or a shape. This presents an interesting situation, and allows further confirmation of our principal hypothesis. The general expectation is that repetition of a picture will lead to a decrease in response times. The question, however, is whether that decrease is modulated by context, and whether that contextual modulation is affected by whether the interleaved trial was an object or a shape. The results of this analysis are plotted in **Figure 5**. As can be seen, collapsing across the factor context, the overall amount of repetition priming is similar when the intervening trial is an object or a shape. However, and of particular interest, is that there is a clear polarity reversal along the dimension of within-category semantic distance according to the nature of the intervening stimulus. Specifically, when the intervening stimulus is a shape, the most repetition priming is observed

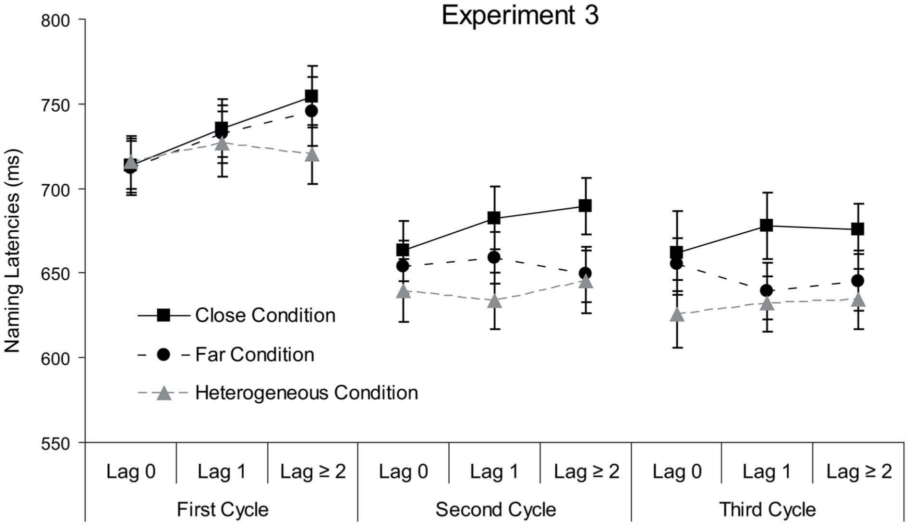


FIGURE 4 | Mean naming latencies by Semantic Context, Cycle, and Lag for Experiment 3.

Als (2005): (a) Damian and Als used eight items per block (four experimental, four filler) while we used only six (three experimental, three filler); (b) in their design, objects were presented under two semantic contexts (homogeneous and heterogeneous) while in our experiment there were three semantic contexts (close, far, and heterogeneous); (c) filler items in their study were from a pool of 16 pictures from diverse semantic categories while we choose 3 pictures corresponding to geometrical shapes.

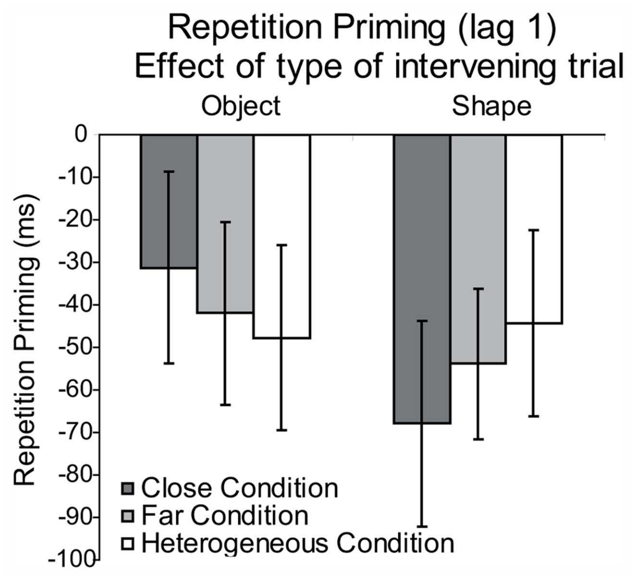


FIGURE 5 | Repetition Priming by Semantic Context and Cycle for Experiment 3 (in Lag 1).

for the within-category semantically close condition, followed by the within-category semantically far and finally the unrelated condition. This pattern is entirely inverted when the intervening trial was an object. These data, while inviting experiments specifically designed to test for these effects, strongly suggest that response times on a given trial in the cyclic naming paradigm are a composition of short-lived semantic facilitation and long-lag semantic interference.

GENERAL DISCUSSION

Several findings have been reported:

1) In the first cycle naming a picture is faster in a categorically related context (with no intervening unrelated trials) compared to an unrelated context. This was observed in the first cycle of Experiment 1 and in the first cycle of the within-category close condition of Experiment 2. This semantic facilitation replicates previous observations (e.g., Abdel Rahman and Melinger, 2007) and is comparable with other picture naming paradigms that report facilitation effects in object naming when the immediately preceding naming trial is categorically related (e.g., Huttenlocher and Kubicek, 1983). Importantly, our findings (Experiment 2) further show that this polarity reversal from facilitation to interference exists when semantic distance is manipulated within-category, as it does when comparing related to unrelated.
2) Further repetition of items within a block generates a semantic interference effect; that is, naming latencies are slower in semantically homogenous blocks than in heterogeneous blocks for later cycles. This is shown by the interference effects in cycles 2 through 4 of Experiment 1, and cycles 2 and 3 of Experiment 2, replicating previous cyclic naming studies (e.g., Belke et al., 2005b).
3) The facilitation effect in the first cycle disappears when unrelated filler naming trials are embedded within the object naming trials, while the semantic interference effect that emerges for later cycles remains stable. This was observed comparing Experiments 2 and 3, and replicates Damian and Als' (2005) study.
4) Increasing the visual similarity among the items within a block slows down the overall naming latencies but does not influence the semantic effects (as reported for the comparisons between Experiments 3a and 3b); the first observation replicates previous studies on visual influences on object naming (Lotto et al., 1999), while the second observation suggests that the semantic effects in the cyclic naming task can be placed at a post-perceptual level of processing (Damian et al., 2001)[3].

In summary, our results show that the way in which speaking is affected by semantic context depends on the joint influence of multiple factors. The theory of lexical selection by competition is not able to explain the fact that in the first cycle, a semantic facilitation effect (semantically related < unrelated) as well as a facilitatory effect of within-category semantic distance (close < far) is observed (Experiments 1 and 2). In this context, some authors have argued that facilitation effects suggest that lexical competition "*needs some potentiation, through repetition of the items, to become observable*" (Belke et al., 2005b, p. 687; for a similar argument see also, Abdel Rahman and Melinger, 2007); others argued that there is some type of a short-term semantic facilitation effect (Wheeldon and Monsell, 1994). But why would this be the case? Merely asserting this amounts to not much more than a redescription of the data.

Another approach is to pursue an explanation in terms of the incremental learning model outlined in the Introduction. Oppenheim et al. (2007, 2010) have proposed an interpretation of semantic interference effects in naming tasks that applies an error-driven learning mechanism. Naming an object would enhance the connectivity between the semantic and the lexical representations corresponding to the target picture, and, at the same time, it would weaken the connections to semantically related lexical units. These two aspects of the model can account, according to Oppenheim et al. (2010), for repetition priming effects and semantic interference effects in object naming tasks (for similar arguments see

---

[3] The differences between the within-category semantically close and within-category semantically far conditions has been referred to as a "graded" effect. So-called graded semantic distance effects in the blocked naming task were investigated by Vigliocco et al. (2002). However, in that study semantic distance was manipulated between category and not within-category as we did here. For instance, target pictures (e.g., items from the category "clothing") were presented in blocks mixed with pictures belonging to a different category that was semantically close to the target category (e.g., "body parts") or semantically far from the target category (e.g., "vehicles"). Vigliocco and colleagues reported that naming latencies were on average slower in the former condition, the semantically close condition, than in the latter, the semantically far condition. In sum, our result would extend Vigliocco and colleagues' results to a within-category experimental design. The authors did not report analyses by the factor cycle (nor by lag). In order to provide a direct comparison with the study of Vigliocco et al. (2002), we conducted paired t-tests on naming latencies between the close and far conditions of Experiment 3 collapsing across cycle. The analysis showed slower naming latencies in the close condition than in the far condition [$t1(23) = 4.24, p < 0.01$; $t2(53) = 5.24, p < 0.01$], converging with Vigliocco and colleagues' findings but using a within-category distance design.

Vitkovitch and Humphreys, 1991; Navarrete et al., 2010). However, it is still unclear whether such a model can account for the semantic facilitation effect as a consequence of having named a semantically related picture in the immediately preceding trial. One possibility is that it takes "some time" for the weakening mechanism to operate. That is, while the first mechanism of strengthening of the semantic-to-lexical connections over the target picture (yielding repetition priming) is immediate, the inhibitory mechanism of weakening the connections between semantic and lexical related words would appear later. In other words, just as for lexical selection by competition, some type of *post hoc* mechanism would have to be envisioned that would effectively prevent incremental learning from occurring on a trial-to-trial basis. If such a mechanism were postulated, then because the incremental learning account is articulated within a theory of lexical selection in which the most highly activated word is selected regardless of the activation levels of non-target words, then the trial-to-trial facilitation would be explained. But then we are left, again, with the question of: Why? – Why does incremental learning not occur on a trial-by-trial basis? This could be an avenue to pursue but at least in the present context, such an account would be no less *post hoc* than the claim that "lexical selection requires potentiation through repetition." Both accounts amount to either redescriptions of the data or *post hoc* and unsupported assertions that deal simply with this particular fact.

Another route to resolving these issues may be suggested by the data reported in **Figure 5**. Perhaps the cyclic naming effect has to be understood as a modulatory effect of repetition priming. What appears to be a "semantic interference effect" is really reduced repetition priming in the more related condition compared to the less related condition. For the first presentation of items there is, if anything, semantic facilitation – both comparing related to unrelated, and within-category close to within-category far. That effect (by hypothesis) is what tells us about lexical selection, as the RT effects for the first presentation have been "contaminated" by neither repetition nor by long-lasting interference effects. How does one get the interference effect from repetition priming? From the initial baseline that is established by the first presentation of the items, all response times speed up with repetition; the amount that latencies speed up with repetition, however, will be affected (by hypothesis) by the strength of the connection that obtains from semantics to their corresponding lexical items. For pictures that appear in the context of semantically more similar items, their semantic-to-lexical connections will be relatively weaker than pictures that appear in the context of semantically more distant items. This is because the degree to which incremental learning weakens non-target semantic-to-lexical representations would be affected by the semantic similarity between the target and non-target words. This type of an explanation would be able to explain the full pattern of findings, including: (1) the observation of initial semantic facilitation (lexical selection is not by competition), (2) the emergence of semantic interference with repetition ("semantic interference" is really less repetition priming), (3) the observation that by the last cycle within-category, semantically close is slower than within-category semantically far (incremental learning is modulated by semantic distance), and (4) having unrelated filler items abolishes semantic facilitation effects in the first cycle but does not affect the emergence of the "semantic interference effect" (the interference effect is caused by differential weakening of semantic-to-lexical connections, and hence is a long-lag effect, while semantic facilitation is a trial-to-trial priming effect).

This type of an explanation is able to explain findings that are otherwise difficult to explain in terms of lexical competition. For instance, the hypothesis of lexical competition would have difficulty explaining semantic interference effects that can be induced merely by a change of instruction – instructions cannot change or modulate structural properties of the system, and lexical competition is, by hypothesis, a structural property of the system. However, instructions can affect the way in which different concepts are conceived as being related to one another, and so could reasonably affect the dynamics of which connections are incrementally weakened. In line with this theoretical approach, Abdel Rahman and Melinger (2011) recently reported that semantic interference in the cyclic naming task can be modulated by providing information to the participants about the relation of the items of the blocks. Besides the homogenous and heterogeneous conditions, Abdel Rahman and Melinger included a condition in which the items come from different semantic categories but belong to a common event or theme. For instance, "coffee," "knife," "stool," "bucket," and "creek" are not categorically related items but they can be potentially integrated into the common theme of a "fishing trip." Two versions of the cyclic naming task were constructed. In one version participants performed the standard task and, while the semantic blocking effect emerged, there was no difference between the thematic and the heterogeneous conditions. Interestingly, in a second version, participants were provided with a title before initiating each block that served to provide a theme for grouping or relating the items within the blocks. For instance, the title "foods" for the homogenous block containing the food items, the title "fishing trip" for the thematic block containing the items ("coffee," "knife," "stool," "bucket," and "creek"*)* and the title "clothing" for a heterogeneous block without clothes items. Under these conditions, a thematic interference effect was reported: naming latencies in the thematic blocks were slower than in the heterogeneous blocks. According to Abdel Rahman and Melinger, "*these findings suggest that semantic activation spread during speech planning can be modulated and flexibly adapted as a given context or situation dictates*" (page 157; for similar arguments, see Experiment 3 in Wheeldon and Monsell, 1994). Our suggestion here is that the incremental learning model developed by Oppenheim et al. (2010) presents a more "flexible" characterization of the dynamics of lexical access during speech production than the hypothesis of lexical selection by competition.

A second finding relevant to the dynamics of flexibility in speech production that can be explained straightforwardly by the incremental learning model was reported by Belke et al. (2005b). In Experiment 3 of that study, Belke and colleagues studied whether there is transfer of the semantic blocking effect, as observed for repeated items, to new (previously unnamed) items. In that experiment, there could be either eight or four cycles of a given item within a block. Specifically, for half of the blocks the same items were presented across the eight cycles, while for the other half of the blocks cycles 5–8 contained different semantic-category exemplars as the items presented in cycles 1–4. Belke and colleagues found that the magnitude of the semantic effect was not modulated according to whether or not the items were old or new in the last four cycles of the block. These findings suggest that the semantic effect generalizes to new, previously unnamed items. This generalization to new, unnamed items was interpreted by Belke and colleagues as evidence that it takes some time to potentiate the lexical competition mechanism in a cyclic paradigm. The finding is also consistent with the view, however, that the semantic effect reflects weakening of semantic-to-lexical connections for non-target items.

In summary, then, our findings indicate that models of the cyclic naming effect based on lexical competition have difficulty explaining the presence of semantic facilitation in the first cycle when there are no interleaved unrelated filler items. We have argued that the incremental learning model of Oppenheim et al. (2010) is able to accommodate the "emergence" of the semantic interference effect in later cycles, if it is assumed that the interference reflects reduced repetition priming. This hypothesis can be directly evaluated with future work. The cyclic naming paradigm is fertile ground for studying the dynamics of word retrieval in speech production and will undoubtedly continue to serve as an important venue for adjudicating among different models of lexical selection.

REFERENCES

Abdel Rahman, R., and Melinger, A. (2007). When bees hamper the production of honey: lexical interference from associates in speech production. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 604–614.

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

Alario, F.-X., and Moscoso del Prado Martín, F. (2010). On the origin of the cumulative semantic inhibition effect. *Mem. Cognit.* 38, 57–66.

Bajo, M. T. (1988). Semantic facilitation with pictures and words. *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 579–589.

Belke, E., Brysbaert, M., Meyer, A. S., and Ghyselinck, M. (2005a). Age of acquisition effects in picture naming: evidence for a lexical-semantic competition hypothesis. *Cognition* 96, B45–B54.

Belke, E., Meyer, A. S., and Damian, M. F. (2005b). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *Q. J. Exp. Psychol.* 58A, 667–692.

Belongie, S., Malik, J., and Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 509–522.

Biggs, T., and Marmurek, H. (1990). Picture and word naming: is facilitation due to processing overlap? *Am. J. Psychol.* 103, 81–100.

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

Bloem, I., van den Boogaard, S., and La Heij, W. (2004). Semantic facilitation and semantic interference in language production: further evidence for the conceptual selection model of lexical access. *J. Mem. Lang.* 51, 307–323.

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

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

Caramazza, A., and Costa, A. (2001). Set size and repetition in the picture-word interference paradigm: implications for models of naming. *Cognition* 80, 291–298.

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

Costa, A., Mahon, B., Savova, V., and Caramazza, A. (2003). Level of categorization effect: a novel effect in the picture-word interference paradigm. *Lang. Cogn. Processes* 18, 205–233.

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., and Als, L. C. (2005). Long-lasting semantic context effects in the spoken production of object names. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 1372–1384.

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.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* 3, 283–321.

Finkbeiner, M., and Caramazza, A. (2006). Now you see it, now you don't: on turning semantic interference into facilitation in a Stroop-like task. *Cortex* 42, 790–796.

Finkbeiner, M., Gollan, T., and Caramazza, A. (2006). Bilingual lexical access: what's the (hard) problem? *Biling. (Camb. Engl.)* 9, 153–166.

Flores d'Arcais, G. B., and Schreuder, R. (1987). Semantic activation during object naming. *Psychol. Res.* 49, 153–159.

Forster, K. I., and Forster, J. C. (2003). DMDX: a windows display program with millisecond accuracy. *Behav. Res. Methods Instrum. Comput.* 35, 116–124.

Hartsuiker, R. J., Pickering, M. J., and de Jong, N. H. (2005). Semantic and phonological context effects in speech error repair. *J. Exp. Psychol. Learn. Mem. Cogn.* 5, 921–932.

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

Humphreys, G. W., Riddoch, M. J., and Quinlan, P. T. (1988). Cascade processes in picture identification. *Cogn. Neuropsychol.* 5, 67–103.

Huttenlocher, J., and Kubicek, L. F. (1983). The source of relatedness effects on naming latency. *J. Exp. Psychol. Learn. Mem. Cogn.* 9, 486–496.

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

Kuipers, J. R., La Heij, W., and Costa, A. (2006). A further look at semantic context effects in language production: the role of response congruency. *Lang. Cogn. Process.* 21, 892–919.

La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Mem. Cognit.* 16, 400–410.

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

Lotto, L., Job, R., and Rumiati, R. (1999). Visual effect in picture and word categorization. *Mem. Cognit.* 27, 674–684.

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

Lupker, S. J. (1988). Picture naming: an investigation of the nature of categorical priming. *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 444–455.

Mahon, B. Z., Milleville, S. C., Negri, G. A., Rumiati, R. I., Caramazza, A., and Martin, A. (2007). Action-related properties shape object representations in the ventral stream. *Neuron* 55, 507–520.

McRae, K., and Boisvert, S. (1998). Automatic semantic similarity priming. *J. Exp. Psychol. Learn. Mem. Cogn.* 24, 558–572.

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

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

Oppenheim, G. M., Dell, G. S., and Schwartz, M. F. (2010). The dark side of incremental learning: a model of cumulative semantic interference during lexical access in speech production. *Cognition* 114, 227–252.

Rapp, B., and Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychol. Rev.* 107, 460–499.

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

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

Sperber, R. D., McCauley, C., Ragain, R. D., and Weil, C. M. (1979). Semantic priming effects on picture and word processing. *Mem. Cognit.* 7, 339–345.

Vigliocco, G., Vinson, D. P., Damian, M. F., and Levelt, W. J. M. (2002). Semantic distance effects on object and action naming. *Cognition* 85, B61–B69.

Vitkovitch, M., Cooper-Pye, E., and Leadbetter, A. (2006). Semantic priming over unrelated trials: evidence for different effects in word and picture naming. *Mem. Cognit.* 34, 715–725.

Vitkovitch, M., and Humphreys, G. W. (1991). Perseverant naming errors in speeded picture naming: its in the links. *J. Exp. Psychol. Learn. Mem. Cogn.* 17, 664–680.

Wheeldon, L. R., and Monsell, S. (1994). Inhibition of spoken word production by priming a semantic competitor. *J. Mem. Lang.* 33, 332–356.

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

Received: 07 August 2011; accepted: 02 February 2012; published online: 21 February 2012.

Citation: Navarrete E, Del Prato P and Mahon BZ (2012) Factors determining semantic facilitation and interference in the cyclic naming paradigm. Front. Psychology ***3****:38. doi: 10.3389/fpsyg.2012.00038*

This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Psychology.

APPENDICES

APPENDIX A

Materials used in Experiment 1 organized by Homogenous and Heterogeneous blocks. Italian translations are provided after the English name.

Homogenous blocks

Donkey (asino), horse (cavallo), pig (maiale), pecora (sheep)
Submarine (sottomarino), airplane (aereo), bicycle (bicicletta), train (treno)
Bed (letto), hammock (amaca), bench (panchina), chair (sedia)
Cabinet (armadio), drawer (cassetiera), sofa (divano), table (tavolo)
Coconut (cocco), banana (banana), pear (pera), grape (uva)
Helicopter (elicottero), motorcycle (motocicletta), tractor (trattore), ship (nave)
Bottle (bottiglia), cup (tazza), chalice (calice), flask (fiasco)
Shirt (camicia), skirt (gonna), sweater (maglione), dress (vestito)
Car (macchina), truck (camion), van (furgone), bus (pullman)
Hat (cappello), glove (guanto), scarf (sciarpa), belt (cintura)
Spider (ragno), goat (capra), owl (gufo), penguin (pinguino)
Drum (tamburo), guitar (chitarra), piano (pianoforte), saxophone (sassofono)
Swan (cigno), cat (gatto), elephant (elefante), lizard (lucertola)
Alligator (coccodrillo), turtle (tartaruga), frog (rana), snake (serpente)
Scoop (paletta), fork (forchetta), spoon (cucchiaio), ladle (mestolo)
Record player (giradischi), radio (radio), stereo (stereo), television (televisione)

Heterogeneous blocks

Goat (capra), car (macchina), drum (tamburo), dress (vestito)
Cabinet (armadio), chalice (calice), elephant (elefante), train (treno)
Banana (banana), helicopter (elicottero), frog (rana), stereo (stereo)
Bicycle (bicicletta), horse (cavallo), glove (guanto), ladle (mestolo)
Fork (forchetta), cat (gatto), bus (pullman), scarf (sciarpa)
Coconut (cocco), lizard (lucertola), bench (panchina), submarine (sottomarino)
Ship (nave), guitar (chitarra), sofa (divano), pig (maiale)
Alligator (coccodrillo), skirt (gonna), motorcycle (motocicletta), pear (pera),
Drawer (cassetiera), belt (cintura), penguin (pinguino), television (televisione)
Bottle (bottiglia), shirt (camicia), turtle (tartaruga), piano (pianoforte)
Truck (camion), flask (fiasco), spider (ragno), table (tavolo)
Spoon (cucchiaio), saxophone (sassofono), pecora (sheep), grape (uva)
Airplane (aereo), hat (cappello), record player (giradischi), snake (serpente)
Owl (gufo), bed (letto), sweater (maglione), cup (tazza),
Donkey (asino), van (furgone), radio (radio), chair (sedia)
Hammock (amaca), swan (cigno), scoop (paletta), tractor (trattore),

APPENDIX B

Materials used in Experiments 2 and 3 organized by Semantic Context (within-category close, within-category far, and unrelated). Within-category close blocks were determined by each row (e.g. dog, wolf, fox). Within-category far related blocks were designed using a "column" restricted to each set (e.g. dog, mouse, lizard). Unrelated blocks contained the homologous object in each set restricted by group (e.g. dog, fork, orange). The shape pictures used in Experiment 3 (unrelated fillers) were circle, triangle, diamond.

Group 1

Animals:	dog, wolf, fox
	mouse, squirrel, chipmunk
	lizard, frog, snake
Tools:	fork, spoon, ladle
	plate, cup, bowl
	knife, razor, scissors
Fruits/vegetables:	orange, lime, lemon
	broccoli, cauliflower, lettuce
	cucumber, asparagus, celery

Group 2

Animals:	horse, donkey, zebra
	ant, mosquito, bee
	eagle, parrot, owl
Tools:	wrench, screwdriver, pliers
	shovel, rake, hoe
	bolt, nail, screw
Vehicles:	truck, van, bus
	bicycle, scooter, motorcycle
	wagon, carriage, sleigh

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 27 May 2011
doi: 10.3389/fpsyg.2011.00114

Event-related brain potential investigation of preparation for speech production in late bilinguals

Yan Jing Wu*[1] *and Guillaume Thierry*[1,2]

[1] *School of Psychology, Bangor University, Bangor, UK*
[2] *Economic and Social Research Council Centre for Research on Bilingualism in Theory and Practice, Bangor University, Bangor, UK*

Edited by:
Kristof Strijkers, University of Barcelona, Spain

Reviewed by:
Cristina Baus, Universitat Pompeu Fabra, Spain
Daan Hermans, Royal Dutch Kentalis, Netherlands

****Correspondence:***
Guillaume Thierry, School of Psychology, Adeilad Brigantia, Bangor University, Bangor, Gwynedd LL57 2AS, UK.
e-mail: g.thierry@bangor.ac.uk

It has been debated how bilinguals select the intended language and prevent interference from the unintended language when speaking. Here, we studied the nature of the mental representations accessed by late fluent bilinguals during a rhyming judgment task relying on covert speech production. We recorded event-related brain potentials in Chinese–English bilinguals and monolingual speakers of English while they indicated whether the names of pictures presented on a screen rhymed. Whether bilingual participants focussed on rhyming selectively in English or Chinese, we found a significant priming effect of language-specific sound repetition. Surprisingly, however, sound repetitions in Chinese elicited significant priming effects even when the rhyming task was performed in English. This cross-language priming effect was delayed by ~200 ms as compared to the within-language effect and was asymmetric, since there was no priming effect of sound repetitions in English when participants were asked to make rhyming judgments in Chinese. These results demonstrate that second language production hinders, but does not seal off, activation of the first language, whereas native language production appears immune to competition from the second language.

Keywords: ERP, bilingualism, language production, cognitive control, inhibition

INTRODUCTION

Speaking is the process of transforming thoughts into speech. Current psycholinguistic models posit that speech production involves multiple stages of information processing including, conceptualization, retrieval of lexical representations, and the motor preparation for articulation (Dell, 1986; Levelt, 1989, 1999; Caramazza, 1997). For bilingual individuals, an additional task is to select words from the appropriate (i.e., intended) language to speak. Experimental psychology and electrophysiological research have made the case that when they speak in one language, information in the other language is also being activated. Therefore, how bilinguals manage to select words in the intended language and prevent interference from the unintended language has been a central question for research on bilingual language production. In the current study, we first review previous attempts to reveal the nature of the cognitive and brain mechanisms that control bilingual language production, with a focus on methodology. We argue that the tasks and measurements used in these studies do not allow teasing apart the interplay of the first and second language; instead, they often mix the effects derived from activations of both languages during speech preparation. As a result, the function of the control mechanism has not been fully specified. We then introduce an alternative paradigm, which provides insights into first and second language activation, respectively, when bilinguals make covert spoken word production.

When speaking in their relatively weaker language (i.e., the second language), bilingual speakers have been shown to make speech errors that are characteristic of their native language (Poulisse and Bongaerts, 1994; Poulisse, 1997, 1999). Although it is difficult to determine the exact source of this cross-language interference, speech errors in bilinguals form the preliminary evidence for access to the unintended language. Further evidence derives from experimental studies that make use of interlingual materials such as cognates (words that share semantic and lexical form across languages; e.g., "café," which the same word in English and French. For example, it has been repeatedly shown that bilingual speakers name pictures faster when their names are cognates as compared to non-cognates. Given that monolinguals do not distinguish cognates from words that only exist in one language, the effect found in bilinguals must relate to representations in both the intended and the unintended languages. It has been proposed that the cognate facilitation effect originates in the additional source of activation afforded by existing representations in both the languages as compared to language-specific words (Costa et al., 2000; Kroll et al., 2000; Christoffels et al., 2006; Hoshino and Kroll, 2008, but see Sanchez-Casas and Garcia-Albea, 2005, for an alternative explanation). The cognate effect has been replicated by studies using event-related potentials (ERPs). ERPs are average recordings of brain potentials associated with mental operations. Christoffels et al. (2007) found an enhanced negativity between 300 and 400 ms when bilingual participants named pictures whose names are cognates. These ERP modulations were correlated with reduced reaction times found in both a blocked and a mixed language experiment.

In addition to the cognate effect, previous studies have shown that picture naming latency is significantly reduced when a picture is followed by or presented together with a distractor word that is the translation of the picture's name in the unintended language (i.e., the picture–word interference paradigm; Costa et al., 1999). Other studies have shown significant increase in picture naming latency when the distractor word in the non-target language is phonologically related to the picture name in the target language (the so-called phonotranslation effect; Hermans et al., 1998; Costa et al., 2003; Hermans, 2004). In both cases, the significant effect of non-target language distractors on naming latencies, whether facilitatory or interfering, is a sign of language co-activation in the course of production. This pattern of results in bilinguals is comparable with the performance of monolinguals when the picture is named in the same language as the distractor word (the so-called phonological interference effect; Lupker, 1979; La Heij et al., 1985), suggesting that planning for speech activates both the intended and the unintended language in bilinguals. Furthermore, an ERP study has also found that translation distractors presented in the non-target language reduce the ERP amplitude as compared to the control condition, when Chinese–English bilinguals are engaged in a covert picture naming task (Guo and Peng, 2006). This cross-language identity effect, which is dependent on relative proficiency in the two languages, suggests that parallel activation of the first and second language during speaking extends beyond the level of lexical representations since there is no overlap between translation equivalents in Chinese and English.

However, cognate and picture–word interference studies do not specify the level of representations that is activated in the unintended language and the time course of its activation in relation to accessing the target language. Since cognates share lexical-semantic as well as phonological features across languages, it is difficult to pinpoint the representation level at which the cognate facilitation effect finds its source (Costa et al., 2005). For example, Strijkers et al. (2010) demonstrated, in an early temporal window (e.g., 180–200), effects of lexical frequency and cognate status on ERPs collected during bilingual speech production. However, the lexical origin of the cognate effect cannot exclude the possibility of phonological priming resulting from additional source of activations due to shared phonological representations. Another issue is that cognates, like other stimuli present in some form in two different languages, are likely to activate representations and processes in the two languages (Hermans et al., 2010; Wu and Thierry, 2010). While for some bilingual individuals (e.g., Spanish–Catalan bilinguals) dual-language is the natural speech context, bilinguals with other language pairs are exposed to a different language context (e.g., Chinese–English bilinguals). Therefore, studies involving cognates have limitations regarding result generalization. Furthermore, cognate effects, as assessed by behavioral performance or with ERPs, cannot tease apart the processes at work when bilingual produce words in each of their languages considered separately. This problem is even more salient when pictures are mixed with distractor words in the unintended language whether they are presented in the visual or the auditory modality. On the basis of effects observed in the picture–word interference paradigm it is only possible to infer that the unintended language is accessed when bilinguals prepare for speech. It is difficult, however, to characterize the independent contribution of language-specific representations or to distinguish semantic and lexical interference from cognate effects. Moreover, in a typical picture–word interference paradigm, the distractor word itself initiates a bottom-up word recognition process that intrudes into speech planning. Consequently, evidence derived from picture–word interference must be regarded as complicated by the interaction between the processing of the word and that of the picture rather than "pure" word production (see alternative evidence for cross-language phonological activations using simple picture naming in Colomé, 2001).

To characterize the nature of the representations from the two languages accessed during speech production in bilingual individuals, the present study manipulated phonological priming in the first and second languages independently. In experiment 1, Chinese–English proficient bilinguals were engaged in a rhyming judgment task in which they had to decide whether the English name of the target picture rhymed with that of a picture prime. Pairs of pictures from four conditions were presented randomly: semantically related, semantically unrelated but rhyming in English, and semantically unrelated but rhyming in Chinese, and semantically unrelated but rhyming in neither English or Chinese (**Figure 1**). We avoided artificial effects by facial movements on ERPs during overt speech, and also enabled measurement of activity in a late time window (i.e., 600 ms+) by engaging participant in a task only requiring button presses. However, this relied on the assumption that rhyming judgment required phonological access to the name of the picture.

Previous behavioral studies are limited to overall effects on reactions times which tell us nothing of the phases of processing preceding the observed response (e.g., voice reaction time). The present study used ERPs to investigate language co-activation during production to provide insights into the time course of priming effect from stimulus presentation to response. Whilst the ERPs elicited by semantically related pictures were expected to reveal the time course of access to meaning, ERPs elicited by target pictures that rhymed in English or in Chinese with the name of the picture prime provided insight into the activation of intended versus unintended phonological representations. Indeed, if naming in English involves phonological access to Chinese (i.e., the unintended language), this process can be characterized by comparison with phonological activation of English (i.e., the intended language) and with semantic priming, since these three processes were tested independently within three different experimental conditions. This paradigm avoids the explicit dual-language context caused by the presentation of distractor words and a language-ambiguous context by the use of cognates. We also tested a group of native English speakers as control participants to (1) obtain a baseline for rhyming effects and (2) ensure that the rhyming manipulation in Chinese picture names did not induce spurious effects in English.

In experiment 2, Chinese–English bilingual participants performed the rhyming judgment task in Chinese to examine possible influences of second language activation during the production of the native language. The majority of studies in the literature have focused on the influence of the stronger language (i.e., the first or native language) on the processing of the weaker language (i.e., the second language). Therefore, it remains unclear whether a second language affects the processing of the native language during production (but see Bloem and La Heij, 2003). To investigate this, the present study tested covert word production in both languages using a fully balanced design. When bilingual participants performed the task in Chinese, priming effects triggered by English rhymes were expected to reveal the potential interference of second language information retrieval during native language

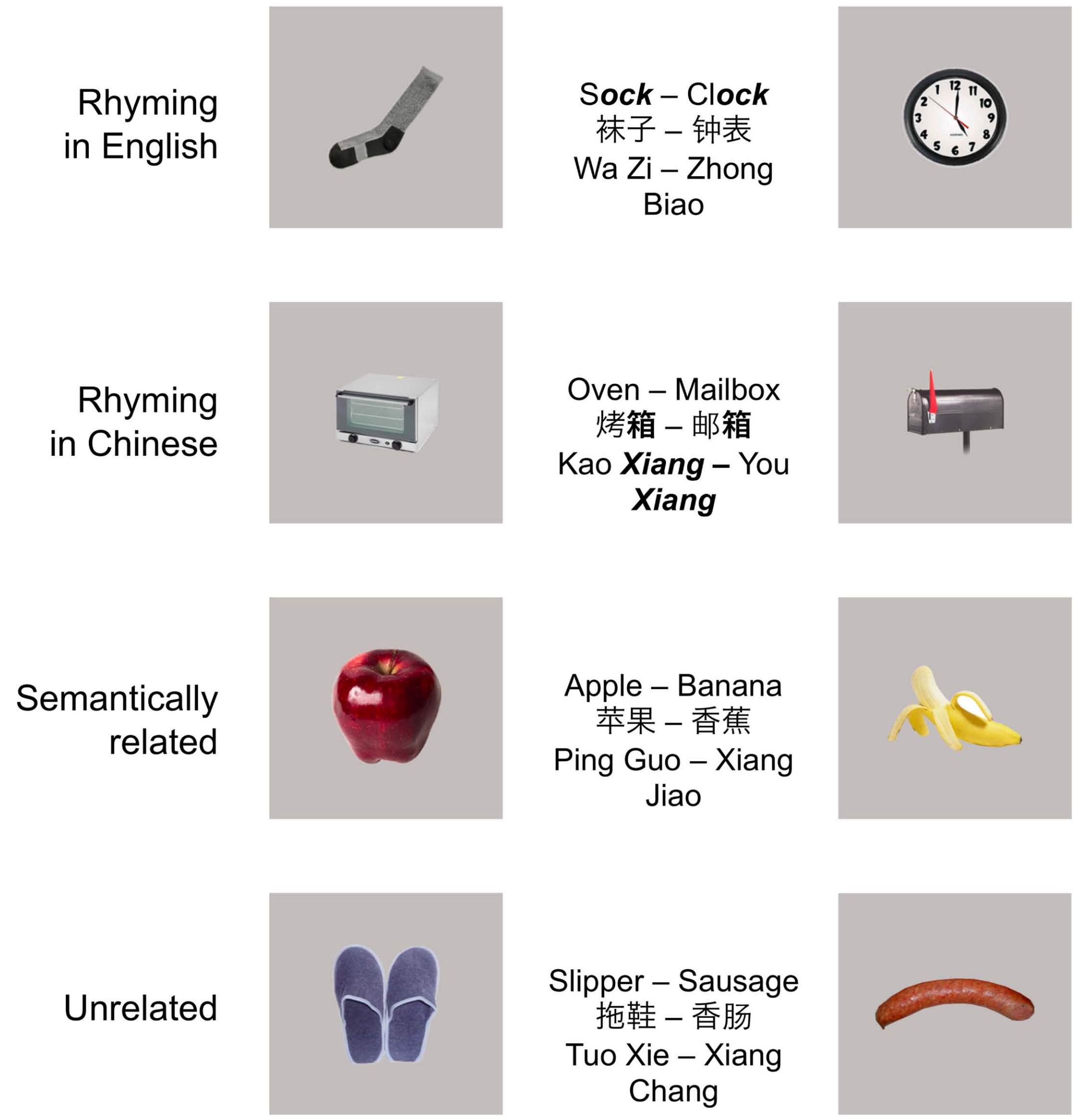


FIGURE 1 | Examples of stimuli used in the rhyming judgment tasks. Each cell contains one example of a picture pair used in the English and Chinese tasks, its English names, its simplified Chinese translations, and the corresponding Chinese Pin Yin (alphabetic transposition of the phonological form). As compared to English, Chinese characters that rhyme seldom bear overlaps in written forms (i.e., orthography). To prevent the potential confounding effects of orthographic variability on picture naming (Weekes et al., 2005; Bi et al., 2009; Zhang et al., 2009), we used Chinese words that shared a character repetition in both phonology and orthography, forming a "rhyming" condition that is comparable to the English control condition. Color pictures of objects from real life situation were used to minimize naming difficulty, since black–white line drawings can sometimes be more difficult to name. The current experiment did not include a "familiarization" procedure in which participants were trained with the desired names of stimuli in advance. Although, as is common practice in picture naming studies, such practice helps reduce error rates, ERPs are particularly sensitive to storage in episodic memory. Also, such a procedure tends to prime a specific lexical candidate for each picture and, therefore, may artificially bias language production.

production. In the same experimental session, priming effects elicited by Chinese rhymes served as a baseline for overt priming effects in the intended native language.

MATERIALS AND METHODS

PARTICIPANTS

Fifteen Chinese students studying at Bangor University who had normal or corrected-to-normal vision and self-reported normal hearing were paid to take part in the experiment. Their age ranged between 19 and 23, and they were controlled for handedness (right) and the level of education (undergraduate). The Chinese–English bilinguals started English formal instruction at the age of puberty (e.g., 12 or 13). At the time of testing, they were living and studying in the UK for an average of 18 (±3.2) months. In terms of English proficiency, all participants had a score of 6 as measured by the International English Language Testing System (IELTS), which is the entrance requirement for non-native speakers to study in most English-speaking institutions (www.ielts.org/

test_takers_information/what_is_ielts.aspx). The IELTS equally covers four fundamental language skills (i.e., reading, listening, writing, and speaking). The maximum score for IELTS is 9 and the majority of test takers obtain a score of between 4 and 7. Fifteen English monolinguals were recruited from students taking a psychology undergraduate course at Bangor University as control participants and they were paid with course credits for their participation. Every participant signed a consent form before taking part in the experiment that was approved by the ethics committee of the School of Psychology, Bangor University.

STIMULI

Two hundred pairs of pictures were equally allocated to four experimental conditions in terms of semantic relatedness (i.e., related or unrelated), and rhyming names (i.e., rhyming in English or Chinese). They were matched between conditions for lexical frequency and concreteness (Coltheart, 1981). The English names were matched for numbers of phonemes across conditions ($P > 0.1$) and the Chinese names were always two characters in length. Semantic relatedness between pictures was rated on a Likert scale from 1 (unrelated) to 5 (strongly related) by two independent groups of native Chinese and native English speakers (**Figure 1**). Differences in semantic relatedness ratings were highly significant between semantically related and unrelated pairs ($P < 0.001$ for all pairwise comparisons). Picture stimuli were matched across conditions for basic visual characteristics (e.g., size, resolution, and background). The variability in point of view, shape, and color of the objects presented was large in all the conditions to avoid a systematic bias in terms of inter-stimulus variance (Thierry et al., 2007). Particular care was taken in the choice of pictorial representations for each target word such that these were not biased toward Chinese or English cultural prototypes (see examples in **Figure 1**). No picture was repeated throughout the experiment.

PROCEDURE

All experiments took place in a sound-proof laboratory where the participant sat on a comfortable armchair 1.5 m away from a computer screen. After signing the consent form and receiving the instruction, participants viewed two blocks of stimuli presented in a pseudo-randomized order. Each trial began with a pre-stimulus interval of 200 ms. A picture was then flashed for 500 ms at fixation followed by the second picture of a pair, which stayed on the screen until a response was made, after a randomly selected inter-stimulus interval of 500, 600, or 700 ms. Participants were instructed to indicate whether the name of the second picture in each pair either rhymed in English (Exp 1) with that of the first picture or shared a phonological component (character) in Chinese (Exp 2) by pressing keys set under their left and right index fingers. Response side and the order of experiments were fully counterbalanced between participants. Naturally, English control participants who have no knowledge of Chinese were only given the English rhyming task. All participants were debriefed orally.

ERP RECORDING

Electrophysiological data were recorded in reference to Cz at a rate of 1 kHz from 64 Ag/AgCl electrodes placed according to the extended 10–20 convention. Impedances were kept <5 kΩ. Electroencephalogram activity was filtered on-line band pass between 0.1 and 200 Hz and refiltered off-line with a 25-Hz, low-pass, zero-phase shift digital filter. Eye blinks were mathematically corrected, and remaining artifacts were manually dismissed. There was a minimum of 30 valid epochs per condition in every subject. Epochs ranged from −100 to 1000 ms after the stimulus onset. Baseline correction was performed in reference to pre-stimulus activity, and individual averages were digitally re-referenced to the global average reference. ERP data were collected simultaneously to behavioral data.

ERP DATA ANALYSIS

Peak detection was carried out automatically, time-locked to the latency of the peak at the electrode of maximal amplitude on the grand-average ERP. Temporal windows for peak detection were determined based on visual inspection of variations of the Global Field Power measured across the scalp (Picton et al., 2000; Luck, 2005). Mean ERP amplitudes elicited by the target picture were subjected to a repeated measures analysis of variance (ANOVA) with rhyming (rhyming in English/rhyming in Chinese/no rhyming but semantically related/completely unrelated) and electrode (63 levels) as within-subject factors, and group as between-subject factor (native English controls/Chinese–English bilinguals) using a Greenhouse–Geisser correction where applicable. We also analyzed ERP data by means of pairwise millisecond-by-millisecond comparisons between conditions considered significant when differences were above threshold ($P < 0.05$) for >30 ms over a minimum of nine clustered electrodes (Guthrie and Buchwald, 1991).

RESULTS

In experiment 1, when native English speakers performed the rhyming task in English, repeated measures ANOVA revealed a significant main effect of condition on reaction times ($F_{3,42} = 2.91$, $P < 0.05$). *Post hoc* analysis (LSD) showed that this difference was driven by faster reaction times for target pictures names that rhymed with prime pictures names in English as compared to all other conditions (**Figure 2**, all Ps < 0.05). We also found that more errors were made for the English rhyming condition ($F_{3,42} = 8.61$, $P < 0.001$) than for the other conditions (all Ps < 0.001). In particular, no effect of rhyming in Chinese names was found on either reaction times or error rates in native English control participants (all Ps > 0.1). In the Chinese–English bilinguals, rhyming in English reduced reaction times ($F_{3,42} = 3.08$, $P < 0.001$) and increased error rates ($F_{3,42} = 4.7$, $P < 0.001$) as compared to semantically related and unrelated picture pairs, but no significant reaction time difference was found between pairs of picture names that rhymed in English and those that rhymed in Chinese ($P > 0.1$). However, picture names rhyming in Chinese (i.e., in the unintended language) also increased error rates as compared to semantically related and unrelated picture pairs ($P < 0.05$).

In experiment 2, Chinese–English bilingual participants making rhyming judgments in Chinese responded significantly faster ($F_{3,42} = 2.98$, $P < 0.05$) to picture pairs with rhyming names in Chinese and semantically related pictures as compared to picture pairs that rhymed in English and unrelated pictures (all Ps < 0.05). Bilingual participants also made more errors ($F_{3,42} = 3.42$, $P < 0.05$) in these two conditions as compared to

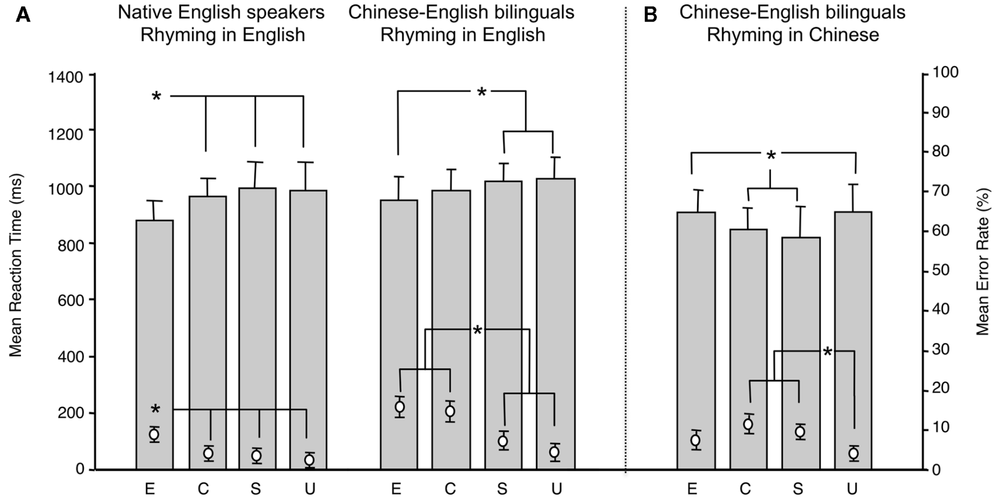


FIGURE 2 | Behavioral results of all groups in the rhyming judgment tasks. Reaction times (bars; left axis) and error rates (bullets; right axis) for the native English speakers and the Chinese–English bilingual speakers **(A)** in the English rhyming judgment task is presented on the left of the vertical line. Results of the Chinese–English bilingual speakers in the Chinese rhyming judgment task are presented to the right of the vertical line **(B)**. Conditions in which the picture pairs had names rhyming in English, rhyming in Chinese, were semantically related, or unrelated are labeled E, C, S, and U, respectively. The stars indicate significant differences ($P < 0.05$). Error bars depict SEM in all cases.

the unrelated condition (both *P*s < 0.05). No effect of rhyming in English was found either on reaction times or error rates in this experiment (all *P*s > 0.1).

The ERP data was collected simultaneously with behavioral data. In native speakers of English performing the English rhyming task, a repeated ANOVA showed a significant effect of condition on ERP mean amplitude ($F_{3,42} = 19.2$, $P < 0.001$). *Post hoc* analysis revealed that this effect was accounted for by two differences (**Figure 3**). Firstly, target pictures that rhymed with prime pictures in terms of English names elicited significantly reduced ERP amplitudes as compared to those rhyming based on Chinese names and unrelated picture names (all Ps < 0.001). A millisecond-by-millisecond comparison revealed that the significant difference in this comparison started as early as 220 ms after the presentation of the target picture. Secondly, target pictures related in meaning to the prime pictures elicited reduced ERP amplitudes as compared to unrelated pairs of pictures ($P < 0.001$), with a similar time course as the priming found for rhyming in English. There was no difference between the ERPs elicited by target pictures whose names rhymed based on Chinese picture names and the ERPs elicited by completely unrelated pictures ($P > 0.1$).

Statistical analysis of ERPs recorded in the Chinese–English bilinguals performing the English rhyming task showed a main effect of condition ($F_{3,42} = 5.52$, $P < 0.001$). Rhyming in English and semantic relatedness of the pictures both reduced the ERP main amplitude against the unrelated condition (all *P*s < 0.001). The priming effect elicited by rhymes in English was significant between 250 and 600 ms and of smaller magnitude than the same effect found in the native English participants. The priming effect of semantic relatedness started at around 250 ms and extended throughout the whole period of analysis (i.e., 1000 ms), showing a comparable time course and magnitude to that of the native English control participants.

Critically, analysis of ERP modulations elicited by pictures whose names rhymed in Chinese also revealed a significant priming effect against the unrelated condition. When compared to unrelated pictures, pictures with rhyming Chinese names reduced mean ERP amplitude from 500 to 800 ms, that is later than the English rhyming effects found in both the Chinese–English bilinguals and the native English speakers. The priming effect elicited by rhyming in Chinese names was also smaller in magnitude as compared to the effects of semantic relatedness and explicit rhyming in English.

The effect of rhyming in the unintended language was further confirmed by means of a between-subject repeated measures ANOVA comparing native English speakers and Chinese–English bilinguals. The main effect of rhyming in Chinese was not significant across groups ($P > 0.1$). However, we found a significant group-by-phonological priming interaction ($F_{1,28} = 4.74$, $P < 0.05$), such that rhyming in Chinese had no effect in the native English speakers ($P > 0.1$) but, in Chinese–English bilinguals, it significantly reduced N400 amplitude for pictures that rhymed in their Chinese names as compared to pictures that were unrelated ($P < 0.001$; **Figure 3**).

When Chinese–English bilingual participants were asked to make rhyming judgment in Chinese (i.e., Exp 2), target pictures that were either semantically related or rhymed in Chinese with the prime pictures induced a significantly smaller N400 than pictures that were unrelated to the primes (both *P*s < 0.05). Both effects became significant at around 280 ms after stimuli presentation and, together, explained the main effect of experimental conditions ($F_{3,42} = 2.77$, $P < 0.05$). Noticeably, rhyming in English yielded no significant effect on any ERP components when compared to the unrelated condition.

DISCUSSION

The purpose of the present study was to examine the mental processes underlying spoken word production in bilingual individuals. This was achieved by having participants name covertly

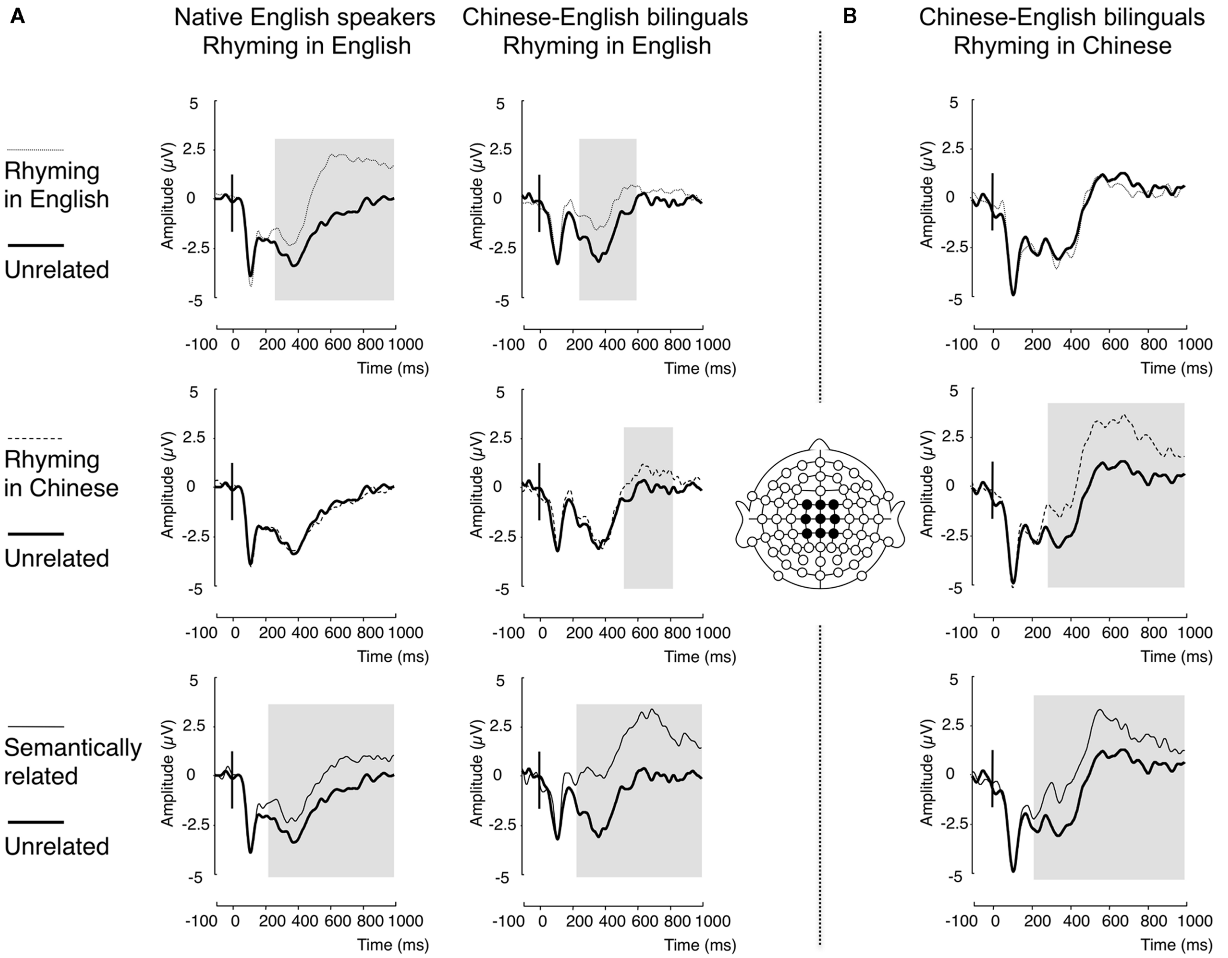


FIGURE 3 | Event-related potential results of all groups in the rhyming judgment tasks. ERP results for the native English speakers and the Chinese–English bilingual speakers in the English rhyming judgment task are presented to the left of the vertical line **(A)**. ERP results for the Chinese–English bilingual speakers in Chinese rhyming judgment task are presented to the right of the vertical line **(B)**. Waveforms depict brain potential variations from nine central electrodes (FC1, FC2, FCz, C1, C2, Cz, CP1, CP2, CPz). The schematic head shows electrode locations. The shaded areas represent significant differences between conditions (e.g., $P < 0.05$) over a minimal period of 30 ms.

pairs of pictures that rhymed in their first and second language, as well as those that were related in meaning. Given that the three experimental conditions were tested separately, the paradigm teases apart relative contributions of these factors to the process of speech preparation.

BEHAVIORAL RESULTS SUGGEST PARALLEL ACTIVATIONS OF BOTH LANGUAGES IN BILINGUALS

When making rhyming judgment on the English names, both native English speakers and Chinese–English bilinguals displayed reduced reaction times and increased error rates for target pictures that rhymed with the prime pictures in English as compared to other conditions. This behavioral pattern might be due to conflicts between the task-dependent expectations and the relatively low proportion of target picture pairs (25%) in the experiment. However, Chinese–English bilinguals also showed an increased error rate for target pictures that rhymed with the prime pictures in Chinese, an effect absent in the native English speakers. The Chinese rhyming effect found in the English experiment suggests that speech preparation in bilinguals is language non-selective. By contrast, when Chinese–English bilinguals made rhyming judgment on the basis of Chinese picture names (i.e., Exp 2), a reduced reaction time and increased error rate was observed only in the Chinese rhyming condition as compared to the unrelated condition. No sign of phonological access to English was found, as rhyming in English names did not affect either reaction time or error rate. These findings reveal an asymmetry in the cross-language interactions during bilingual word production: Speaking in the second language activates phonological representations in the first language, but not *vice versa*.

ERPS DISSOCIATE ACCESS TO THE INTENDED FROM THE UNINTENDED LANGUAGE

In the English rhyming task (i.e., Exp 1), ERP amplitude modulation was observed in the N400 range when the target picture was semantically related or rhymed with the prime pictures in English in both the English monolinguals and the Chinese–English bilinguals. However, target pictures with names that rhymed with prime picture names in Chinese, the unintended language, also modulated ERPs in the Chinese–English bilinguals, suggesting that phonological representations of the native language are accessed during the planning of speech production in the second language. English monolinguals did not show any ERP modulation for pictures with names that rhymed in Chinese, indicating that the character repetition in Chinese did not spuriously interact with other conceptual or lexical variables involved in spoken word production. Therefore, the Chinese rhyming effect observed in ERPs when bilingual participants make rhyming judgment in English can only be accounted for by spontaneous access to phonological representations in the unintended language, i.e., the same conclusion as that drawn from the behavioral findings. However, unlike mean reaction times, which are the final product of a convolution of cognitive processes, the high temporal resolution of ERPs allows the analysis of millisecond-by-millisecond unfolding of mental functions. This analysis reveals that the ERP effect elicited by rhyming in English became significant 150 ms before the effect in Chinese, despite the fact that the two effects were comparable in direction and magnitude. This suggests that phonological retrieval of the intended language begins earlier than that of the unintended language during speech production in the second language.

In the Chinese rhyming task (i.e., Exp 2), target pictures that are semantically related or rhyme via Chinese names elicited reduced ERP amplitude as compared to unrelated pictures. In this instance, since Chinese was the intended language, rhyming effects emerged as early as in the case of the English rhyming task. The time course of the explicit Chinese rhyming effect suggests that the relatively late effect of Chinese phonological repetition in the English rhyming task is not due to processing differences between the two languages (Liu and Perfetti, 2003); it indeed reflects a cognitive mechanism that dissociates phonological retrieval of the intended from that of the unintended language during spoken word production. Moreover, rhyming in English names did not have an impact on the ERPs of Chinese rhyming judgment, suggesting that, consistent with the behavioral results, spoken word production in the native language does not involve access to phonological representations of the second language.

In addition to the behavioral evidence of non-selective access in covert speech production of the second language, a critical finding of the current study is that access to the intended and unintended languages involves different time courses. This novel finding provides a basis to contrast two hypotheses regarding lexical selection mechanism in bilinguals. Previous research has established that bilinguals activate both languages, to a dynamic level of representations, while speaking in one language only (for a review see Kroll et al., 2006). One explanation as to how bilinguals prevent cross-language interference posits that an inhibitory mechanism suppresses lexical competition from the unintended language that is activated initially to allow for the selection of words from the intended language (Green, 1998). Such cognitive control mechanism would not only account for bilingual lexical selection at both the behavioral and neuroanatomical levels (Abutalebi and Green, 2007; Abutalebi et al., 2008), but would also help explain bilinguals' superior performance in a range of non-verbal tasks (Bialystok et al., 2005; Costa et al., 2008; Emmorey et al., 2008). However, a contrasting view posits that language cues or the intention to speak in one language serves to differentially activate bilinguals' two languages so that the intended language receives stronger activation than the unintended language at the conceptual level (Finkbeiner et al., 2006; see also Poulisse and Bongaerts, 1994; La Heij, 2005). According to this differential activation proposal, the observed lexical access to the unintended language reflects only a natural flow of activation, but does not functionally compete for selection with the intended language (Costa et al., 1999; Costa, 2005). While one possibility is that the two accounts represent the selection mechanisms of bilinguals at different levels of second language proficiency (Costa and Santesteban, 2004), there has been, so far, little evidence that directly supports this "selection-by-language proficiency" account.

The Chinese rhyming effect observed here in the English task suggests that the intention to speak in one language does not suffice to eliminate activation of the other language. Consistent with this view, the activation of the unintended language also influenced bilinguals' behavioral performance, which *may* involve inhibition as the underlying mechanism. Furthermore, differential activation levels of the intended and unintended languages were manifested as temporally separated ERP modulations. Overall, these results are compatible with co-activation of language representations from the two languages although there may be a temporal dynamic aspect of activation-inhibition processes that will need to be specified in the future.

Here, to avoid contaminations arising from muscle movement, spoken word production was tested via covert naming (i.e., rhyming judgment of picture names) rather than overt production (But see Costa et al., 2009; Strijkers et al., 2010, and Hoshino and Thierry, 2011). The rhyming task was chosen because previous studies in monolinguals have shown that rhyming reflects phonological analysis during spoken word preparation: it is associated with reduced negativity in the N400 range during both reading (Grossi et al., 2001) and picture naming (Barrett and Rugg, 1990a,b). Furthermore, an auditory study has shown that target words spoken in different voices than prime words elicit the same pattern of ERP variations, indicating that the rhyming is not significantly affected by physical-acoustic variables, but rather reflects a phonological matching process (Praamstra and Stegeman, 1993). However, the judgment task used here arguably involved a matching process that is not part of spoken word production in everyday life. Bilingual participants might have involuntarily named the picture in the unintended language during reanalysis of the stimuli, despite the fact that the instructions did not encourage them to access both languages. Also, the reprocessing of the picture names may have happened as part of the speech monitoring process and it could account for the delayed ERP effect in the case of Chinese rhyming. In other words, bilingual participants could have accessed Chinese picture names as they were checking for possible sources of errors and preparing for the response relative to rhyming in English, but not in the initial stage of lexical selection. In addition, the fact that, in the current study, bilingual participants were tested both in the Chinese and the English tasks may have encouraged this monitoring process. As a result, while rhyme-based priming is an index of phonological retrieval, its ERP correlates might have been influenced by task-dependent components[1]. A potential solution is to record brain potentials while participants produce speech overtly. Despite a reduced window of reliable measurement, this methodology has been successfully applied in several studies recently (Christoffels et al., 2007; Costa et al., 2009; Strijkers et al., 2010). For example, Hoshino and Thierry (2011) showed that, when Spanish–English bilinguals name pictures in English, visually presented English distractor words phonologically related to the name of the picture in Spanish (phonotranslation condition) significantly modulated ERP mean amplitude in two temporal windows (e.g., 200–260 ms and 350–400 ms). Consistent with the current study, Hoshino and Thierry's (2011) findings suggest that phonological representations of the unintended language are accessed and that they compete for selection during second language production.

Another finding of the current study worth considering is that no effect of rhyming in English was observed in the Chinese rhyming task, suggesting that lexical selection may be resolved shortly after the conceptual level and before access to phonological forms. This idea is consistent with previous studies showing language-selective access of spoken word production in the native language (Bloem and La Heij, 2003; La Heij, 2005; Ivanova and Costa, 2008; Colomé and Miozzo, 2010; but also see Rodriguez-Fornells et al., 2005). It also argues that rhyming judgment does not necessarily lead to artificial activation of both languages, which may be seen as a criticism of the English rhyming experiment. Further studies will elucidate the nature of representations and the relative timing of access in bilingual speech production, for instance by combining the design implemented here with overt speech.

ACKNOWLEDGMENTS

Yan Jing Wu and Guillaume Thierry are funded by the Economic and Social Research Council (RES-000-23-0095) and the European Research Council (ERC-StG-209704).

REFERENCES

Abutalebi, J., Annoni, J. M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., Lazeyras, F., Cappa, S. F., and Khateb, A. (2008). Language control and lexical competition in bilinguals: an event-related FMRI study. *Cereb. Cortex* 18, 1496–1505.

Abutalebi, J., and Green, D. (2007). Bilingual language production: the neurocognition of language representation and control. *J. Neurolinguistics* 20, 242–275.

Barrett, S. E., and Rugg, M. D. (1990a). Event-related potentials and the phonological matching of picture names. *Brain Lang.* 38, 424–437.

Barrett, S. E., and Rugg, M. D. (1990b). Event-related potentials and the semantic matching of pictures. *Brain Cogn.* 14, 201–212.

Bi, Y., Xu, Y., and Caramazza, A. (2009). Orthographic and phonological effects in the picture-word interference paradigm: evidence from a logographic language. *Appl. Linguist.* 30, 637–658.

Bialystok, E., Craik, F. I., Grady, C., Chau, W., Ishii, R., Gunji, A., and Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: evidence from MEG. *Neuroimage* 24, 40–49.

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

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

Christoffels, I. K., De Groot, A. M. B., and Kroll, J. F. (2006). Memory and language skills in simultaneous interpreting: the role of expertise and language proficiency. *J. Mem. Lang.* 54, 324–345.

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

Colomé, A. (2001). Lexical activation in bilinguals' speech production: language-specific or language-independent? *J. Mem. Lang.* 45, 721–736.

Colomé, A., and Miozzo, M. (2010). Which words are activated during bilingual word production? *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 96–109.

Coltheart, M. (1981). The MRC psycholinguistic database. *Q. J. Exp. Psychol.* 33, 497–505.

Costa, A. (2005). "Lexical access in bilingual production," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. De Groot (New York: Oxford University Press), 308–325.

Costa, A., Caramazza, A., and Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *J. Exp. Psychol. Learn Mem. Cogn.* 26, 1283–1296.

Costa, A., Colomé, A., Gomez, O., and Sebastian-Galles, N. (2003). Another look at cross-language competition in bilingual speech production: lexical and phonological factors. *Biling. (Camb. Engl.)* 6, 167–179.

Costa, A., Hernandez, M., and Sebastian-Galles, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86.

Costa, A., Miozzo, M., and Caramazza, A. (1999). Lexical selection in bilinguals: do words in bilingual's two lexicons compete for selection? *J. Mem. Lang.* 41, 365–397.

Costa, A., and Santesteban, M. (2004). Bilingual word perception and production: two sides of the same coin? *Trends Cogn. Sci.* 8, 253; author reply 254.

Costa, A., Santesteban, M., and Cano, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain Lang.* 94, 94–103.

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.

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

Emmorey, K., Luk, G., Pyers, J. E., and Bialystok, E. (2008). The source of enhanced cognitive control in bilinguals: evidence from bimodal bilinguals. *Psychol. Sci.* 19, 1201–1206.

Finkbeiner, M., Gollan, T., and Caramazza, A. (2006). Bilingual lexical access: what is the (hard) problem? *Biling. (Camb. Engl.)* 9, 153–166.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. (Camb. Engl.)* 1, 67–81.

Grossi, G., Coch, D., Coffey-Corina, S., Holcomb, P. J., and Neville, H. J. (2001). Phonological processing in visual rhyming: a developmental ERP study. *J. Cogn. Neurosci.* 13, 610–625.

Guo, T., and Peng, D. (2006). Event-related potential evidence for parallel activation of two languages in bilingual speech production. *Neuroreport* 17, 1757–1760.

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

Hanulovà, J., Davidson, D. J., and Indefrey, P. (2010). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Lang. Cogn. Proc.* doi: 10.1080/01690965.2010.509946. [Epub ahead of print].

Hermans, D. (2004). Between-language identity effects in picture-word interference tasks: a challenge for language-nonspecific or language-specific lexical access? *Int. J. Biling.* 8, 115–125.

Hermans, D., Bongaerts, T., de Bot, K., and Schreuder, R. (1998). Producing words in a foreign language: can speakers prevent interference from their first language? *Biling. (Camb. Engl.)* 1, 213–229.

Hermans, D., Ormel, E., Besselaar, R. V., and van Hell, J. (2010). Lexical activation in bilinguals' speech production is dynamic: how language ambiguous words can affect crosslanguage activation. *Lang. Cogn. Proc.* doi: 10.1080/01690965.2010.530411. [Epub ahead of print].

Hoshino, N., and Kroll, J. F. (2008). Cognate effects in picture naming: does cross-language activation survive a change of script? *Cognition* 106, 501–511.

Hoshino, N., and Thierry, G. (2011). Language selection in bilingual word production: electrophysiological evidence for cross-language competition. *Brain Res.* 1371, 100–109.

Ivanova, I., and Costa, A. (2008). Does bilingualism hamper lexical access

---

[1] See a similar discussion regarding the N2 effect as the index of phonological or phonetic segment access in bilingual speakers engaged in the dual-choice go/nogo paradigm (Hanulovà et al., 2010).

in highly-proficient bilinguals? *Acta Psychol.* 127, 277–288.

Kroll, J. F., Bobb, S. C., and Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: arguments against a fixed locus of language selection in bilingual speech. *Biling. (Camb. Engl.)* 9, 119–135.

Kroll, J. F., Dijkstra, T., Janssen, N., and Schriefers, H. (2000). Selecting the language in which to speak: experiments on lexical access in bilingual production. *Paper Presented at the 41st Annual Meeting of the Psychonomic Society*, New Orleans, LA.

La Heij, W. (2005). "Selection processes in monolingual and bilingual lexical access," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. B. De Groot (New York: Oxford University Press), 189–307.

La Heij, W., van der heijden, A. H. C., and Schreuder, R. (1985). Semantic priming and Stroop-like interference in word-naming tasks. *J. Exp. Psychol. Hum. Percept. Perform.* 11, 62–80.

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

Levelt, W. J. (1999). Models of word production. *Trends Cogn. Sci.* 3, 223–232.

Liu, Y., and Perfetti, C. A. (2003). The time course of brain activity in reading English and Chinese: an ERP study of Chinese bilinguals. *Hum. Brain Mapp.* 18, 167–175.

Luck, S. (2005). *An Introduction to Event-Related Potentials Technique*. Cambridge, MA: The MIT Press.

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

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R. Jr., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., and Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 37, 127–152.

Poulisse, N. (1997). "Language production in bilinguals," in *Tutorials in Bilingualism: Psycholinguistic Perspectives*, eds A. M. B. De Groot and J. F. Kroll (Mahwah, NJ: Lawrence Erlbaum Associates), 201–224.

Poulisse, N. (1999). *Slips of the Tongue: Speech Errors in First and Second Language Production*. Amsterdam/Philadelphia: John Benjamins.

Poulisse, N., and Bongaerts, T. (1994). First language use in second language production. *Appl. Linguist.* 15, 36–57.

Praamstra, P., and Stegeman, D. F. (1993). Phonological effects on the auditory N400 event-related brain potential. *Brain Res. Cogn. Brain Res.* 1, 73–86.

Rodriguez-Fornells, A., van der Lugt, A., Rotte, M., Britti, B., Heinze, H. J., and Munte, T. F. (2005). Second language interferes with word production in fluent bilinguals: brain potential and functional imaging evidence. *J. Cogn. Neurosci.* 17, 422–433.

Sanchez-Casas, R. M., and Garcia-Albea, J. E. (2005). "The representation of cognate and noncognate words in bilingual memory," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. De Groot (New York: Oxford University Press), 226–250.

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.

Thierry, G., Martin, C. D., Downing, P., and Pegna, A. J. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nat. Neurosci.* 10, 505–511.

Weekes, B., de Zubicaray, G., McMahon, K., Eastburn, M., Bryant, M., and Wang, D. (2005). Orthographic effects on picture naming in Chinese: a 4T erfMRI study. *Brain Lang.* 95, 14–15.

Wu, Y., and Thierry, G. (2010). Investigating bilingual processing: the neglected role of language processing contexts. *Front. Psychol.* 1:178. doi: 10.3389/fpsyg.2010.00178

Zhang, Q., Chen, H. C., Weekes, B. S., and Yang, Y. (2009). Independent effects of orthographic and phonological facilitation on spoken word production in Mandarin. *Lang. Speech*, 52(Pt 1), 113–126.

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

Received: 25 February 2011; paper pending published: 11 March 2011; accepted: 15 May 2011; published online: 27 May 2011.
*Citation: Wu YJ and Thierry G (2011) Event-related brain potential investigation of preparation for speech production in late bilinguals. Front. Psychology **2**:114. doi: 10.3389/fpsyg.2011.00114*
This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 09 December 2011
doi: 10.3389/fpsyg.2011.00371

Early goal-directed top-down influences in the production of speech

***Kristof Strijkers*[1] *, *Yen Na Yum*[2], *Jonathan Grainger*[3] and *Phillip J. Holcomb*[2]**

[1] *Department of Psicologia Bàsica, Universitat de Barcelona, Barcelona, Spain*
[2] *Department of Psychology, Tufts University, Medford, MA, USA*
[3] *CNRS, Aix-Marseille University, Marseille, France*

Edited by:
Albert Costa, University Pompeu Fabra, Spain

Reviewed by:
Niels O. Schiller, University of Leiden, Netherlands
Bradford Mahon, University of Rochester, USA

****Correspondence:***
Kristof Strijkers, Department of Psicologia Bàsica, Universitat de Barcelona, Vall D'Hebron 171, 08035 Barcelona, Spain.
e-mail: kristof.strijkers@gmail.com

It was recently reported that the conscious intention to produce speech affects the speed with which lexical information is retrieved upon presentation of an object (Strijkers et al., 2011). The goal of the present study was to elaborate further on the role of these top-down influences in the course of planning speech behavior. In an event-related potentials (ERP) experiment, participants were required to overtly name pictures and words in one block of trials, while categorizing the same stimuli in another block of trials. The ERPs elicited by the naming task started to diverge very early on (~170 ms) from those elicited by the semantic categorization task. Interestingly, these early ERP differences related to task intentionality were identical for pictures and words. From these results we conclude that (a) in line with Strijkers et al. (2011), goal-directed processes play a crucial role very early on in speech production, and (b) these task-driven top-down influences function at least in a domain-general manner by modulating those networks which are always relevant for the production of language, irrespective of which cortical pathways are triggered by the input.

Keywords: speech production, goal-directed processing, lexical access, ERPs, categorization

INTRODUCTION

A substantial body of research in the last two decades has provided compelling evidence that top-down processes can have a powerful influence on certain early low-level perceptual processes (e.g., Corbetta et al., 1990; Posner and Dehaene, 1994; Desimone and Duncan, 1995; Luck et al., 1997; Hillyard et al., 1998; Kastner et al., 1999; Bar, 2003; Bar et al., 2006; Gilbert and Sigman, 2007). These findings have been crucial with respect to the way we conceptualize visual processing, in that the traditional automatic bottom-up view has changed dramatically when taking into account goal-directed behavior related to a visual act. That is, recognizing visual input does not solely proceed in a unidirectional manner from lower to higher levels of representation, but rather is achieved through the dynamical interplay of stimulus-driven processes with early top-down influences facilitating access to those representations that are relevant to the desired behavior. However, the role top-down processes play in regulating higher goal-directed behaviors, such as producing speech, has not received nearly as much attention. This would appear to be an important issue to explore though, since speaking is in principle an intention-driven activity. A speaker knows beforehand, at least most of the time, whether he/she wants to verbalize or not certain ideas, thoughts, and stimuli in his environment. Consequently, it seems reasonable to assume that these *a priori* goal-directed settings should exert an important influence on the way speech planning proceeds. Nevertheless, most speech production models assume that the influence of task intentions come about relatively late, namely after word selection (e.g., Dell, 1986; Caramazza, 1997; Levelt et al., 1999). Such a view is clearly in disagreement with the more dynamical accounts of brain processing that are emerging in the fields of vision and object recognition. One of the reasons for this discrepancy might be that the temporal role of top-down influences of task intentionality has never been systematically investigated for language behavior. To this end, we aimed at increasing our understanding of when goal-directed influences associated with speech production affect the course of processing. In particular, we compared the brain's electrical response (event-related potentials; ERPs) in a task where one has the conscious intention to articulate the name of the stimulus (naming task) vs. a task where no naming intention is present semantic categorization task (SCT) both for non-linguistic stimuli (pictures) and linguistic stimuli (written words). By doing so, we aimed to see when the brain starts differentiating the same visual input as a function of task intention, and whether these top-down influences vary depending on the type of processing required for different input modalities.

There is a long-standing tradition in the field of language production that has explored whether items that a speaker does not intend to name nevertheless activate lexical representations. Within this tradition there have been a large number of studies investigating whether distractor items (items we do not wish to utter) affect the speed with which we name a target item, as well as a smaller number of studies assessing the presence of linguistic effects in non-verbal tasks (e.g., Kroll and Potter, 1984; Glaser and Glaser, 1989; Schriefers et al., 1990; Levelt et al., 1991; Roelofs, 1992, 2003, 2008; Jescheniak and Levelt, 1994; Starreveld and La Heij, 1996; Meyer et al., 1998, 2007; Cutting and Ferreira, 1999; Costa and Caramazza, 2002; Morsella and Miozzo, 2002; Navarrete and Costa, 2005; Bles and Jansma, 2008). However, unlike the current study, these studies did not aim at uncovering how goal-directed speech behavior influences the course of processing within the system. Instead, they sought to assess whether activated concepts automatically transmit information to the lexical system and up to which level of representations this spreading activation extends. For instance, in linguistic Stroop-like tasks it is frequently reported that distractors having a semantic or phonological relationship with a target affect the speed with which that target is named. From these findings it has been concluded that both speech intended and speech non-intended concepts activate the lexicon in parallel. Such a conclusion is in accordance with models of lexical access that embrace the principle of spreading activation, according to which activated concepts percolate to the lexical system regardless of the intention of a speaker (e.g., Dell, 1986; Caramazza, 1997; Levelt et al., 1999; but see Bloem and La Heij, 2003). However, what has not been addressed is when and how intentional processes do come into play in order to eventually produce the desired behavior. In other words, previous comparisons between items we intend to verbalize and items we do not intend to verbalize have mainly served the purpose of improving our understanding of how concepts activate lexical representations (and beyond) during language production, but this line of research did not aim to uncover how this lexicalization process or even earlier processes leading up to lexical access may interact with task intentional processing. Therefore, in order to develop accounts of language production which do not solely accommodate language-related processing, but instead try to incorporate this within the broader spectrum of human information processing, it is important to study how language processing progresses in relation to goal-directed behavior.

Recently, Strijkers et al. (2011) investigated this question explicitly and reported evidence that the top-down intention to speak seems to affect the language system in a proactive manner; that is, prior to the spreading activation between concepts and words. Specifically, these authors compared the brain's electrical response for a variable known to affect lexical access, namely word frequency, during overt object naming and non-verbal object categorization. They found that during naming, ERPs elicited for objects with low frequency names started to diverge from those with high frequency names as early as 152 ms after stimulus onset (pP2), while during non-verbal categorization the same frequency comparison appeared 200 ms later at a qualitatively distinct component (N400). Two important conclusions were drawn from these findings: first, in line with spreading activation models of lexical access, activation in semantic representations percolates to the lexical system regardless of a speaker's intention; Second, and in contrast to the predictions of most language production models, initial access to the lexicon is instigated by the top-down intention to speak. Put differently, when there is conscious intention to name an object the brain will engage substantially faster in lexical access compared to when no such intention is present due to top-down signals pre-activating the lexical system prior to the conceptually driven activation of words. To our knowledge, this was the first study to directly demonstrate the vital and early role top-down processes play in facilitating the retrieval of words one intends to utter. This result places novel constraints on language production models in that initial access to the lexicon from *activated* concepts is not as automatic as originally thought, but can be better seen as a dynamical process driven by goal-directed intentions.

Here we aimed to extend the findings of Strijkers et al. (2011) by examining when the goal-directed operations determined by task intention penetrate stimulus-evoked processing. Like Strijkers et al. (2011) we compared the time-course of ERP effects in a task requiring speech production (naming) vs. a task where speech is not necessary (semantic categorization). However, rather than exploring how the intention to speak may interact with a particular production operation, here we sought to directly ascertain the temporal dissociations related to task intention. Exploring the electrophysiological signature of task intention in isolation, should enable us to pinpoint when the most evident ERP deflections associated with naming and categorization intention occur. This is relevant for two reasons: first, it can provide an independent test of the conclusions reached by Strijkers et al. (2011). In that study, the authors argued that top-down intention to speak penetrates the lexico-semantic system prior to concept selection. If this conclusion is correct, ERP differences elicited by task intention in general (naming vs. categorization) should occur prior to, or in the temporal vicinity of the time-course uncovered by Strijkers and colleagues (i.e., around 150–200 ms after stimulus onset). Second, given that no particular linguistic operation such as lexical access is targeted, we will be able to provide a more general assessment of the processing differences between speech production and semantic categorization. In this way, potentially important ERP effects other than those that are lexically driven (for instance, already during visual and/or conceptual processing) may become apparent.

An important aspect of the current experiment is that we also explored whether top-down involvement differs as a function of presentation modality. By comparing effects across modalities we should be able to ascertain whether the top-down intention to speak affects the language network in a modality-general manner or whether this process elicits distinct modulations depending on the input. Compared to picture naming, where the production of speech entails active retrieval from memory, word naming directly conveys the linguistic information which has to be uttered (i.e., written words are automatically associated with the required output). As a consequence, early top-down processes that facilitate lexical retrieval may be especially relevant for more demanding processing situations such as picture naming compared to the more predominantly stimulus-driven processing associated with word naming. Given these differences in linguistic vs. non-linguistic input, and consequently, differences in the dynamics for retrieving the same words between the two tasks, it will be interesting to see whether goal-directed influences display the same or distinct time-courses.

In two blocks of trials undergraduate participants were presented with a combination of black and white pictures of common objects (picture condition) and the written words that were the names of objects (word condition). In one block participants were asked to rapidly name (naming task block), and in another block they were asked to categorize (SCT) the words or objects. ERPs time-locked to the onset of words and pictures in both blocks were recorded along with the overt response. This approach has only recently been employed successfully to study language production (e.g., Christoffels et al., 2007; Koester and Schiller, 2008; Chauncey et al., 2009; Costa et al., 2009; Dell'Acqua et al., 2010; Strijkers et al., 2010). The design was a two (Modality: words vs. pictures) by two (Task: naming vs. semantic categorization) factorial. In light of the results of Strijkers et al. (2011) we predicted electrophysiological deflections within 200 ms of stimulus onset when depicted objects had to be named compared to being categorized. For printed word naming, if the top-down mechanisms related to task intention are qualitatively similar across modalities and processing dynamics, we expect to see the same early, and perhaps even slightly earlier, ERP dissociations between naming and categorization as for pictures. If, however, the influences generated by the higher goal-directed processes operate as a function of a particular input and/or whether the task-relevant representations have to be retrieved internally (pictures), then distinct and potentially later ERP modulations for written words compared to pictures might be obtained.

MATERIALS AND METHODS

PARTICIPANTS

Twelve native English speakers (six females, mean age 18.67 years) participated and were compensated for their time. All participants were right-handed, with normal or corrected-to-normal visual acuity and no history of neurological insult or language disability.

STIMULI

The picture stimuli consisted of 184 black and white line drawings of common objects, selected from a standardized inventory (Snodgrass and Vanderwart, 1980). The word stimuli were 184 English words that corresponded to the line drawings. Of these, 24 words and 24 images were so-called "probe" items. The ERPs to probe items were not analyzed. All stimuli were presented in white on a black background. Both pictures and words were presented together in two mixed blocks (see below), arranged in a pseudo-random order to prevent expectation and priming effects.

PROCEDURE

Participants were seated in a comfortable armchair facing a computer monitor in a sound-attenuated room for electrode placement. Each trial started with a fixation cross in the middle of the screen for 500 ms and a blank screen for 500 ms. The stimulus (a picture or a word) then appeared for 400 ms, followed by a 1100 ms blank screen, and a blink signal for 1700 ms. This was followed by another blank screen for 500 ms and the fixation cross for the next trial (see **Figure 2** for examples of both trial types). Participants were asked to blink during the blink signal if necessary, and minimize eye movements for the rest of the time. There were two scheduled 1-min breaks during each block of the experiment.

In the first block participants engaged in a go/no–go SCT, in which they were instructed to press a button whenever they saw either a picture or a word referring to a human body part (so-called "probe" items). Probes made up 12% of trials with equal numbers of picture and word body parts. No response was required to the remaining 88% of critical stimuli which were averaged in the ERPs reported here.

In a second block of trials participants were told to rapidly name all picture and word stimuli. The exact same trial structure was used in this second block, although the word and picture stimuli were switched such that all items that were formerly pictures were now words and vice versa. Stimuli that appeared in picture format for one participant for a given task appeared in word format for another participant, and vice versa, and the order of the two blocks (semantic categorization and naming) was counterbalanced across participants.

EEG RECORDINGS

Electroencephalograms were collected using 32-channel caps (Electro-cap International). The tin electrodes were arranged according to International 10–20 system (see **Figure 1**). In addition, an electrode below the left eye (LE) was used to monitor for blinks and vertical eye movements and an electrode beside the right eye (HE) was used to monitor for horizontal eye movements. Two electrodes were placed behind the ears on the mastoid bone; the left mastoid site (A1) was used as a reference for all electrodes, and the right mastoid site (A2) was recorded to evaluate differential mastoid activity. Impedance was kept at less than 5 kΩ for all electrode sites except the lower eye channel, which was below 10 kΩ. The EEG was amplified using an SA Bioamplifier (SA Instruments, San Diego, CA, USA) operating on a bandpass of 0.01 and 40 Hz. The digitizing computer continuously sampled

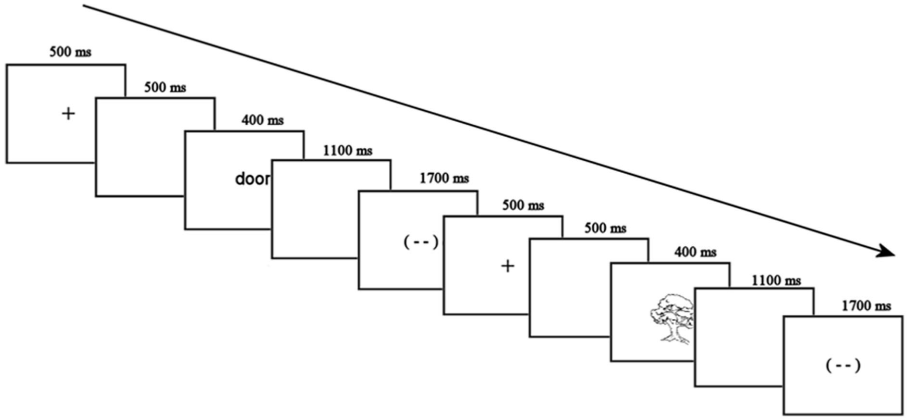


FIGURE 1 | Two sample trials, one with a word stimulus and one with a picture stimulus.

the EEG at a rate of 200 Hz while a stimulus computer simultaneously presented stimuli to a 19″ CRT monitor located 54″ in front of the participant (all stimuli subtended less than 7° of horizontal visual angle).

DATA ANALYSIS

Averaged ERPs were computed for all word and picture stimuli for each participant in each block of trials (task) at the 29 scalp sites shown in **Figure 2**. Epochs with eye movement artifacts between −100 and 600 ms post stimulus onset were excluded prior to averaging. The resulting ERPs were baselined between −100 and 0 ms and re-referenced to the average of the 29 scalp sites (average reference – Joyce and Rossion, 2005). The resulting ERP data were measured by calculating mean amplitudes within three latency windows: 150–250 ms and 250–350 ms for picture and word stimuli and also between 350 and 450 ms for picture stimuli (the late window was not used with the word stimuli because a substantial number of trials in the naming task had significant speech artifact starting just after 350 ms post-word onset – see the anterior sites in **Figure 5**).

The data were analyzed with repeated measures ANOVAs with four independent variables: TASK (SCT vs. Naming), MODALITY (words vs. pictures), ANTERIOR–POSTERIOR (prefrontal vs. frontal vs. temporal–parietal vs. occipital for average reference data and frontal, vs. central vs. parietal vs. occipital for mastoid reference data), and LATERALITY (left vs. midline vs. right – see **Figure 2** for the electrode sites included in the analysis). Significant interactions involving the TASK and MODALITY factors were followed up with planned ANOVAs breaking down the interaction.

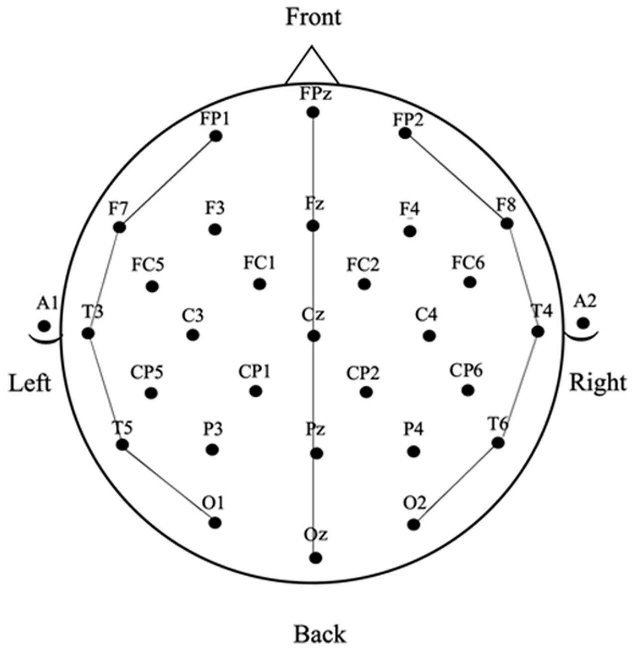


FIGURE 2 | Thirty-two-channel electrode montage.

RESULTS

BEHAVIORAL ANALYSES

The behavioral analyses were restricted to the naming data (given that for the categorization experiment the critical trials – no–go trials – did not require a behavioral response). Here it was observed that word naming resulted in faster reaction times (520 ms; SD: 69 ms) compared to picture naming (790 ms; SD: 78 ms). A paired t-test between word and picture naming confirmed that this difference was significant [$t(10) = 29.12$, $p < 0.001$].

ERP ANALYSES

150–250 ms. As can be seen in **Figure 3**, in this epoch there are substantial differences between ERPs to stimuli presented during a SCT compared to the same items presented during a naming task. This observation is supported by both a main effect of TASK [$F(1,11) = 5.41$, $p = 0.04$] as well as an interaction between TASK and the two scalp site variables [TASK × ANT–POST × LATERALITY: $F(8,88) = 4.39$, $p = 0.0002$]. **Figures 4** and **5** reveal that the effects of TASK differed as a function of MODALITY (words vs. pictures) and scalp site [TASK × MODALITY × ANT–POST × LATERALITY: $F(8,88) = 3.23$, $p = 0.0029$]. To better understand this interaction we conducted follow-up analyses examining the two modalities separately. In the analyzes of ERPs recorded to picture stimuli we found significant differences between the TASKS as a function of scalp site [TASK × ANT–POST: $F(4,44) = 3.37$, $p = 0.0172$]. This interaction was due to the naming task producing more negative-going ERPs over posterior sites but more positive-going ERPs over the most anterior sites than the SCT (see the FP row of sites in **Figure 4**). It is important to note that there was no trend for this effect to differ as a function of laterality (i.e., TASK × ANT–POST × LATERALITY, and TASK × LATERALITY, ps > 0.1). For the word stimuli there was a similar pattern of posterior negativity and anterior positivity for naming compared to semantic categorization [TASK × ANT–POST: $F(4,44) = 3.66$, $p < 0.07$], but this trend tended to be greater over the left hemisphere and midline toward the back of the head [O1, Oz, and T5, TASK × ANT–POST × LATERALITY: $F(8,88) = 7.06$, $p < 0.0001$ – see voltage maps in the bottom left panel of **Figure 6**]. In summary, an N170-like negativity at posterior sites was larger for naming than semantic categorization, and this effect tended to be larger over the left than right hemisphere, but only for word stimuli.

250–350 ms. Differences between conditions continued into this second measurement window. There was not, however, a significant main effect of MODALITY as in the previous window, although this factor did interact with the TASK variable [TASK × MODALITY: $F(1,11) = 5.27 = 0.0423$]. There were also significant [TASK × MODALITY × LATERALITY and TASK × MODALITY × ANT–POST × LATERALITY interactions ($F(2,22) = 7.14$, $p = 0.0041$; $F(8,88) = 4.51$, $p = 0.0001$, respectively]. To better understand these interactions we conducted follow-up analyses for the two modalities separately. For the picture modality we found a significant interaction of the TASK variable with both of the scalp site variables [TASK × ANT–POST × LATERALITY: $F(8,88) = 3.33$, $p = 0.0022$]. Pictures in the naming task tended to produce more negative-going ERPs over left posterior and central sites, but more positive-going ERPs

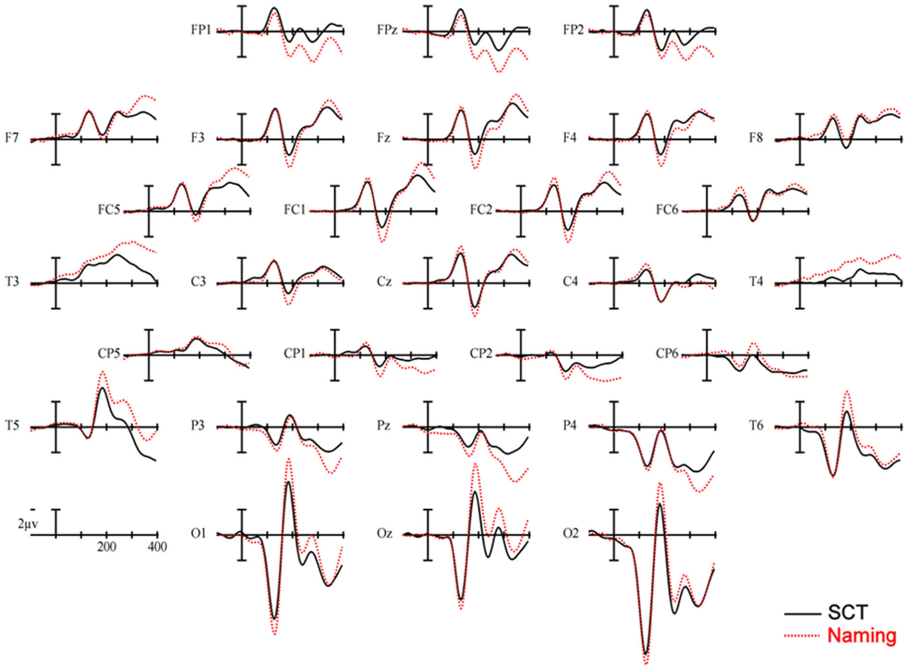


FIGURE 3 | Main effect of TASK. ERPs collapsed across item type. Solid lines are ERPs from all 28 scalp sites during a semantic categorization task (SCT) and dashed lines are ERPs recorded to the same stimuli in a naming task. Stimulus onset is the vertical calibration bar and each tic mark on the *x*-axis is 100 ms.

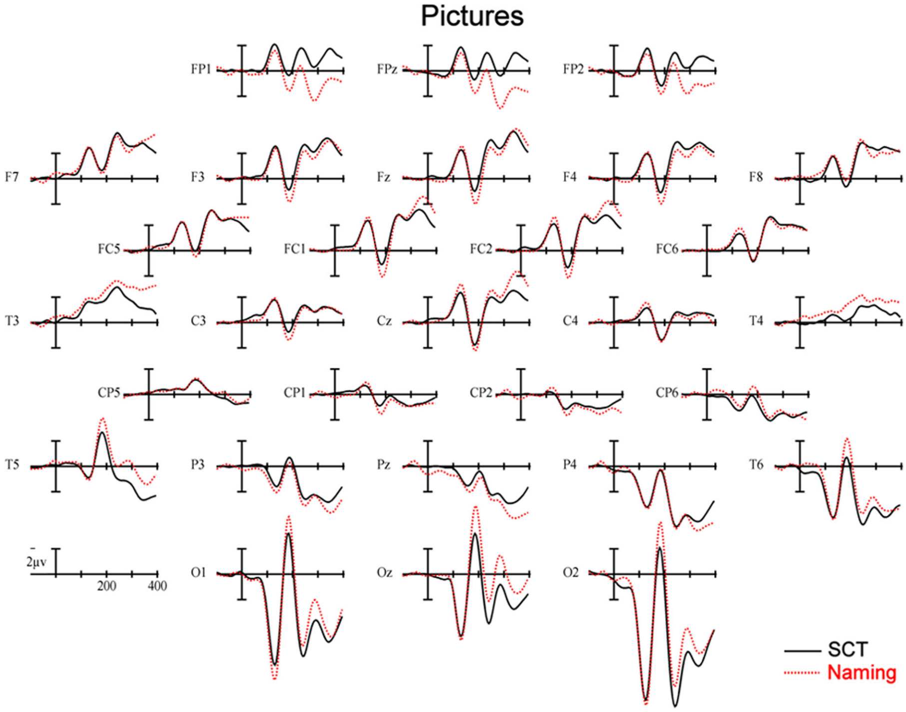


FIGURE 4 | Event-related potentials s recorded to pictures of objects, solid lines are ERPs during a semantic categorization task (SCT) and dashed lines are ERPs recorded in a picture naming task.

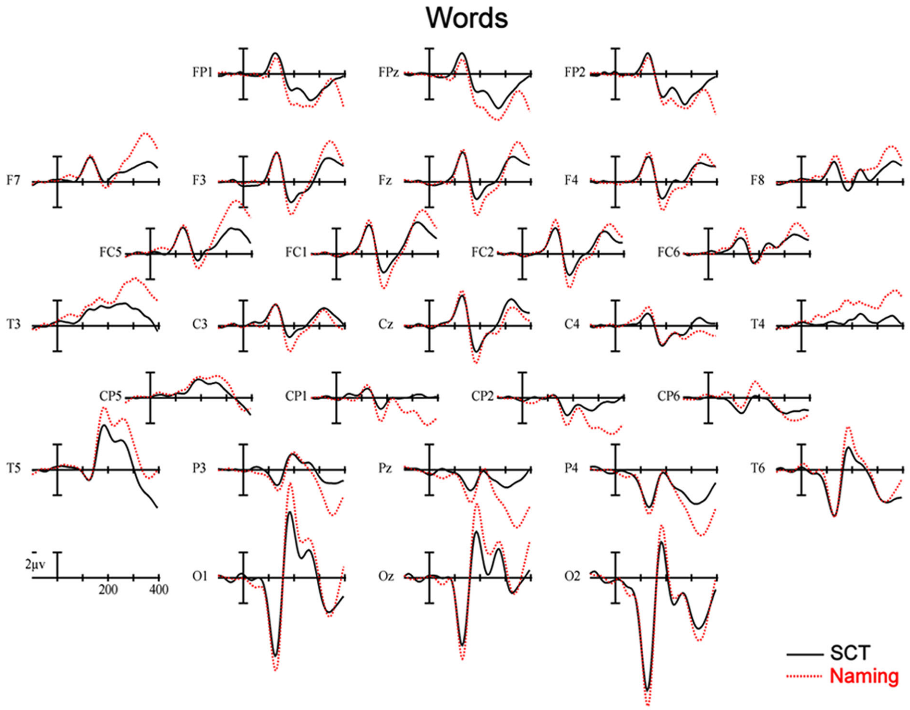


FIGURE 5 | Event-related potentials recorded to words, solid lines are ERPs during a semantic categorization task (SCT), and dashed lines are ERPs recorded in a word naming task.

over midline and left hemisphere sites than did ERPs to the same items during the SCT (see **Figure 4**). While the basic pattern of TASK effects was similar for words, the left lateralized nature of the posterior negativity and anterior positivity was stronger than for pictures [TASK × LATERALITY: $F(2,22) = 5.79$, $p = 0.012$; TASK × ANTERIOR–POSTERIOR × LATERALITY: $F(8,88) = 8.12$, $p < 0.00001$].

350–450 ms. As previously mentioned, because of the earlier onset of articulation in word naming, we only analyzed the ERP data for picture stimuli in this epoch. As can be seen in **Figure 4**, ERPs to pictures in this epoch tended to be more negative-going in the naming task than the SCT at most scalp sites with the notable exception of the most anterior electrodes (FP1, FPz, and FP2) and Pz. At these sites the naming tasks generated clearly more positive-going ERPs than did the SCT [TASK × ANT–POST × LATERALITY interaction: $F(8,88) = 4.29$, $p = 0.0002$]. As can be seen in **Figure 6**, the pattern for picture stimuli in this epoch is similar to the pattern seen in the previous epoch for words (compare **Figure 6** bottom right with top right).

DISCUSSION

We found that ERPs elicited by the naming task started to diverge very early on (~170 ms) from those elicited by the SCT, and these effects were qualitatively similar for pictures and words, aside from some hemispheric differences potentially reflecting the distinct visual input. Concretely, naming elicited enhanced negative brain responses peaking around 170 ms at occipital electrodes (N170) and enhanced positive amplitudes peaking around 200 ms at anterior sites (P2) compared to semantic categorization. In later time-windows some expected differences between pictures and words became apparent: (a) picture naming implied additional frontal ERP enhancements, arguably reflecting a more effortful word retrieval for the former; (b) word naming resulted in an earlier onset of motor preparation than picture naming did, nicely mimicking the behavioral naming latencies. The results demonstrate that within 200 ms, top-down processes associated with task intentionality start influencing the course of brain processing. The data extends the increasing body of evidence from other modalities demonstrating that cognitive processing arises as a consequence of the early interplay between externally and internally driven processes (e.g., Desimone and Duncan, 1995; Ullman, 1995; Kastner and Ungerleider, 2000; Engel et al., 2001; Gilbert and Sigman, 2007) to the realm of language processing, and more specifically is in agreement with recent work arguing that the brain's fast engagement in lexical access is driven by the intention to speak (Strijkers et al., 2011).

In addition, the current results suggest that these early goal-directed modulations play a central part in the flow of activation regardless of the modality in which participants had to perform the task. This finding is consistent with the hypothesis that the mechanism behind these top-down influences is domain-general. That is, the manner in which the task intentional mechanism guides activity rapidly toward relevant representations is independent of the input processes[1]. A straightforward manner in which to interpret

---

[1] Note that, under the right conditions, this does not exclude the potential involvement of input specific top-down processes in speech production as well.

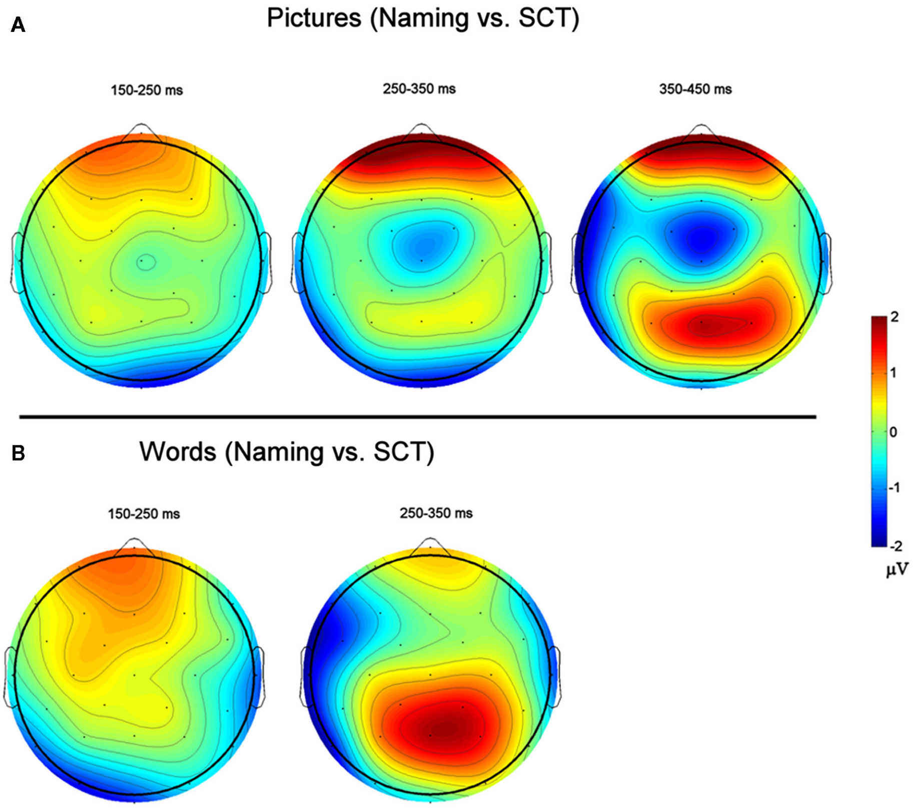


FIGURE 6 | Voltage Maps based on difference waves calculated by subtracting ERPs recorded to items during the semantic categorization task (SCT) from ERPs recorded to items recorded during naming. (A) (Top) are for picture stimuli, and **(B)** (bottom) are for word stimuli.

this finding is to assume that the intention to speak sets the brain in a general naming state, enhancing those representations which are always relevant for speech production regardless which cortical pathways are triggered by the input. A similar scenario unfolds for the SCT in that the system is tuned toward rapid activation of those representations that are crucial for classifying items as belonging to a particular category. In support of this notion is the fact that ERP task differences are similar for the two input modalities, even before the deflections associated with the different inputs become apparent. A final conclusion that can be derived from the similar early enhancements caused by task intention across input modalities is that top-down processes are not simply a helpful asset for the brain to calculate a motor program for speech production. Rather it suggests that early top-down processing forms an integral part of any speech production act. To summarize, the main contribution of the present study is the demonstration that the intentional act of producing speech requires the early involvement of domain-general top-down processes enhancing information relevant for producing speech irrespective of the input and the processing dynamics associated with that input. In what follows we will tentatively suggest potential loci where these goal-directed influences may take place in the course of speech planning.

A first potential locus for the early influence of top-down processes identified in the present study is at the lexical level. As mentioned in the Introduction, Strijkers and colleagues already demonstrated that in picture naming, lexical access is facilitated in a proactive manner by the intention to speak. Given the overlapping time-course (∼150–200 ms) between the ERP effects reported in that study and the current study, a similar conclusion as proposed by Strijkers et al. (2011) could be entertained here. In the case of naming, a task intentional mechanism will *a priori* increase the baseline activity of the lexical system so that when the actual stimulus is presented, there is facilitated and privileged access to word representations in order to rapidly and efficiently produce speech. In the case of semantic categorization, there is no need for top-down signals associated with task intention to enhance speech-related representations. Instead, goal-directed activity will "push" the sensory-driven processing of the incoming stimulus toward those semantic features that are relevant for deciding to which category the input belongs to, allowing us to categorize objects independently from the lexical information associated with that object (e.g., Dell'Acqua and Grainger, 1999).

Within this view, the ERP components we identified (or at least one of them), an anterior P2 and a posterior N170, can be conceived as reflecting the task intentional mechanism itself. That is, if the ERP modulations would reflect the top-down signals acting upon a particular process, lexical access in the present framework, we should have seen a similar pP2 (descriptively labeled *production* P2–pP2) modulation as reported by Strijkers et al. (2011), at least in the case of pictures. This posteriorly distributed positive-going component, which has been shown to be sensitive to a range of lexical variables (Costa et al., 2009; Strijkers et al., 2010), was selectively modulated for naming intention (Strijkers et al., 2011). Strijkers et al. proposed that its modulation likely is engendered by the interaction between top-down processing and lexical activation levels. Here, no such pP2 modulation between naming and categorization was found, but instead a frontally maximal P2 and an occipital N170 were elicited. Thus, if we wish to maintain that these cortical deflections represent top-down influences which affect lexicalization, we must assume that the particular electrophysiological expression(s) observed here are associated with the top-down processes in isolation. If so, naming appears to require increased top-down involvement, which seems reasonable given that naming is a process requiring a much more specific behavioral response compared to semantic categorization.

Alternatively, the task driven ERP differences uncovered here may not be reflecting the same top-down signals as in the study of Strijkers et al. (2011). In that study, a particular linguistic stage was targeted by manipulating lexical frequency, whereas in the present study goal-directed influences might be acting upon other levels of processing. This conclusion finds support in a comparison of the effects seen in two early components affected by task in the present study, the anterior P2 and posterior N170, with the results from studies in other fields shown to modulate these components. In visual search paradigms the amplitude of the P2 is larger for attended stimuli and target-relevant information, which has been proposed to reflect attention-driven enhancements of the perceptually-relevant features of the input (e.g., Hillyard and Munte, 1984; Luck and Hillyard, 1994; Mangun and Hillyard, 1995). In language comprehension, similar P2 amplitude increases are found when words or pictures are highly expected in a given sentence, or follow related prime words (e.g., Federmeier and Kutas, 2002, 2005; Federmeier et al., 2005, 2007; Federmeier, 2007). As in vision, these effects are thought to index top-down driven anticipatory activation of the perceptual features related to an expected upcoming word. Given that the current P2 modulation, and in contrast to the pP2 encountered by Strijkers et al. (2011), has a similar frontal distribution as the P2 modulations reported in the above studies, it is possible that we are dealing with a similar effect as encountered in visual perception and language comprehension. In this case, our findings would suggest that the intention to speak can already affect processes related to the input. Such findings would be highly intriguing since they would indicate that the mere intention to engage in speech behavior not only influences the manner in which we retrieve lexical representations (Strijkers et al., 2011), the crucial units which must be retrieved for engaging in a speech act, but already alters the way our brain reacts to the input. Put differently, the intention to speak would change the manner in which we "perceive" an object and a word very early on compared to the visuo-conceptual processing of the same stimuli when the perceptual and/or semantic goals are different from those required for a speech act.

Interestingly, the fact that another early component, the N170, was affected by task intention, fits nicely with this conclusion. Typically, this occipito-temporal negative going component is enhanced for objects with which we have great expertise, such as faces (e.g., Bentin et al., 1996; Tanaka and Curran, 2001; Gauthier et al., 2003; Rossion et al., 2003; but see Liu et al., 2002; Thierry et al., 2007). It is worth mentioning, especially in the current context, that a few studies also reported similar N170-like modulations for expertise with written words (e.g., McCandliss et al., 2003; Yum et al., 2011). If the N170 modulation reported here can be considered similar to the one thought to be sensitive to visual expertise, this would mean that naming leads to more specific visual identification processes compared to categorization, which is not that surprising given the differences in task demands. To sum up, the fact that the input-related N170 was affected by differences in task intention rather than differences between stimuli provides support for our tentative proposal that the intention to speak may already affect the manner in which we visually and/or semantically process stimuli. This is a highly interesting possibility which clearly merits further investigation in future work.

CONCLUSION

In the present study, by comparing brain responses elicited by the same visual input in two different tasks, we were able to establish that goal-directed top-down influences penetrate and affect early processing of words and objects. These results concur with the evidence reported in a recent study (Strijkers et al., 2011) emphasizing the crucial role of early intention-driven processes for the production of speech. In addition, the results of the current study showed that these top-down signals are, at least partly, generated by a domain-general system, likely to be functional each time we want to speak. Whether these influences operate at the level of word representations, as demonstrated in previous work, or even during visuo-conceptual processing could not be determined explicitly on the basis of the present results. Nevertheless, the current results contribute some important insights regarding a topic which has received little attention in the literature so far, and paves the way for future research that will help to gradually narrow down the functional nature and sources of goal-directed processing in language production.

ACKNOWLEDGMENTS

Kristof Strijkers was supported by a pre-doctoral fellowship from the Spanish government (FPU; 2007–2011) and the project Consolider-Ingenio 2010 (CSD 2007–00012). Phillip Holcomb was supported by Grants HD25889 and HD043251.

REFERENCES

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

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

Bentin, S., McCarthy, G., Perez, E., Puce, A., and Allison, T. (1996). Electrophysiological studies of face perception in humans. *J. Cogn. Neurosci.* 8, 551–565.

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

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

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

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

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

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., and Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* 248, 1556–1559.

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

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

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

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

Dell'Acqua, R., and Grainger, J. (1999). Unconscious semantic priming from pictures. *Cognition* 73, B1–B15.

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

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

Engel, A. K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* 2, 704–716.

Federmeier, K. D. (2007). Thinking ahead: the role and roots of prediction in language comprehension. *Psychophysiology* 44, 491–505.

Federmeier, K. D., and Kutas, M. (2002). Picture the difference: electrophysiological investigations of picture processing in the twocerebral hemispheres. *Neuropsychologia* 40, 730–747.

Federmeier, K. D., and Kutas, M. (2005). Aging in context: age-related changes in context use during language comprehension. *Psychophysiology* 42, 133–141.

Federmeier, K. D., Mai, H., and Kutas, M. (2005). Both sides get the point: hemispheric sensitivities to sentential constraint. *Mem. Cognit.* 33, 871–886.

Federmeier, K. D., Wlotko, E., De Ochoa-Dewald, E., and Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Res.* 1146, 75–84.

Gauthier, I., Curran, T., Curby, K. M., and Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nat. Neurosci.* 6, 428–432.

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

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

Hillyard, S. A., and Munte, T. F. (1984). Selective attention to color and location: an analysis with event-related brain potentials. *Percept. Psychophys.* 36, 185–198.

Hillyard, S. A., Vogel, E. K., and Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 353, 1257–1267.

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

Joyce, C., and Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: the effect of reference electrode site. *Clin. Neurophysiol.* 116, 2613–2631.

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

Kastner, S., and Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annu. Rev. Neurosci.* 23, 315–341.

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

Kroll, J. F., and Potter, M. C. (1984). Recognizing words, pictures, and concepts: a comparison of lexical, object, and reality decisions. *J. Verbal Learn Verbal Behav.* 23, 39–66.

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

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

Liu, J., Harris, A., and Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. *Nat. Neurosci.* 5, 910–916.

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

Luck, S. J., and Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology* 31, 291–308.

Mangun, G. R., and Hillyard, S. A. (1995). "Mechanisms and models of selective attention," in *Electrophysiology of Mind*, eds M. D. Rugg and M. G. H. Coles (Oxford: Oxford University Press), 40–85.

McCandliss, B. D., Cohen, L., and Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci.* 7, 293–299.

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

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

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

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

Posner, M. I., and Dehaene, S. (1994). Attentional networks. *Trends Neurosci.* 17, 75–79.

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

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

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

Rossion, B., Joyce, C. A., Cottrell, G. W., and Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage* 20, 1609–1624.

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

Snodgrass, J. G., and Vanderwart, M. (1980). A standardized set of 260 pictures: norm for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol. Hum. Learn. Mem.* 6, 174–215.

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

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

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

Tanaka, J. W., and Curran, T. (2001). A neural basis for expert object recognition. *Psychol. Sci.* 12, 43–47.

Thierry, G., Martin, C. D., Downing, P., and Pegna, A. J. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nat. Neurosci.* 10, 505–511.

Ullman, S. (1995). Sequence seeking and counter streams: a computational model for bidirectional information flow in the visual cortex. *Cereb. Cortex* 1, 1–11.

Yum, Y. N., Holcomb, P. J., and Grainger, J. (2011). Words and pictures: an electrophysiological investigation of domain specific processing in native Chinese and English speakers. *Neuropsychologia* 49, 1910–19220.

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

Received: 25 May 2011; accepted: 24 November 2011; published online: 09 December 2011.

Citation: Strijkers K, Yum YN, Grainger J and Holcomb PJ (2011) Early goal-directed top-down influences in the production of speech. Front. Psychology ***2****:371. doi: 10.3389/fpsyg.2011.00371*

This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 16 September 2011
doi: 10.3389/fpsyg.2011.00237

A generative model of speech production in Broca's and Wernicke's areas

***Cathy J. Price*[1]*, *Jenny T. Crinion*[2] *and Mairéad MacSweeney*[2]**

[1] *Wellcome Trust Centre for Neuroimaging, University College London, London, UK*
[2] *Institute of Cognitive Neuroscience, University College London, London, UK*

Edited by:
Albert Costa, University Pompeu Fabra, Spain

Reviewed by:
Xing Tian, New York University, USA
Pascale Tremblay, The University of Trento, Italy

****Correspondence:***
Cathy J. Price, Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, UK.
e-mail: c.price@fil.ion.ucl.ac.uk

Speech production involves the generation of an auditory signal from the articulators and vocal tract. When the intended auditory signal does not match the produced sounds, subsequent articulatory commands can be adjusted to reduce the difference between the intended and produced sounds. This requires an internal model of the intended speech output that can be compared to the produced speech. The aim of this functional imaging study was to identify brain activation related to the internal model of speech production after activation related to vocalization, auditory feedback, and movement in the articulators had been controlled. There were four conditions: silent articulation of speech, non-speech mouth movements, finger tapping, and visual fixation. In the speech conditions, participants produced the mouth movements associated with the words "one" and "three." We eliminated auditory feedback from the spoken output by instructing participants to articulate these words without producing any sound. The non-speech mouth movement conditions involved lip pursing and tongue protrusions to control for movement in the articulators. The main difference between our speech and non-speech mouth movement conditions is that prior experience producing speech sounds leads to the automatic and covert generation of auditory and phonological associations that may play a role in predicting auditory feedback. We found that, relative to non-speech mouth movements, silent speech activated Broca's area in the left dorsal pars opercularis and Wernicke's area in the left posterior superior temporal sulcus. We discuss these results in the context of a generative model of speech production and propose that Broca's and Wernicke's areas may be involved in predicting the speech output that follows articulation. These predictions could provide a mechanism by which rapid movement of the articulators is precisely matched to the intended speech outputs during future articulations.

Keywords: speech production, auditory feedback, PET, fMRI, forward model

INTRODUCTION

Speech production is a complex multistage process that converts conceptual ideas into acoustic signals that can be understood by others. The stages include conceptualization of the intended message, word retrieval, selection of the appropriate morphological forms, sequencing of phonemes, syllables, and words, phonetic encoding of the articulatory plans, initiation, and coordination of sequences of movements in the tongue, lips, and laryngeal muscles that vibrate the vocal tract, and the control of respiration for vowel phonation and prosody. In addition to this feed forward sequence, auditory, and somatosensory processing of the spoken output is fed back to the motor system for online correction of laryngeal and articulatory movements (Levelt et al., 1999; Guenther et al., 2006). This self monitoring process is thought to be essential for learning to speak in a first (native) or second language but also plays a role in adult/fluent speech production, particularly when the auditory feedback is distorted. The sensorimotor interactions involved in monitoring the spoken response require an internal model of the intended speech to which the output can be matched (Borden, 1979; Paus et al., 1996; Heinks-Maldonado, 2005). The aim of the current study was to identify brain responses related to the internal model and to consider how these responses might predict auditory output prior to auditory or sensorimotor feedback.

The concept of internal models that predict the sensory consequences of an action is not specific to speech production. In the motor system, internal models that finesse motor control are referred to as "forward models" (Miall, 1993; Wolpert et al., 1995). More generally, forward models are examples of generative models that the brain may use for both perception (Helmholtz, 1866; MacKay, 1956; Gregory, 1980; Ballard et al., 1983; Friston, 2001, 2005) and active inference (Friston, 2010). The underlying principle of a generative model of brain function is that higher-level systems predict the inputs to lower-levels; and the resulting prediction error is then used to optimize future predictions – a scheme known as predictive coding.

Recent accounts of forward models in speech production have varied in how the auditory predictions and feedback are implemented (see **Figure 1**). For example, in the model proposed by Tian and Poeppel (2010), a motor efference copy is generated during motor planning and fed into a forward model of motor

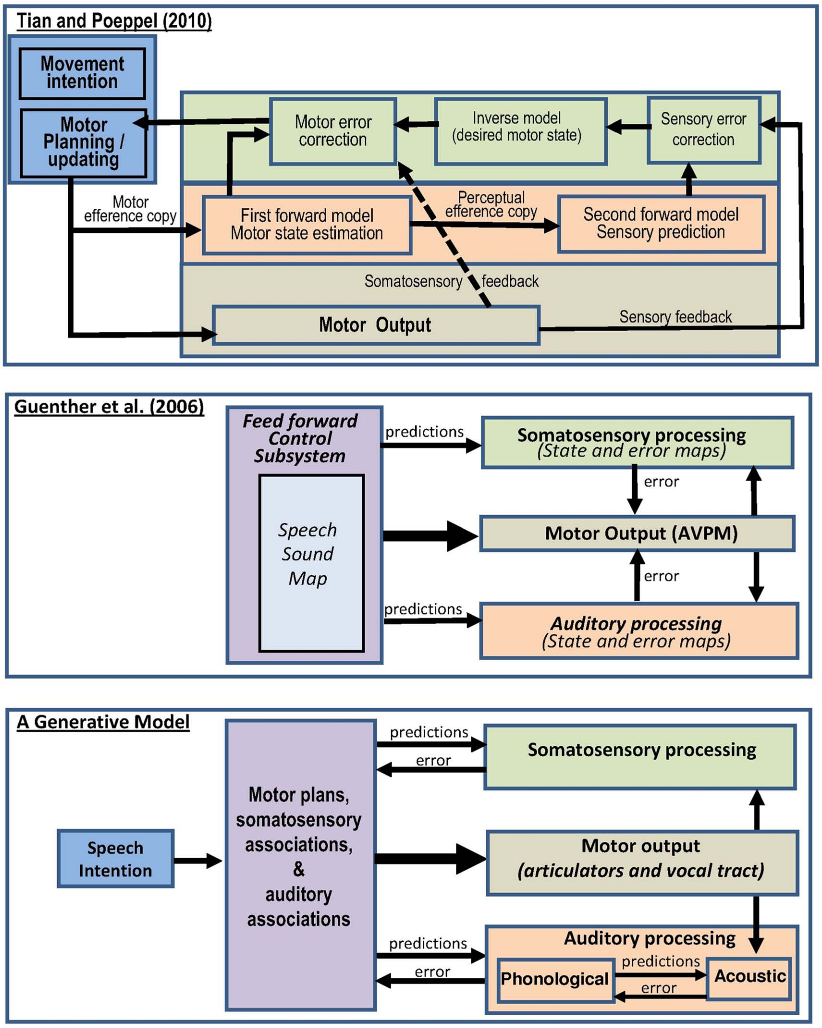


FIGURE 1 | Three different implementations of internal models of speech production. Top: Tian and Poeppel (2010) proposed model of motor control based on internal forward models and feedback. Middle: hypothesized processing stages involved in speech acquisition and production according to the DIVA model (directions into velocities of articulators), adapted from Guenther et al. (2006). AVPM, articulatory velocity and position maps. Below: a proposed generative model of speech production that is more consistent with the free energy and predictive coding framework (Friston, 2010).

processing which in turn feeds into a second forward model of sensory (auditory) processing (see first panel in **Figure 1**). This perspective differs from that proposed by Guenther et al. (2006) in which auditory and sensorimotor predictions are generated in parallel with motor commands (rather than in series as in the Tian and Poeppel, 2010 model), see second panel in **Figure 1**. The parallel processing of predictions and motor commands in Guenther et al. (2006) is more consistent with predictive coding accounts in generative models of active inference (Friston, 2010; Friston et al., 2010) where higher-level representations (i.e., prior knowledge of movements and their associations) drive the motor commands and predict the sensory responses in parallel (see third panel in **Figure 1**). However, in the Guenther et al. (2006)model, mismatches between the sensory response and the predicted sensory response (i.e., the prediction errors) are fed back to the motor system. This differs to the predictive coding in generative models (third panel of **Figure 1**) where the prediction errors are fed back to the source of the predictions (i.e., the high level representations) in order to optimize future predictions and minimize future prediction error. In addition, predictions in generative models are propagated in a hierarchical fashion through the system. For example, the third panel of **Figure 1** shows that higher-level representations predict phonological associations of words and phonological processing predicts acoustic associations of words, with potentially many intervening stages that are not illustrated.

Although, the importance of a forward model of speech output during articulation is well recognized (Heinks-Maldonado, 2005; Christoffels et al., 2007; Hawco, 2009), no previous functional imaging study has attempted to identify the anatomical location of brain activation related to the forward model of speech output during articulation. This requires an experimental paradigm that activates speech production but controls for processing related to (a) auditory feedback and (b) movement of the articulators. Instead, previous functional imaging studies that have investigated the self monitoring of speech have primarily focused on activation related to auditory feedback rather than auditory predictions. This has involved altering rather than eliminating the auditory feedback (Paus et al., 1996; Hashimoto and Sakai, 2003; Ford et al., 2005; Fu et al., 2006; Christoffels et al., 2007; Toyomura et al., 2007; Tourville et al., 2008; Takaso et al., 2010). The results have highlighted activation changes in the superior temporal gyri but do not distinguish activation related to predicting speech from activation related to changes in auditory feedback. In contrast to this prior work, our study used a speech task that did not involve the generation of sound or auditory feedback because our aim was to identify brain activation that might be related to the internal model that predicts speech output during articulation.

To isolate brain activation associated with the internal model of speech output, we compared the production of speech to the production of non-speech mouth movements. The key difference between these conditions is that articulation of speech typically results in auditory speech processing whereas the production of non-speech mouth movements is not associated with auditory speech, although there may be some degree of acoustic association. In the speech condition, participants repeatedly articulated the words "one" and "three" without generating any sounds. This task places minimal demands on conceptualization of the intended message, word retrieval, the selection of the appropriate morphological forms, sequencing, respiration control, prosody, and auditory processing of the spoken output. However, silent articulation of words does not eliminate the experience of previously learnt auditory associations that have been tightly coupled with movement in the articulators during speech production (i.e., we have auditory imagery of the words "one" or "three" as they are silently articulated). These auditory images of speech may play a role in predicting the auditory consequences of speech production (Tian and Poeppel, 2010).

The words "one" and "three" were chosen because they have very distinct muscle movements that could be approximately matched in the non-speech mouth movement condition. Articulating "one" primarily involves lip pursing whereas articulating "three" primarily involves tongue protrusion and retraction. In the non-speech mouth movement condition, participants either pursed their lips (in a kissing action), protruded, and retracted their tongue or alternated between these movements. By including three different levels of non-speech mouth movements (lips repeatedly, tongue repeatedly, lips alternating with tongue), we were able to compare activation for different types of articulators (lips versus tongue) and also manipulate the complexity of the movements. For example, we were able to check whether increased activation for speech compared to non-speech was observed in areas where activation increased with the complexity of the movements (i.e., for alternating between different movements compared to repeatedly making the same movement).

Having controlled for auditory feedback and movement of the articulators, we predicted that activation related to the forward/generative model of auditory processing during speech production would be observed in the left ventral premotor cortex and/or the superior temporal gyrus/sulcus. These predictions are made on the basis of prior proposals by Guenther et al. (2006) who link the internal model of speech sound maps to the ventral premotor cortex; and Hickok et al. (2011) who link an internal model of motor processing to the premotor cortex; an internal model of auditory processing to the superior temporal gyrus/sulcus; and the translation between auditory and motor processing to the area they refer to as Spt (in the Sylvian fissure between the planum temporale (PT) and ventral supramarginal gyrus).

In addition to dissociating brain activation for speech and non-speech mouth movements, we also looked for activation that was common to both speech and non-speech mouth movements relative to finger tapping and visual fixation. Previous imaging studies have distinguished different systems involved in the motor control of speech: An articulatory "preparatory loop" that includes the inferior frontal, anterior insula, supplementary motor area, and superior cerebellum; an executive loop including the motor cortex, thalamus, putamen, caudate, and inferior cerebellum (Riecker et al., 2005) and a feedback loop including the postcentral gyri, the supratemporal plane, and the superior temporal gyri (Dhanjal et al., 2008; Peschke et al., 2009). The involvement of these regions in non-speech as well as speech mouth movements has already been demonstrated. For example, Chang et al. (2009) compared speech to non-speech orofacial movements and vocal tract gestures (whistle, cry, sigh, cough) and found common activations in the inferior frontal gyrus, the ventral premotor cortex, the supplementary motor area, the superior temporal gyrus, the insula, the supramarginal gyrus, the cerebellum, and the basal ganglia. This suggests a general role for these regions in orofacial movements and their auditory consequences.

By including a visual fixation baseline, we could also identify activation that was common to both finger and mouth movements; and control for inner speech that occurs independently of mouth movements during free thought.

MATERIALS AND METHODS

Functional imaging data were acquired using positron emission tomography (PET). For the current study of speech production there are two advantages of using PET rather than fMRI: the PET scanning environment is quieter for recording the presence or absence of speech output; and the regional cerebral blood flow (rCBF) signals are not distorted by air flow through the articulators. The study was approved by the local hospital ethics committee.

PARTICIPANTS

We scanned 12 right handed, native English speakers who had normal or corrected vision and hearing and no history of neurological disease or mental illness. All gave written informed consent. One participant was subsequently excluded for reasons given below. The remaining 11 subjects (10 male) had a mean age of 34 years (range 19–68). The predominance of male participants is a consequence of using PET scanning which is not appropriate for women of child bearing age. Our results did not change when the one female was removed ($n = 10$; mean age = 32 years, age range = 19–52) therefore we did not exclude the female participant. Inter-subject variability in our results was investigated and reported (see **Figure 2**) to ensure consistency across participants, despite the wide range of ages and unequal distribution of males and females.

PARADIGM

There were four conditions: silent speech, non-speech mouth movements, finger tapping, and visual fixation. Each condition was repeated in three different blocks (with one block equivalent to one 90 s PET scan). In all 12 scans, a black circle, presented every 750 ms, was used as an external stimulus to pace movement production. During the three speech scans, participants were instructed to articulate the word "one" or "three" in time with the stimulus. They were specifically instructed to move their mouths as if they were speaking but without generating any sound (i.e., silent mouth movements). In one of the three speech scans, they articulated the word "one" on every trial; in a second, they articulated the word "three" on every trial and in the third, they alternated the articulation of "one" and "three," with one speech utterance per stimulus. In the three non-speech mouth movement scans, participants pursed their lips in time with the stimulus, protruded, and retracted their tongue, or alternated between pursing their lips and protruding and retracting their tongue. In the three-finger tapping scans, participants made a two-finger movement in one scan, a three-finger movement in another scan and alternated between the two-finger movement and three-finger movement in the third scan. The two-finger movement involved a tap of their index finger followed by a tap of their middle finger on a table placed under their arm in the scanner. The three-finger movement involved a tap of their index finger followed by a tap of their middle finger followed by a tap of their fourth finger. Participants practiced these movements before the scan and they were referred to as "double drum" and "triple drum" respectively. In the three baseline scans, participants were instructed "Please look at the flashing dot and try to empty your mind."

All responses, during all conditions were video recorded to ensure that the data collected were consistent with the experimental aims (e.g., mouth movements without sound during the speech condition). A scan/condition was repeated if the participants did not follow the instructions correctly. This only happened once for three different participants and in each case the repeated scan replaced the faulty scan. One subject (20-year-old male) did not follow the instructions in two different scans and was therefore excluded from the final analyses ($n = 11$). There was no further behavioral analysis because, in the final data sets, each condition was accurately performed (i.e., error free). Moreover, the functional imaging data showed no activation in the primary auditory cortex during any condition. This is consistent with the participants performing all conditions silently.

DATA ACQUISITION

Functional activation images were acquired using a SIEMENS/CPS ECAT EXACT HR+ (model 962) PET scanner (Siemens/CTI, Knoxville, TN, USA). Each participant had 12 or 13 PET scans (see previous section), to measure rCBF using bolus infusion of radioactively labeled water ($H_2{}^{15}O$). The dose received was 9 mCi per measurement, as approved by the UK Administration of Radioactive Substances Advisory Committee (ARSAC). Using statistical parametric mapping (SPM99), scans from each subject were realigned using the first as a reference, transformed into a standard MNI space (Ashburner and Friston, 1997) and smoothed

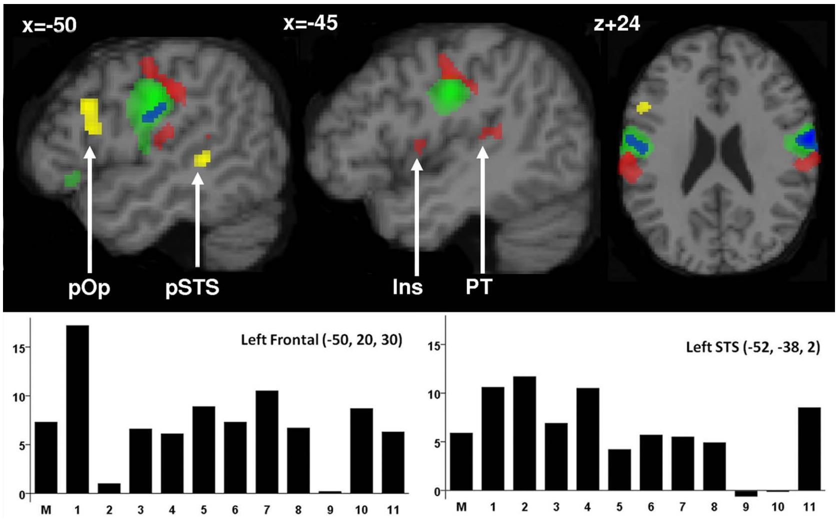


FIGURE 2 | Activation during silent articulation of speech. Top: Activation for speech more than non-speech mouth movements is illustrated in yellow in the pars opercularis (pOp) and the left posterior superior temporal sulcus (pSTS). Activation for speech and non-speech mouth movements relative to finger movements and fixation is illustrated in green. The blue area within this system corresponds to the location where activation was greater for tongue movements relative to lip movements. Activations for all movement tasks (mouth and finger) relative to fixation are illustrated in red. Within the red areas, we have marked activation that was located in the insula (INS) and the left planum temporale (PT). Statistical threshold was set at $p < 0.05$ after FWE correction for multiple comparisons across the whole brain in extent, see **Table 1** for details. Below: Activation for speech relative to non-speech mouth movements (percentage signal change on the *y* axis) in each participant (1–11 on the *x* axis) and the mean (M) at the peak co-ordinates for group activation in the frontal and temporal regions. This illustrates the consistency of the effect in the same voxels.

with a Gaussian kernel of 8 mm FWHM. Structural MRI images for each subject were obtained for coregistration with the PET data.

STATISTICAL ANALYSIS

Statistical analysis used standardized procedures (Friston et al., 1995). This involved ANCOVA with subject effects modeled and global activity included as a subject specific covariate. The condition and subject effects were estimated according to the general linear model at each voxel. The statistical model included 10 conditions: Fixation (summed over three scans), the three-finger tapping conditions, the three non-speech mouth movement conditions and the three speech conditions. The statistical contrasts of interest identified activation that was greater for (1) all speech than all non-speech mouth and finger conditions; (2) all speech than all non-speech mouth movements; (3) all speech and all non-speech mouth movements relative to all finger movements; and (4) all movement conditions relative to fixation; (4) non-speech tongue movements relative to non-speech lip movements or vice versa; and (5) alternating between movements or the same type (e.g., non-speech lip/tongue/lip) versus repetition of the same movement (e.g., non-speech tongue/tongue/tongue). The statistical threshold was set at $p < 0.05$ after family wise error (FWE) correction for multiple comparisons across the whole brain in height or extent. To ensure that activation in contrast (3) reflected common activation for all types of movement, we used the inclusive masking option on SPM to exclude voxels that were not significantly activated (at $p < 0.001$ uncorrected) by (6) speech > fixation, (7) non-speech mouth > fixation, and (8) finger movements > fixation. As the inclusive masking removes voxels from activation maps that are highly significant ($p < 0.05$ corrected), they make the results more conservative rather than less.

RESULTS

GREATER ACTIVATION FOR SILENT SPEECH THAN NON-SPEECH MOUTH MOVEMENTS

There were two areas where activation was significantly higher for silent speech than non-speech mouth movements: the left posterior superior temporal sulcus (pSTS) and the left dorsal pars opercularis within the inferior frontal gyrus extending into the left middle frontal gyrus. In each of these areas, activation was also higher for speech than finger movements and for speech relative to the visual fixation baseline. The loci and significance of these effects are shown in **Table 1** and **Figure 2**.

OTHER EFFECTS

Both speech and non-speech mouth movements resulted in extensive activation in bilateral pre-central gyri relative to finger tapping and visual fixation (see **Table 1** and green areas in **Figure 2** for details). In addition, activation that was common to speech, non-speech mouth movements, and finger tapping (relative to the

Table 1 | Location of activation for speech relative to non-speech mouth movements and finger movement; and for all movement tasks relative to fixation; at peaks that were significant at $p < 0.05$ after correction for multiple comparisons across the whole brain in height ($Z > 4.7$) or extent ($Z > 90$ voxels).

Location of speech activations	Speech > mouth and fingers					Speech > mouth	
	Co-ordinates (*x,y,z* in MNI)			*Z* score	*k*	*Z* score	*k*
Left posterior superior temporal sulcus	−52	−38	2	5.2	103	4.5	110
Left dorsal pars opercularis	−50	20	30	4.9	153	4.3	153
	Speech and mouth > fingers					**Tongue > lips**	
Left pre-central gyrus	−54	6	6	5.8	1658		
	−58	−2	14	7.4			
	−62	−6	26	7.5		4.5	139
	−48	−12	32	8.0			
Right pre-central gyrus	58	−4	10	6.0	1632		
	64	−6	26	8.1		5.8	234
	58	−8	30	8.0			
	Speech, mouth and fingers > fixation						
Left pre/post-central gyrus	−56	−12	16	6.7	1212		
	−58	−8	30	8.0			
	−48	−20	36	7.4			
	−46	−12	42	7.4			
Right post-central gyrus	+64	−14	30	5.6	576		
Left posterior cerebellum	−16	−60	−24	6.9	513		
Right posterior cerebellum	+28	−62	−24	7.7	903		
Left anterior insula	−38	2	+4	5.5	67		
Left planum temporale	−46	−38	+14	5.7	53		

k = number of voxels significant at p < 0.001 uncorrected.

visual fixation baseline) was observed bilaterally in the postcentral gyri, superior cerebellum, inferior cerebellum, putamen, with left lateralized activation in the thalamus, insula, supratemporal plane, and supplementary motor area (see **Table 1** and **Figure 2** which represents a subset of these regions in red). Common activation in these areas may relate to shared processing functions. For example, it has been proposed that activation in the anterior insula is related to the voluntary control of breathing during speech production (Ackermann and Riecker, 2010). It might therefore be the case that all three motor tasks (speech, non-speech mouth movements, and finger tapping) involve voluntary control of breathing in time with the motor activity. Alternatively, common activation might reflect different functions that could not be anatomically distinguished in the current study. As the current study is concerned with differential activation for speech relative to non-speech mouth movements, we do not discuss the common activations further.

The only other significant effect was observed when non-speech tongue movements were compared to non-speech lip movements. These effects are shown in blue in **Figure 2**. The MNI coordinates of this effect ($x = +64$, $y = -6$, $z = 26$; Z score $= 5.9$; and $x = -58$, $y = -6$, $z = 26$; Z score $= 4.1$) correspond to those previously reported for tongue movements (Corfield et al., 1999; Pulvermuller et al., 2006). The consistency of this effect with recent functional imaging (Takai et al., 2010) and early electrocortical mapping (Penfield and Rasmussen, 1950) provides reassuring support that our study had sufficient power to identify effects of interest with high precision. We did not see significantly increased activation for non-speech lip relative to mouth movements; nor did we see differential activation between any of the conditions that alternated between two movements (e.g., lips/mouth/lips) were compared to the corresponding conditions when the same movements was repeated continuously (e.g., lips/lips/lips or mouth/mouth/mouth).

DISCUSSION

Silently articulating the words "one" and "three" strongly activated left inferior frontal and superior temporal language regions compared to lip pursing, tongue movements, finger tapping, and visual fixation. The left inferior frontal activation was located in the left dorsal pars opercularis and therefore corresponds to classic Broca's area. The left superior temporal activation was located in the left pSTS and therefore corresponds to classic Wernicke's area. We suggest that, during speech production, activation in these classic language areas are related to covertly generated auditory associations that are evoked automatically, and in synchrony, with highly familiar mouth movements, previously intimately associated with sound production, and thus auditory feedback. In contrast, lip pursing, tongue, and finger movements are less practiced actions that are not intimately associated with speech sounds although they may have acoustic associations. The location and function of these activations is discussed below, in the context of generative models of perception and active inference (Friston, 2010; Friston et al., 2010). These data lead us to propose that Broca's and Wernicke's areas may play a role in predicting the auditory response during articulation, even in the absence of auditory feedback.

The activation in the dorsal pars opercularis extended anteriorly into the left inferior frontal sulcus (see **Figure 2**). It does not, therefore, correspond to the ventral premotor site of the speech sound maps proposed in the model by Guenther et al. (2006). It is also anterior to the more posterior premotor areas that respond during the observation of hand actions (Caspers et al., 2010), speech perception (Skipper et al., 2007; Callan et al., 2010), mirror neurons (Morin and Grezes, 2008; Kilner et al., 2009), and phonetic encoding during speech production (Papoutsi et al., 2009). Nevertheless, it does correspond to the area that is activated during both inner and overt speech tasks, for example, silent phonological decisions on written words (Poldrack et al., 1999; Devlin et al., 2003), lip reading (Fridriksson et al., 2009; Turner et al., 2009), overt speech production (Jeon et al., 2009; Whitney et al., 2009; Holland et al., 2011), and sentence comprehension (Bilenko et al., 2009; Mashal et al., 2009; Tyler et al., 2010). Moreover, it is not differentially activated by articulating words silently (as in the current study) or saying them aloud (see Price et al., 1996). Therefore the activation is more likely to reflect a fundamental property of speech production than atypical task-specific processing (e.g., the act of inhibiting the production of sounds following instructions to articulate silently). Given the minimal demands on conceptual, lexical, and auditory processing in the current study, we suggest that increased activation in the left dorsal pars opercularis for silently articulating words relative to non-speech mouth movements is related to higher-level representations of learnt words that predict the auditory consequences of well learnt speech articulations. Further we propose that these "predictions" are sent to auditory processing regions in the PT and the pSTS. Confirmation of this hypothesis requires a functional connectivity study with high temporal resolution to determine how activation in the left dorsal pars opercularis interacts with that in the superior temporal gyrus and sulcus.

The left pSTS activation that we observed during the silent articulation of speech is associated with phonological processing of speech sounds (Scott et al., 2000). The same STS area is also activated by written words in the absence of auditory inputs (Booth et al., 2003; Richardson et al., 2011). In addition, Leech et al. (2009) associated the left pSTS with learnt auditory associations. Specifically, they used a video game to train participants to associate novel acoustically complex, artificial non-linguistic sounds to visually presented aliens. After training, viewing aliens alone, with no accompanying sound, activated the left pSTS with activation in this area proportional to how well the auditory categories representing each alien had been learnt. As Leech et al. (2009) point out, part of what makes speech special is the extended experience that we have with it throughout development and this includes acoustic familiarity, enhanced audio–visual associations, and auditory memory in addition to the higher-level processing that is specific to speech (e.g., phonology and semantics). The activation that we observe in left pSTS may therefore reflect auditory associations of the articulated words. This might either be seen as a consequence of auditory predictions from the left dorsal pars opercularis and left pSTS may, in turn, play an active role in generating the predicted acoustic input during articulation (see the generative model in **Figure 1**). As acknowledged above, future functional connectivity studies using data with high temporal resolution will be required to distinguish these alternatives.

We did not find speech-selective activation in the lower bank of the Sylvian fissure that has been referred to as the PT, left supratemporal plane (SPT), or Sylvian parietal temporal junction (Spt). The Sylvian fissure is the sulcus above the superior temporal gyrus but our speech-selective activation was in the pSTS which is the sulcus below the superior temporal gyrus. We did, nevertheless, confirm the involvement of PT/STP/Spt in speech production because we found common PT/STP/Spt activation for speech, mouth movements, and finger movements, relative to fixation. In other words, as shown previously (Binder et al., 2000), PT/STP/Spt was activated by speech but activation in this region was not specific to speech.

The observation of activation in PT/STP/Spt during finger tapping movements is surprising. Traditionally, PT has been considered to be an auditory association area that is important for speech but not more activated for speech than tone stimuli (Binder et al., 2000). More recent proposals suggest that the PT/STP/Spt region is an interface for speech perception and speech production (Wise et al., 2001; Hickok et al., 2009) and involved in anticipating the somatosensory consequences of movements in the articulators (Dhanjal et al., 2008). Our finding that PT/STP/Spt activation is observed for finger tapping and mouth movements might suggest an even more general role in sensorimotor processing. Alternatively, it might be the case that finger tapping and non-speech mouth movements have low level acoustic associations that are predicted during the movements that have previously been associated with such sounds. In other words, we are proposing that, during speech production, auditory predictions are generated at (a) the level of acoustic associations of any type of movement (in PT/STP/Spt) and (b) the phonology associated with learnt words (in pSTS), see lower part of **Figure 1**.

How do our results fit with the models illustrated in **Figure 1**? As emphasized above, the full answer to this question requires techniques with higher temporal resolution that can characterize how all the speech production areas interact and influence one another during articulation. Nevertheless, our data do allow us to test the anatomical hypotheses from the different models. Specifically, the Tian and Poeppel (2010) model suggests that the forward model of auditory processing is in the sensory cortex and the Guenther et al. (2006) model suggests their speech sound maps are in the ventral premotor cortex. In contrast, the effects that we observed for speech processing in the left dorsal pars opercularis and pSTS are in higher-level association areas, not in sensory areas or the ventral premotor cortex. The Spt activation that we observed for speech, non-speech, and finger tapping movements might plausibly correspond to the model proposed by Hickok et al. (2011) in which Spt translates an internal model of motor processing to an internal model of auditory processing. However, the Hickok et al. (2011) model does not provide an interpretation of our speech-selective activation in left dorsal pars opercularis or pSTS. Thus, the anatomical predictions of the previous models do not explain our data. We therefore propose a new anatomical model. Within the framework of the generative model, illustrated in **Figure 1**, we suggest that the activation we observed in the left dorsal pars opercularis corresponds to processing in higher-level areas that predicts the auditory and motor consequences of speech; and the pSTS activation corresponds to phonological processing that may be involved in predicting the auditory response in PT/STP/Spt. Future studies are now required to investigate the validity of this proposal and test how higher-level systems predict inputs to lower-levels; and how prediction error is used to optimize future predictions (Friston, 2010; Friston et al., 2010). We speculate that, during overt speech production, top-down predictions from higher-level areas optimize auditory processing of the heard response by minimizing the prediction error (i.e., the mismatch between the produced and predicted response). In parallel, the prediction error is fed back to the higher-level regions and used to optimize future motor commands and auditory predictions.

In conclusion, we found that regions corresponding to distinct parts of Broca's and Wernicke's areas were activated for mouth movements that have previously been learnt as words and therefore have well established auditory associations. We therefore suggest that the dorsal pars opercularis part of Broca's area and pSTS part of Wernicke's area are involved in predicting the auditory consequences of well rehearsed articulations. In addition, we propose that the left dorsal pars opercularis and pSTS areas may be involved in generating and maintaining a forward generative model of expected speech which can be used as a template for auditory prediction. Mismatches between the auditory predictions and auditory feedback can then be fed to the articulators to improve the precision of subsequent output. These audio–motor interactions are particularly important during speech acquisition in childhood, in those with hearing loss or when adults learn a new language. They are also needed to modify the intensity of speech output in noisy environments and when auditory feedback is altered (e.g., by delay on the telephone). We speculate that the devastating impact of damage to Broca's and Wernicke's areas on speech production may in part be related to the importance of dorsal pars opercularis and pSTS for auditory–motor integration of speech.

ACKNOWLEDGMENTS

This work was funded by the Wellcome Trust. We would like to thank Karl Friston for useful discussions and our radiographers (Amanda Brennan and Janice Glensman) for their help scanning the participants.

REFERENCES

Ackermann, H., and Riecker, A. (2010). The contribution(s) of the insula to speech production: a review of the clinical and functional imaging literature. *Brain Struct. Funct.* 214, 419–433.

Ashburner, J., and Friston, K. (1997). Multimodal image coregistration and partitioning – a unified framework. *Neuroimage* 6, 209–217.

Ballard, D. H., Hinton, G. E., and Sejnowski, T. J. (1983). Parallel visual computation. *Nature* 306, 21–26.

Bilenko, N. Y., Grindrod, C. M., Myers, E. B., and Blumstein, S. E. (2009). Neural correlates of semantic competition during processing of ambiguous words. *J. Cogn. Neurosci.* 21, 960–975.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., and Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–528.

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., and Mesulam, M. M. (2003). Relation between brain activation and lexical performance. *Hum. Brain Mapp.* 19, 155–169.

Borden, G. J. (1979). An interpretation of research of feedback interruption in speech. *Brain Lang.* 7, 307–319.

Callan, D., Callan, A., Gamez, M., Sato, M. A., and Kawato, M. (2010). Premotor cortex mediates perceptual performance. *Neuroimage* 51, 844–858.

Caspers, S., Zilles, K., Laird, A. R., and Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage* 50, 1148–1167.

Chang, S. E., Kenney, M. K., Loucks, T. M., Poletto, C. J., and Ludlow, C. L. (2009). Common neural substrates support speech and non-speech vocal tract gestures. *Neuroimage* 47, 314–325.

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

Corfield, D. R., Murphy, K., Josephs, O., Fink, G. R., Frackowiak, R. S., Guz, A., Adams, L., and Turner, R. (1999). Cortical and subcortical control of tongue movement in humans: a functional neuroimaging study using fMRI. *J. Appl. Physiol.* 86, 1468–1477.

Devlin, J. T., Matthews, P. M., and Rushworth, M. F. (2003). Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *J. Cogn. Neurosci.* 15, 71–84.

Dhanjal, N. S., Handunnetthi, L., Patel, M. C., and Wise, R. J. (2008). Perceptual systems controlling speech production. *J. Neurosci.* 28, 9969–9975.

Ford, J. M., Gray, M., Faustman, W. O., Heinks, T. H., and Mathalon, D. H. (2005). Reduced gamma-band coherence to distorted feedback during speech: when what you say is not what you hear. *Int. J. Psychophysiol.* 57, 143–150.

Fridriksson, J., Moser, D., Ryalls, J., Bonilha, L., Rorden, C., and Baylis, G. (2009). Modulation of frontal lobe speech areas associated with the production and perception of speech movements. *J. Speech Lang. Hear. Res.* 52, 812–819.

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. Biol. Sci.* 360, 815–836.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138.

Friston, K. J. (2001). Dynamic representations and generative models of brain function. *Brain Res. Bull.* 54, 275–285.

Friston, K. J., Daunizeau, J., Kilner, J., and Kiebel, S. J. (2010). Action and behavior: a free-energy formulation. *Biol. Cybern.* 102, 227–260.

Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S., and Turner, R. (1995). Analysis of fMRI time-series revisited. *Neuroimage* 2, 45–53.

Fu, C. H., Vythelingum, G. N., Brammer, M. J., Williams, S. C., Amaro, E. Jr., Andrew, C. M., Yágüez, L., van Haren, N. E., Matsumoto, K., and McGuire, P. K. (2006). An fMRI study of verbal self-monitoring: neural correlates of auditory verbal feedback. *Cereb. Cortex* 16, 969–977.

Gregory, R. L. (1980). Perceptions as hypotheses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 290, 181–197.

Guenther, F. H., Ghosh, S. S., and Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang.* 96, 280–301.

Hashimoto, Y., and Sakai, K. L. (2003). Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: an fMRI study. *Hum. Brain Mapp.* 20, 22–28.

Hawco, C. S. (2009). Control of vocalization at utterance onset and mid-utterance: different mechanisms for different goals. *Brain Res.* 1276, 131–139.

Heinks-Maldonado, T. H. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology* 42, 180–190.

Helmholtz, H. (1866). "Concerning the perceptions in general," in *Treatise on Physiological Optics,* Vol. III, 3rd Edn. (translated by J. P. C. Southall 1925 Opt. Soc. Am. Section 26, reprinted New York: Dover, 1962).

Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69, 407–422.

Hickok, G., Okada, K., and Serences, J. T. (2009). Area Spt in the human planum temporale supports sensory-motor integration for speech processing. *J. Neurophysiol.* 101, 2725–2732.

Holland, R., Leff, A. P., Josephs, O., Galea, J. M., Desikan, M., Price, C. J., Rothwell, J. C., and Crinion, J. (2011). Speech facilitation by left inferior frontal stimulation. *Curr. Biol.* 21, 1–5.

Jeon, H. A., Lee, K. M., Kim, Y. B., and Cho, Z. H. (2009). Neural substrates of semantic relationships: common and distinct left-frontal activities for generation of synonyms vs. antonyms. *Neuroimage* 48, 449–457.

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., and Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *J. Neurosci.* 29, 10153–10159.

Leech, R., Holt, L. L., Devlin, J. T., and Dick, F. (2009). Expertise with artificial nonspeech sounds recruits speech-sensitive cortical regions. *J. Neurosci.* 29, 5234–5239.

Levelt, W. J., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–38; discussion 38–75.

MacKay, D. M. (1956). "The epistemological problem for automata," in *Automata Studies*, eds C. E. Shannon and J. McCarthy (Princeton, NJ: Princeton University Press), 235–251.

Mashal, N., Faust, M., Hendler, T., and Jung-Beeman, M. (2009). An fMRI study of processing novel metaphoric sentences. *Laterality* 14, 30–54.

Miall, R. C. (1993). Is the cerebellum a Smith predictor. *J. Mot. Behav.* 25, 203–216.

Morin, O., and Grezes, J. (2008). What is "mirror" in the premotor cortex? A review. *Neurophysiol. Clin.* 38, 189–195.

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

Paus, T., Perry, D. W., Zatorre, R. J., Worsley, K. J., and Evans, A. C. (1996). Modulation of cerebral blood flow in the human auditory cortex during speech: role of motor-to-sensory discharges. *Eur. J. Neurosci.* 8, 2236–2246.

Penfield, W., and Rasmussen, T. (1950). *The Cerebral Cortex of Man.* New York: The Macmillan Company.

Peschke, C., Ziegler, W., Kappes, J., and Baumgaertner, A. (2009). Auditory-motor integration during fast repetition: the neuronal correlates of shadowing. *Neuroimage* 47, 392–402.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., and Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* 10, 15–35.

Price, C. J., Moore, C. J., and Frackowiak, R. S. (1996). The effect of varying stimulus rate and duration on brain activity during reading. *Neuroimage* 3, 40–52.

Pulvermuller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., and Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7865–7870.

Richardson, F. M., Seghier, M. L., Leff, A. P., Thomas, M. S., and Price, C. J. (2011). Multiple routes from occipital to temporal cortices during reading. *J. Neurosci.* 31, 8239–8247.

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.

Scott, S. K., Blank, C. C., Rosen, S., and Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123, 2400–2406.

Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., and Small, S. L. (2007). Speech-associated gestures, Broca's area, and the human mirror system. *Brain Lang.* 101, 260–277.

Takai, O., Brown, S., and Liotti, M. (2010). Representation of the speech effectors in the human motor cortex: somatotopy or overlap? *Brain Lang.* 113, 39–44.

Takaso, H., Eisner, F., Wise, R. J., and Scott, S. K. (2010). The effect of delayed auditory feedback on activity in the temporal lobe while speaking: a positron emission tomography study. *J. Speech Lang. Hear. Res.* 53, 226–236.

Tian, X., and Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Front. Psychol.* 1:166. doi: 10.3389/fpsyg.2010.00166

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

Toyomura, A., Koyama, S., Miyamaoto, T., Terao, A., Omori, T., Murohashi, H., and Kuriki, S. (2007). Neural correlates of auditory feedback control in human. *Neuroscience* 146, 499–503.

Turner, T. H., Fridriksson, J., Baker, J., Eoute, D. Jr., Bonilha, L., and Rorden, C. (2009). Obligatory Broca's area modulation associated with passive speech perception. *Neuroreport* 20, 492–496.

Tyler, L. K., Shafto, M. A., Randall, B., Wright, P., Marslen-Wilson, W. D., and Stamatakis, E. A. (2010). Preserving syntactic processing across the adult life span: the modulation of the frontotemporal language system in the context of age-related atrophy. *Cereb. Cortex* 20, 352–364.

Whitney, C., Weis, S., Krings, T., Huber, W., Grossman, M., and Kircher, T. (2009). Task-dependent modulations of prefrontal and hippocampal activity during intrinsic word production. *J. Cogn. Neurosci.* 21, 697–712.

Wise, R. J., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K., and Warburton, E. A. (2001). Separate neural subsystems within "Wernicke's area." *Brain* 124, 83–95.

Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882.

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

Received: 02 February 2011; accepted: 30 August 2011; published online: 16 September 2011.

Citation: Price CJ, Crinion JT and MacSweeney M (2011) A generative model of speech production in Broca's and Wernicke's areas. Front. Psychology ***2****:237. doi: 10.3389/fpsyg.2011.00237*

This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 30 September 2011
doi: 10.3389/fpsyg.2011.00253

Motor response selection in overt sentence production: a functional MRI study

Pascale Tremblay*[1]* and Steven L. Small[2]

[1] Center for Mind and Brain Sciences, Università degli Studi di Trento, Trento, Italy
[2] Department of Neurology, University of California Irvine, Irvine, CA, USA

Edited by:
Albert Costa, University Pompeu Fabra, Spain

Reviewed by:
Peter Indefrey, University of Dusseldorf, Germany
Analia Arevalo, East Bay Institute for Research and Education, USA
Denise Klein, McGill University, Canada

***Correspondence:**
Pascale Tremblay, Center for Mind and Brain Sciences, University of Trento, Via delle Regole, 101, I-38060 Mattarello, Trento, Italy.
e-mail: pascale.tremblay@unitn.it

Many different cortical areas are thought to be involved in the process of selecting motor responses, from the inferior frontal gyrus, to the lateral and medial parts of the premotor cortex. The objective of the present study was to examine the neural underpinnings of motor response selection in a set of overt language production tasks. To this aim, we compared a sentence repetition task (externally constrained selection task) with a sentence generation task (volitional selection task) in a group of healthy adults. In general, the results clarify the contribution of the pre-SMA, cingulate areas, PMv, and pars triangularis to the process of selecting motor responses in the context of sentence production, and shed light on the manner in which this network is modulated by selection mode. Further, the present study suggests that response selection in sentence production engages neural resources similar to those engaged in the production of isolated words and oral motor gestures.

Keywords: premotor cortex, supplementary motor area, response competition, sentence generation, spoken language, fMRI

INTRODUCTION

How are our innermost thoughts converted into an articulated verbal message? The neural mechanisms that underlie this fascinating conversion include the selection of words to express an intended meaning, and the selection and sequencing of motor programs to realize them. Motor response selection in the context of spoken language production can be broadly construed as the process by which a set of lexical units forming a message is transformed into a sequence of motor programs; it is a complex process that links cognitive, linguistic, and sensorimotor systems.

Despite the importance of motor response selection, attempts to incorporate this process into contemporary biological models of language remain scarce (but see for example Crosson et al., 2001). Most models of speech and/or language (e.g., Levelt, 1999; Hickok and Poeppel, 2004; Indefrey and Levelt, 2004; Riecker et al., 2005; Guenther et al., 2006) postulate a lexical selection stage, which is a non-motor, language-specific process that can, with some difficulty, be integrated into a broader action execution framework. However, these models postulate that competition for selection occurs only at lexical stage, and thus never incorporate *motor* response selection. Although cascaded models of spoken language production (e.g., Morsella and Miozzo, 2002), do not postulate a motor selection stage *per se*, they do assume that lexical competition spreads to phonological representations, thereby supporting the idea that competition occurs at different levels of representation.

Notwithstanding the lack of a theoretical framework for response selection in spoken language production, several recent studies suggest a role for frontal premotor regions in this process. For example, results of a recent electroencephalographic (EEG) study comparing volitional and externally cued word selection demonstrate modulation of medial frontal activity, suggesting a role for these areas to response selection (Tremblay et al., 2008). Consistent with this finding, several fMRI studies have shown that manipulating response selection during overt or covert single word production modulates large brain networks including the pre-SMA (Brodmann's area 6 m; supplementary motor cortex; SMA), but also the adjacent cingulate motor area (CMA), the inferior fontal gyrus, and the ventral premotor (PM) cortex (Thompson-Schill et al., 1997, 1998; Crosson et al., 2001; Zhang et al., 2004; Alario et al., 2006; Tremblay and Gracco, 2006, 2010; Nagel et al., 2008). One important finding is that the pre-SMA appears to be involved not only in selecting single words (Alario et al., 2006; Tremblay and Gracco, 2006) but also in selecting non-communicative oral motor gestures (Tremblay and Gracco, 2010). Further support for a role for this region is provided by results of a repetitive TMS study (Tremblay and Gracco, 2009), which showed that pre-SMA is *essential* for volitional motor response selection, but not for stimulus-based selection, and that this pattern is similar for selecting words and non-communicative oral motor gestures. Further evidence for a role for pre-SMA in motor response selection was shown by Braun et al. (2001), who found that production of self-organized sequences of lip, jaw, and tongue movements, as well as the production of language, are both associated with activation in pre-SMA. Taken together, these results suggest that the pre-SMA may be playing a central role in selecting motor response during spoken language production.

It could be argued that pre-SMA activation in these studies is related to other linguistic or cognitive processes associated with the production of spoken language. However, there is some evidence to suggest that this is not the case. First, in some of these studies, non-linguistic actions, such as oral gestures (Braun et al., 2001; Tremblay and Gracco, 2010) and hand actions (Tremblay et al., 2008) were compared to word production tasks and similar patterns of neural activity were found across domains (linguistic, non-linguistic). Furthermore, in the realm of motor control *per se*, several neuroimaging studies have examined the process of selecting motor responses and shown that the magnitude of activation in pre-SMA increases commensurate with demands on response selection. For instance, activation in pre-SMA is enhanced when participants are free to choose a motor response from among several alternatives (i.e., "volitional" selection) compared to when they are required to execute a specific, stimulus-driven, motor response (e.g., Deiber et al., 1996; Van Oostende et al., 1997; Hadland et al., 2001; Ullsperger and Von Cramon, 2001; Weeks et al., 2001; Lau et al., 2004, 2006). Despite along-standing tendency to conceptualize language as "unique" or "special," that is, as being independent from other behaviors, it is becoming increasingly accepted that language relies on largely distributed (that is, presumably non-language-specific) neural networks, though the degree and nature of the overlap between language and other functional systems needs to be further characterized. At the behavioral level, several experiments have demonstrated a connection between speech and hand gestures (Gentilucci et al., 2001; Gentilucci, 2003), and between language and oral motor gestures (Alcock et al., 2000; Alcock, 2006). In this context, the finding of similar neural circuits engaged in motor response selection across domains is not surprising.

Taken together, these findings are consistent, at least in part, with a hypothesis that is referred to as the "medio-lateral gradient of control" hypothesis, according to which the more an action requires internal (volitional) control, the more the involvement of medial premotor areas (which corresponds to the medial portion of Brodmann area 6). In contrast, externally (stimulus) driven actions tend to rely on lateral (rather than medial) premotor areas (Goldberg, 1985). Traditionally, the medial portion of Brodmann area 6 was considered to be a single area, the supplementary motor area (Penfield and Welch, 1951; Woolsey et al., 1952). However, it is now widely accepted that this large cortical area divides into at least two distinct areas approximately at the level of the anterior commissure (see for example Rizzolatti et al., 1998; and Luppino and Rizzolatti, 2000, for reviews), with the SMA-proper forming the caudal part of the region, posterior to the VAC line, and the pre-SMA forming the anterior part. The pre-SMA has a connectivity pattern that is ideal for linking cognitive and motor processes, a *sine qua non-for* the implementation of motor response selection, with important projections from the prefrontal cortex, particularly the dorsolateral prefrontal cortex (Luppino et al., 1993; Lu et al., 1994; Wang et al., 2005), and connections with several premotor areas such as the SMA-proper and the lateral PM (Luppino and Rizzolatti, 2000), for controlling motor output. In addition to the pre-SMA, the lateral premotor cortex has also been discussed in the context of response selection, particularly in relation to stimulus-based hand movement selection (Goldberg, 1985; Mushiake et al., 1991; Deiber et al., 1996; Dirnberger et al., 1998), though evidence of distinct pathways for volitional and stimulus-based selection remains scarce.

In sum, a review of the current literature suggests an important contribution of the pre-SMA, along with potential contribution of the adjacent CMA, the inferior frontal gyrus (IFG), and the ventral PM, in selecting motor programs for single words, single oral non-communicative gestures, and finger movements. One important question that follows from these findings is whether the pattern of results in isolated single word processing bears any resemblance to the pattern associated with production of phrases, sentences, and discourse that characterize naturalistic spoken language. Given a heavy reliance on selection, and the accelerated pace at which selection occurs – considering that adult speakers may produce as many as 14 phonemes per second, i.e., up to six to nine syllables per second (e.g., Kent, 2000) – it is reasonable to ask whether selection in this setting relies on the same neural mechanisms as in isolated single word production. The objective of the present study was to test the generalizability of previous results by examining the neural underpinnings of motor response selection in a set of sentence production tasks. To this aim, we compared a sentence repetition task with a sentence generation task in a group of 21 healthy adults. Based on the literature, we predicted a stronger involvement of pre-SMA and possibly ventral PM (PMv) in sentence generation than sentence repetition, reflecting the increased requirements for selection during generation. We also expected regions involved in response selection to be active in both production modes, as both require selection.

MATERIALS AND METHODS

PARTICIPANTS

Twenty-one healthy right-handed (Oldfield, 1971) native speakers of English (mean 25 ± 4.4; 10 males), with a mean of 15.4 years of education participated in the fMRI experiment. All participants had normal hearing sensitivity, as measured by normal pure-tone thresholds and normal speech recognition scores (92.3% accuracy on the Northwestern University auditory test number 6). The Institutional Review Board for the Division of Biological Sciences at The University of Chicago approved the study.

BEHAVIORAL TASKS

To evaluate spontaneous production of words under restricted search conditions, a category fluency task was administered to participants prior to the fMRI session. Participants were instructed to produce as many animal and vegetable words as possible in 1 min (in two separate trials). To examine participants' verbal comprehension skills, an auditory memory span task was administered to participants (an auditory version of the reading span task developed by Daneman and Carpenter, 1980). Participants' responses were recorded and stored to disk for offline analysis. A research assistant naive to the purpose of the study transcribed all the responses.

EXPERIMENTAL PROCEDURES

Participants underwent five different tasks while in the scanner (1) passive observation of object pictures, (2) passive sentence listening, (3) listening and repeating sentences, (4) generating sentences from object pictures, and (5) passive observation of short action movies. The comparison of the language tasks and the non-language tasks has been reported elsewhere (Tremblay and Small, 2011). Each condition was acquired in separate runs, and alternated with "rest" epochs during which the participants were asked to relax. For each condition, the order of the conditions and number of rest trials was optimized using

OPTseq2 (http://surfer.nmr.mgh.harvard.edu/optseq/). Stimuli were presented using Presentation Software (Neurobehavioral Systems).

The tasks of interest for this study were the two sentence production tasks (sentence repetition and sentence generation). During sentence repetition, participants heard a set of 80 sentences (40 action, 40 object sentences) interleaved with 30 rest trials; their task was to repeat the sentence. Both stimulus presentation and response occurred while the gradients were switched off for a 4.5-s of silence ("sparse sampling"). At the beginning of the silent interval, a Go cue was presented, instructing participants to start repeating the sentence. Participants' responses were recorded and stored to disk for offline analysis. In sentence generation, participants were asked to generate 80 sentences (40 action, 40 object) from a set of 40 object pictures interleaved with 28 rest trials. The pictures were simple black-and-white line drawings representing common man-made objects selected from the International Picture Norming Project corpus from the Center for Research in Language at the University of California San Diego (Bates et al., 2003; Szekely et al., 2003). In each experimental trial, a picture was presented for 1 s and was followed, after 500 ms, by a visual cue ("go") instructing participants to start generating the sentence. As noted, all speaking occurred while the MR gradients were switched off.

In addition to these two sentence production tasks, we included two passive tasks, sentence listening and picture observation, as controls for sentence repetition and sentence generation, respectively. During sentence listening 80 short sentences (0.9–1.3 s) interleaved with 30 rest trials were presented to participants. Half of these sentences described manual object-directed actions and the other half described visual properties of the same set of objects. The sentence stimuli were presented while the gradients were switched off which ensured ease of auditory processing for participants. During picture observation, a set of 40 simple black-and-white line drawings was presented one per trial for 1 s and interleaved with 37 rest trials (crosshair fixation). Participants were instructed simply to attend to the pictures.

IMAGE ACQUISITION AND ANALYSIS

Image acquisition

The data were acquired on a 3 T General Electric (Milwaukee, WI) Signa HDx imager with EXCITE. Participants wore MR compatible headphones and goggles (NordicNeuroLab Audio/Visual system). 34 axial slices ($3.125\,\text{mm} \times 3.125\,\text{mm} \times 3.6\,\text{mm}$, no gap, FOV = $256\,\text{mm} \times 256\,\text{mm}$, matrix = 64×64) were acquired in 1.5 s using a multislice EPI sequence with parallel imaging (ASSET = 2; TE = 26 ms; FOV = 20 cm; 64×64 matrix; Flip angle: 73). To eliminate movement artifacts associated with speaking, and to ensure that participants could hear the auditory stimuli, a sparse image acquisition technique was used during all the language tasks. A silent period (1.5 s for listening, 4.5 s for repetition and generation) was interleaved between each volume acquisition. Trials containing errors[1] (corresponding to 1.2% of the trials in sentence repetition and 13.5% in sentence generation) were excluded from the analysis of the behavioral and fMRI data. High-resolution T1-weighted volumes were acquired for anatomical localization.

Timeseries analyses

The timeseries were spatially registered, motion-corrected (within and across runs), de-spiked and converted to percentage of signal change using AFNI (Cox, 1996). A linear least squares model was used to establish a fit to each time point of the hemodynamic response function for each of these conditions. There were separate regressors for each of the experimental conditions. Additional regressors were the mean, linear, and quadratic trend components, as well as the six motion parameters (x, y, z, roll, pitch, yaw). We modeled the entire trial duration (i.e., 6 s), which included stimulus presentation and speech production. Event-related signals were calculated by linear interpolation, beginning at stimulus onset, and continuing for 12 s, using AFNI's tent function (i.e., a piecewise linear spline model). The fit was examined at two different time lags (0–6 s, and 6–12 s) to identify the time point showing the strongest hemodynamic response in our regions of interest (ROI). All subsequent analyses focused on the beta values from the first 6 s post-stimulus onset time lag.

Participants' anatomical scan was aligned to the registered EPI timeseries (Saad et al., 2009). FreeSurfer (Dale et al., 1999; Fischl et al., 1999) was used to create surface representations of each participant's anatomy. Once these surfaces were created, they were exported into SUMA (Saad et al., 2004), which was used to project the functional data resulting from the first-level analysis onto two-dimensional surfaces. Prior to running the group analyses, we applied a 6-mm smoothing kernel to increase the signal-to-noise ratio. Smoothing data on the surface instead of the volume ensures that smoothing avoids inclusion of white matter data, and it prevents averaging data across sulci and gyri (Argall et al., 2006). The group analyses were performed using SUMA on the smoothed beta values. First, we examined the main effect of each condition (repetition, generation) compared to their respective baselines (sentence listening, picture observation). We then examined the difference between sentence generation and sentence repetition. These standard subtraction-type analyses were complemented by a "conjunction" analysis (Nichols et al., 2005) to uncover brain regions commonly active across the speaking tasks. In particular, we identified a task-independent speech production network by computing the intersection (or *conjunction*) of brain activity for repetition ∩ generation. The conjunction analysis only includes regions that survived correction for multiple comparisons in both repetition and generation. For each analysis, a permutation approach (Nichols and Holmes, 2002) was used to identify significant clusters of activated vertices, with an individual vertex threshold of $p < 0.005$, corrected for multiple comparisons to achieve a family-wise error (FWE) rate of $p < 0.05$ (clusters ≥ 168 vertices).

Anatomical region of interest analysis

In addition to the whole brain analyses, further analyses were conducted on two sets (frontal lateral and fronto-medial) of anatomical ROI selected *a priori*. The lateral ROIs included the rostral and caudal portions of PMv (rostral PMv: precentral sulcus; caudal PMv: precentral gyrus), and the pars opercularis and pars

---

[1] Errors included misses, as well as incomplete and/or ungrammatical sentences, and, in sentence repetition, inaccurate repetition.

triangularis of the IFG. The medial ROIs included the pre-SMA and SMA-proper, as well as the rostral and caudal parts of the cingulate gyrus. Each of the ROIs was identified on the individual's cortical surface representation using an automated parcellation scheme as implemented in FreeSurfer (Fischl et al., 2002, 2004; Desikan et al., 2006). This procedure uses a probabilistic labeling algorithm that incorporates the anatomical conventions of Duvernoy (1991), and thus is based on macroanatomical landmarks, not on cytoarchitectonic maps, and therefore represents an approximation to the actual motor and premotor areas. Such anatomical approach is very robust as it takes into account individual participant's anatomy; moreover, it avoids the common problem of selection bias in fMRI research, whereby only those voxels exhibiting a particular pattern are chosen for further analyses (for a discussion of this issue, see for example Vul and Kanwisher, 2009): here, all the voxels in each pre-determined region is selected for analysis.

The ROIs were defined as follows: (1) Rostral PMv: this region was operationalized as the ventral part of the precentral sulcus, defined as the part of the sulcus below the junction of the inferior frontal sulcus with the precentral sulcus. The resulting rostral PMv was bounded rostrally by pars opercularis, caudally by the precentral gyrus, and dorsally by the dorsal PM. (2) Caudal PMv: this region was defined as the part of the precentral gyrus below the junction of the inferior frontal sulcus with the precentral gyrus. The resulting caudal PMv was bounded rostrally by the rostral PMv, caudally by the central sulcus, and dorsally by the dorsal PM. (3) Pars triangularis was defined as the gyrus immediately anterior to pars opercularis; bounded caudally by pars opercularis, and rostrally by pars orbitalis, not including the inferior frontal sulcus. (4) Pars opercularis was defined as the part of the IFG immediately anterior to the precentral gyrus, bounded caudally by the precentral sulcus, and rostrally by pars triangularis, and not including the inferior frontal sulcus. (5) Pre-SMA was defined as the portion of the medial superior frontal gyrus that is anterior to the VAC line, which is a (virtual) vertical line passing through the anterior commissure, and posterior to a virtual line passing through the genu of the corpus callosum. The ventral boundary of the pre-SMA is the cingulate sulcus. (6) SMA-proper was defined as the portion of the medial superior frontal gyrus posterior to the VAC line, and anterior to the medial precentral gyrus. (7) The rostral cingulate region was defined as the part of the cingulate gyrus anterior to the VAC line, and posterior to a virtual line passing through the genu of the corpus callosum. (8) The caudal cingulate was defined as the portion of the cingulate gyrus posterior to the VAC line, and anterior to the medial precentral gyrus.

The mean percentage of BOLD signal change was extracted for each ROI and each condition. We then calculated two difference scores to isolate the effects specific to producing language, over and above perception of the stimuli: (1) repetition (sentence repetition – sentence listening), and (2) generation (sentence generation – picture observation). These scores were entered in a three-way ANOVA with repeated measurement on Task (Repeat, Generate), Hemisphere (Left, Right), and ROI. We conducted this analysis separately for each ROI group (lateral, medial). We used FDR corrected two-tailed comparisons to examine whether the activation magnitude in each ROI was significantly different from zero (positively or negatively) for repetition and generation. When a region showed significant activation in either of the tasks, we also performed an FDR corrected two-tailed pairwise comparison to examine a potential task effect.

In addition to these analyses, we also examined the relation between regional activation and behavior. Specifically, we correlated the mean activation in each ROI during sentence generation and a set of five behavioral measures: (i) accuracy during the sentence generation task (percentage of correct sentences produced); (ii) number of words produced; (iii) number of syllables generated; (iv) category fluency score (total number of words produced for animal and vegetable fluency combined); and (v) verbal working memory score (reading span; total word recalled per participant). We postulated that these last two measures would be highly related to performance on the sentence generation task, because, like the sentence generation task, they involve word search and response selection. Using partial correlations (with participants as a covariate of no interest), we investigated potential linear relationships between the magnitude of brain signal in each of our ROIs and these measures.

ONLINE BEHAVIORAL DATA ANALYSES

Participants' responses during the fMRI session were recorded online using LabVIEW (National Instruments, Austin, TX, USA) and stored to disk. The responses for two participants could not be stored due to technical difficulty. A research assistant naive to the purpose of the study transcribed the responses for the 19 remaining participants. For each sentence, we verified accuracy (whether or not it conformed to task instructions) and grammaticality (whether the sentence was correctly formed). In addition, we calculated the number of syllables and words for each sentence. Finally, we calculated the number of departures from the primed sentence structure.

Trials containing errors were removed from the analysis of the behavioral and fMRI data.

RESULTS

ONLINE BEHAVIORAL DATA

Complete details on the analysis of the behavioral data have been reported elsewhere (Tremblay and Small, 2011). Of particular importance, the sentence repetition and sentence generation tasks did not differ from each another on any of the online measures (number of words, number of syllables, accuracy).

Moreover, as was expected, without having been instructed to do so, participants spontaneously imitated the structure they had been exposed to (primed) during the sentence generation and sentence listening tasks, as anticipated based on known "structural persistence" in sentence production (Bock, 1986). The primed sentence structures were simple sentences containing a subject and a predicate. Half the sentences consisted of a noun subject and a simple predicate such as "The drawer is open" or "The scissors are sharp" (the object-related sentences). The other half of the sentences consisted of the first person pronoun ("I") followed by a predicate, such as "I drag the suitcase" (the action-related sentences). All action sentences used the present tense. Results show that participants employed the primed sentence structures in the majority of the trials, with "novel" sentence structures occurring in only 156 (of 1200 total) trials, representing fewer than 13% of all uttered sentences. Most of these novel structures were simple modifications of the primed structure, such as a change from the present to the past tense (representing 49% of all novel structures), deletion of the pronoun (representing 9% of all novel structures), or deletion of the determiner (representing 8% of all novel structures). The details of the deviations from the primed syntactic structure are reported in **Table 1**.

NEUROIMAGING DATA

Whole brain analyses

Figure 1 reveals the brain areas jointly activated for sentence repetition and sentence generation, after removal of baseline activation (sentence listening and picture observation, respectively). These areas included the precentral gyrus and central sulcus bilaterally, as well as the transverse temporal gyrus and sulcus bilaterally. An exhaustive list of all regions is presented in **Table 2**.

Figure 2A shows task-related activation during sentence repetition, after removing the effect of sentence listening. As can be seen in the Figure, activation was largely bilateral and included clusters of activated nodes along the precentral gyrus and central sulcus covering both the ventral primary motor cortex and the PMv, as well as clusters of activation in the medial frontal area, the bilateral transverse temporal gyrus, and the planum temporale bilaterally. **Figure 2B** shows task-related activation during sentence generation, after removing the effect of picture observation. Activation was distributed across a large network of bilateral brain areas, including primary and secondary visual areas, the precentral gyrus and central sulcus covering both the ventral primary motor cortex and the ventral premotor cortex, in the medial frontal area, in the bilateral transverse temporal gyrus and bilateral planum temporale, and the left IFG. Compared to sentence repetition, in which activation was equally distributed across both hemispheres, activation in sentence generation was stronger on the left than on the right hemisphere. An exhaustive list of all task-related activation for the basic contrasts (repetition – listening and generation – picture observation) is presented in **Table 3**. Direct comparison of the repetition and generation tasks is shown in **Figure 3**. This contrast revealed activation in the left pre-SMA, as well as activation in the left IFG and in the primary visual cortex bilaterally. These results are detailed in **Table 4**.

ROI analyses

Medial ROIs. The three-way omnibus ANOVA (task, hemisphere, ROI) for the medial regions (pre-SMA, SMA-proper, rostral and caudal cingulate gyrus) revealed a significant main effect of ROI [$F_{(3,60)} = 13.19$, $p = 0.000001$], as well as a significant main effect of hemisphere [$F_{(1,19)} = 14.45$, $p = 0.001$], but no main effect of task (Repeat, Generate). There were several significant two-way interactions: ROI by hemisphere [$F_{(3,60)} = 3.66$, $p = 0.018$], ROI by task [$F_{(3,60)} = 5.948$, $p = 0.001$], and hemisphere by task [$F_{(1,19)} = 18.883$, $p = 0.000348$]. The three-way interaction just failed to reach significance [$F_{(1,19)} = 2.434$, $p = 0.074$]. Interestingly, only the pre-SMA exhibited an overall positive pattern of activation; the SMA-proper, rostral and caudal cingulate regions exhibited overall activity that was either significantly decreased from baseline or that was not different from zero.

To further examine these results, we tested the activation level in each of the ROIs against zero using a set of FDR corrected pairwise comparisons. These comparisons revealed that overall the left pre-SMA was significantly more active than all other medial regions. Activations in the caudal cingulate gyri and SMA-proper bilaterally were not significantly different from zero in either production task. Activation in the rostral cingulate gyrus was lower than zero (relative deactivation) for the generation task, in both the left [$t_{(20)} = -2.11$, $p = 0.048$] and the right hemisphere [$t_{(20)} = -2.12$, $p = 0.047$], though it did not survive an FDR correction. In the left pre-SMA, activation was significantly greater than zero for sentence repetition [$t_{(20)} = 3.36$, $p = 0.003$] and sentence generation [$t_{(20)} = 2.74$, $p = 0.012$]. In the right pre-SMA, activation was greater than zero only for sentence repetition [$t_{(20)} = 2.55$, $p = 0.02$], but not for sentence generation [$t_{(20)} = 1.26$, $p = 0.22$]. The left pre-SMA was the only medial region for which sentence generation was

Table 1 | Departures from primed sentence structure.

Type of structural change	Sentence type	Total number of occurrences	Mean number of occurrences per subject	Percentage of all novel structures
Verbchoice[a]	Object	43	2.69	0.28
Object[b]	Object	10	0.63	0.06
Determiner deletion[c]	Object	12	0.75	0.08
Pronoun deletion[d]	Action	14	0.88	0.09
Verb tense[e]	Action	77	4.81	0.49
Total	N/A	156	1.95	1.00

a = Verb choicerefers to a change in verb used in the object sentences. The primed verbs are "is" and "are" ("The book is think"). A common departure to this pattern is to use the verb has instead of the verb is, as in "The paper has lines."

b = Object refers to a change in the focus of the sentence, from describing an attribute of an object to naming an object, such as in " There's a radio" (Subject5). c = Determiner deletion refers to the dropping of the determiner in the object sentences, such as in "Tweezers are small" (Subject 17). d = Pronoun deletion refers to dropping the pronoun "I" in the action sentences such as in "Open the box" (Subject 9). e = Verb tense refers to switching from the present to the past tense in the action sentences such as in "I measured this piece of paper" (Subject 14).

associated with significant stronger activation than sentence repetition [$t_{(20)} = 2.54$, $p = 0.02$]. These results are illustrated in **Figure 4**.

Lateral ROIs. For the lateral ROIs (pars opercularis and triangularis, rostral and caudal PMv), the three-way omnibus ANOVA (task, hemisphere, ROI) revealed a significant main effect of ROI [$F_{(3,60)} = 3.19$, $p = 0.03$], as well as a significant main effect of hemisphere [$F_{(1,20)} = 19.28$, $p \leq 0.001$], but no main effect of task. The ROI by hemisphere two-way interaction was significant [$F_{(3,60)} = 3.053$, $p = 0.035$], as were the hemisphere by task two-way interaction [$F_{(1,20)} = 15.82$, $p = 0.001$] and the three-way ROI by task by hemisphere interaction [$F_{(3,60)} = 2.97$, $p = 0.039$]. In contrast to the medial ROIs, none of the lateral ROIs exhibited relative deactivation. To further examine the activation patterns in the lateral frontal ROIs, we tested the activation level in each of the ROIs against zero using a set of FDR corrected pairwise comparisons. These analyses revealed strong activation in the left pars opercularis for both repetition [$t_{(20)} = 4.936$, $p = 0.0001$] and generation [$t_{(20)} = 3.55$, $p = 0.002$]; activation in the right pars opercularis was not significant in either condition. Activation in the left pars triangularis was only significantly different from zero during generation [$t_{(20)} = 3.67$, $p = 0.0015$], and in the right pars triangularis, it was not significant in either condition. The rostral PMv was significantly active for repetition [$t_{(20)} = 3.211$, $p = 0.004$] and generation [$t_{(20)} = 5.14$, $p = 0.00005$] on the left but not the right hemisphere, while the caudal PMv showed bilaterally significant activation for both repetition [left: $t_{(20)} = 6.16$, $p = 0.00001$; right: $t_{(20)} = 5.67$, $p = 0.00002$] and generation [left: $t_{(20)} = 5.59$, $p = 0.00002$; right: $t_{(20)} = 3.18$, $p = 0.005$]. There was a tendency for all the left lateral ROIs to show a task effect (generation > repetition), but this effect only survived an FDR correction in the left rostral PMv [$t_{(20)} = 4.74$, $p = 0.0001$] and in the left caudal PMv [$t_{(20)} = 3.67$, $p = 0.002$]. None of the right lateral ROIs showed a task effect. These results are illustrated on **Figure 5**.

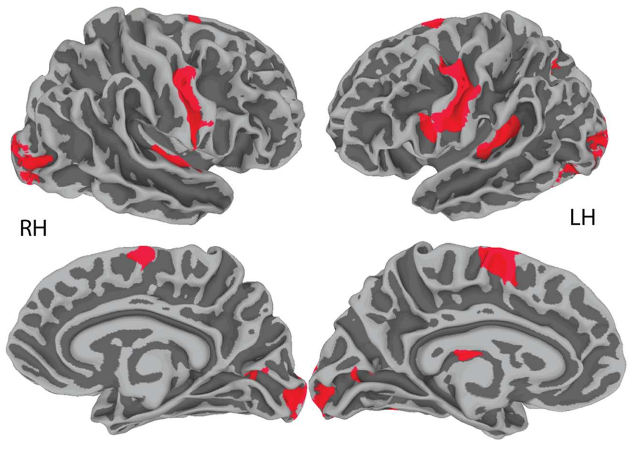


FIGURE 1 | Group-level ($N = 20$) conjunction map (sentence repetition ∩ sentence generation). FWE-corrected joint activation is shown on the group average smoothed white matter folded surface.

Brain–behavior correlations

In addition to examining the activation patterns in the ROIs, we also examined the relationship between activation magnitude during the sentence generation task and a set of five behavioral measures (accuracy during the sentence generation task, number of words produced, number of syllables produced, category

Table 2 | Family-wise error -corrected group-level ($N = 20$), cortical surface results for intersection of sentence generation and sentence repetition.

Anatomical description	Hemi	x	y	z	Cluster size in nodes
Ventral precentral sulcus, extending into the precentral gyrus, central sulcus, and postcentral gyrus	Left	47	15	23	4818
Caudal calcarine fissure, extending laterally into the occipital pole and inferior occipital gyrus		−7	−92	5	1403
Transverse temporal gyrus and sulcus, extending caudally into the planum temporale		−43	−41	20	1369
Medial frontal gyrus (pre-SMA, SMA-proper)		−11	−4	69	1301
Body of the calcarine sulcus		−12	−68	2	343
Inferior temporal sulcus		−43	−77	−5	462
Intra-occipital sulcus		−29	−60	45	357
Ventral central sulcus, extending rostrally into the precentral gyrus, precentral sulcus, and caudally into the postcentral gyrus	Right	63	−7	27	3218
Caudal calcarine fissure, extending laterally into the occipital pole and inferior occipital gyrus		12	−101	12	1349
Body of the calcarine sulcus		4	−76	16	303
Transverse temporal gyrus and sulcus		49	−25	7	468
Medial frontal gyrus (pre-SMA, SMA-proper)		8	14	58	463
Caudal fusiform gyrus and inferior occipital sulcus		17	−94	−10	253
Inferior occipital gyrus		30	−95	−10	341

All coordinates are in Talairach space and represent the centroid surface node for each of the cluster (minimum cluster size: 168 contiguous surface nodes, each significant at $p < 0.005$).

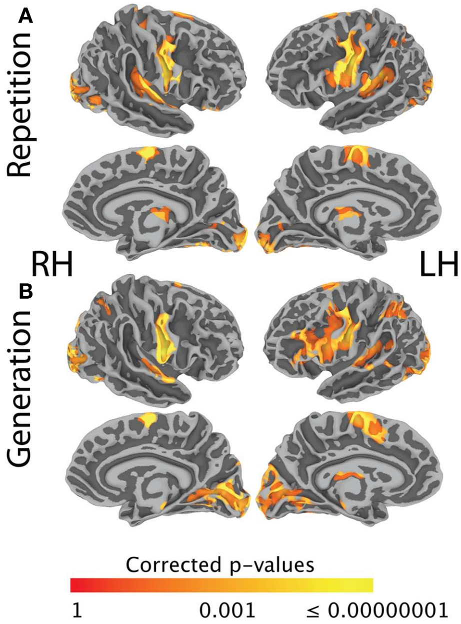


FIGURE 2 | (A) Shows FWE-corrected group-level (*N* = 20) task-related activation during sentence repetition, after removing the effect of sentence listening. **(B)** Shows FWE-corrected group-level (*N* = 20) task-related activation during sentence generation, after removing the effect of picture observation. Activation is shown on the group average smoothed white matter folded surface.

fluency, and verbal working memory). In the animal fluency task, participants generated an average of 25.3 (±6.09 SD; range: 15–37) words. In the vegetable fluency task, they generated on average 14.4 (±14.4 SD; range: 8–23) words. We used the total number of words generated as our measure of fluency. In the auditory span task, participants were able to recall a mean of 53/100 words (±10.4 SD; range: 33–67). The average number of words produced in the sentence generation task was 4.49 (±0.55 SD; range: 4–7); the average number of syllables was 5.62 (±0.65 SD; range: 4–8). The results of the correlation analyses are detailed in **Table 5**. Participants' verbal working memory, as measured by the auditory span task, did not correlate with activation during sentence generation in any of the ROIs. One interesting finding is that activation in the left or right pre-SMA did not correlate with any of the online or offline language measures. In PMv (rostral and caudal) and IFG (pars triangularis and opercularis), activation was negatively correlated with the number of words produced; that is, the more words produced, the less activation was found in these regions.

DISCUSSION

The objective of the present study was to test the generalizability of previous results related to the neural basis of motor response selection by examining the neural underpinnings of this process during a sentence production task, focusing on premotor areas of the cerebral cortex. As discussed in the Introduction, previous studies of hand and finger response selection suggest the existence of a response buffer in which candidate motor programs are co-activated and compete for selection during response planning (e.g., Deiber et al., 1996; Van Oostende et al., 1997; Hadland et al., 2001; Ullsperger and Von Cramon, 2001; Weeks et al., 2001; Lau et al., 2004, 2006). In addition, previous imaging studies (Braun et al., 2001; Tremblay et al., 2008; Tremblay and Gracco, 2010) provide some evidence that this motor response selection mechanism may also be involved during speech production. In the current study we wanted to examine whether such a mechanism could play a role in the production of connected speech. Indeed, most of the research reported in the literature focuses on single word production. However, it is unclear if single word production is an adequate proxy for more complex forms of language, which involve the production of connected speech. To address the question of response selection in a more natural production context, we compared sentence repetition with sentence generation in a group of healthy adults. Sentence generation requires selection of a set of words to express meaning, and the selection of motor programs to realize them, and thus relies heavily on response selection mechanisms; sentence repetition, in contrast, relies less heavily on selection because it involves producing a set of pre-defined words.

While sentence generation, in addition to requiring semantic processing, also requires syntactic processing, the demands on the syntactic system are limited in our study by the fact that participants had just listened to over 150 sentences with similar syntactic structure prior to sentence generation. We used this design to take advantage of structural persistence (Bock, 1986, 1990, see Pickering and Branigan, 1999, for a review), the priming phenomenon in which people tend to use syntactic constructions they have most recently encountered. Indeed, this part of our design was successful: the sentences participants generated were largely identical to those they had heard, thus controlling for the syntactic complexity of the repetition and generation tasks. Hence, while both sentence generation and sentence repetition required selection of motor programs, the generation task included a competition/selection component minimized during sentence repetition. Our hypothesis was that competing words are associated with competing motor programs. Thus, in this context, we expected regions involved in motor response selection to be modulated (generation > repetition), but, critically, we also expected such regions to be active in both sentence production tasks since both require selection and sequencing of motor programs. Based on the literature, we expected to find such pattern in the pre-SMA and possibly ventral premotor cortex (PMv).

TASK-RELATED ACTIVATION AND DEACTIVATION IN MEDIAL CORTICAL AREAS

Our findings demonstrate that a region of the left medial wall, the pre-SMA, was active in both sentence repetition and sentence generation, and showed a unique and significant task-related modulation, suggesting a role in response selection at the sentence level, and henceforth extending previous results at the single word level. Interestingly, this effect was restricted to the left pre-SMA and did not extend into the right pre-SMA, suggesting a degree of functional specialization of the left pre-SMA. This pattern of activation is consistent with previous reports of a selection mode effect in the left but not the right pre-SMA (Tremblay and Gracco, 2010). It is also consistent with results of a study in which participants were required to generate sentences aloud from incomplete stimuli

Table 3 | Family-wise error-corrected group-level ($N = 20$), cortical surface results for the contrast of (A) Sentence Repetition against Sentence listening, and (B) Sentence generation against picture observation.

Anatomical description	Hemi	Max *t*-value	*p*	*x*	*y*	*z*	num
A. SENTENCE REPETITION (SENTENCE LISTENING REMOVED)							
Ventral central sulcus, extending rostrally into the precentral gyrus, precentral sulcus, and pars orbitalis of the inferior frontal gyrus. The cluster also covers the ventral postcentral gyrus	Left	10.986	0.000000001	−42	−8	32	6649
Caudal calcarine sulcus, lingual gyrus, caudal fusiform gyrus, extending laterally into the superior, middle, and inferior occipital gyri		8.116	0.0000001	−16	−94	−7	3770
Transverse temporal gyrus and sulcus, extending caudally into the planum temporale		8.668	0.00000003	−38	−37	20	3725
Medial frontal gyrus (pre-SMA, SMA-proper).		6.745	0.0000015	−5	2	65	1646
Dorsal central sulcus		5.467	0.0000237	−14	−31	60	661
Body of the calcarine sulcus		4.472	0.0002335	−12	−72	10	343
Superior parietal lobule extending downward into the intra-occipital sulcus		4.648	0.0001551	−23	−60	48	562
Intraparietal sulcus		3.836	0.0010321	−27	−53	45	176
Caudal calcarine sulcus, caudal fusiform gyrus, extending laterally into the superior, middle, and inferior occipital gyri	Right	8.781	0.00000003	17	−98	−3	4686
Ventral central sulcus, extending rostrally into the precentral gyrus, precentral sulcus, and caudally into the ventral postcentral gyrus		10.281	0.000000002	60	−7	22	4428
Transverse temporal gyrus and sulcus, extending caudally into the planum temporale		5.728	0.0000132	43	−23	9	3962
Medial frontal gyrus (pre-SMA, SMA-proper).		9.283	0.00000001	6	4	63	964
Dorsal central sulcus		4.789	0.0001118	19	−29	69	815
Body of the calcarine sulcus		4.911	0.0000844	22	−69	4	303
Posterior superior temporal sulcus and gyrus		5.053	0.0000609	57	−32	6	377
Dorsal precentral sulcus		4.721	0.0001309	29	2	46	309
Posterior edge of the middle temporal sulcus		5.455	0.0000244	43	−75	−1	201
Inferior frontal gyrus pars orbitalis		4.455	0.0002430	46	31	−14	274
Superior temporal sulcus extending into the superior temporal gyrus		4.203	0.0004376	60	−17	−1	226
Parahippocampal gyrus		4.101	0.0005555	34	−16	−24	187
B. SENTENCE GENERATION (PICTURE OBSERVATION REMOVED)							
Calcarine fissure (caudal and body), lingual gyrus, cuneus, middle occipital gyrus, fusiform gyrus, collateral sulcus, inferior temporal sulcus, extending laterally and dorsally into the occipital gyrus	Left	4.246	0.0003958	−31	−91	−5	8949
Ventral central sulcus, extending caudally into the postcentral gyrus, and rostrally into the precentral gyrus, precentral sulcus, inferior frontal gyrus, including both pars opercularis and triangularis		4.756	0.0001207	−48	−12	30	9531
Intra-occipital sulcus, intraparietal sulcus		4.085	0.0005767	−30	−69	38	3643
Medial frontal gyrus (pre-SMA, SMA-proper)		4.428	0.0002588	−5	1	64	2313
Transverse temporal gyrus and sulcus, extending caudally into the planum temporale		4.185	0.0004564	−33	−35	15	1516
Inferior frontal gyrus pars orbitale, extending caudally into the orbital gyrus		3.505	0.0022292	−40	34	1	728
Posterior STS		3.722	0.0013464	−56	−46	4	616
Posterior planum temporale		3.775	0.0011899	−57	−46	24	234
Mid superior temporal gyrus		3.795	0.0011357	−63	−26	6	221
Calcarine fissure (caudal, body, and rostral), lingual gyrus, cuneus, middle occipital gyrus, fusiform gyrus, collateral sulcus, extending laterally and dorsally into the occipital pole and onto the occipital gyrus	Right	4.274	0.0341345	33	−91	−2	6577
Ventral central sulcus, extending caudally into the postcentral gyrus, and rostrally into the precentral gyrus		5.871	0.0000096	52	−2	26	4005
Transverse temporal gyrus and sulcus		3.738	0.0012972	59	−13	1	1041

(Continued)

Table 3 | Continued

Anatomical description	Hemi	Max *t*-value	*p*	*x*	*y*	*z*	num
Posterior superior temporal sulcus		3.861	0.0009735	49	−35	7	889
Medial frontal gyrus (pre-SMA, SMA-proper)		3.851	0.0009965	6	8	60	541
Collateral sulcus		3.708	0.0013911	13	−89	−9	334
Posterior inferior temporal gyrus		3.751	0.0012584	45	−62	−11	636
Intraparietal sulcus		3.557	0.0019761	19	−62	55	378
Occipital gyri		3.512	0.0021933	15	−86	41	359

All coordinates are in Talairach space and represent the peak surface node for each of the cluster (minimum cluster size: 168 contiguous surface nodes, each significant at $p < 0.005$).

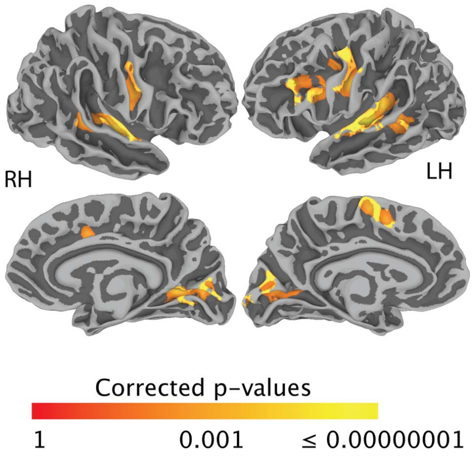


FIGURE 3 | Family-wise error-corrected group-level ($N = 20$) task difference (generation > repetition). Activation is shown on the group average smoothed white matter folded surface.

("the child throws the ball" from "throw child ball"). The comparison of this task, which places a high demand on selection and sequencing mechanisms, with a sentence-reading task, which is less taxing, revealed activation in the left pre-SMA (Haller et al., 2005). It could be argued that activation in pre-SMA is related to semantic processing, though this would be surprising given the known involvement of this region in tasks requiring volitional selection without semantic processing. For instance, Tremblay and Gracco (2010) recently showed that when participants freely choose a word or a non-speech oral motor gesture from a pool of potential responses, activation in left pre-SMA is stronger than when they produce a word or a non-speech oral motor gesture based on specific instructions. In this task, semantic processing is minimal, and importantly, in the free selection condition, selection is not based on semantics. Moreover, the fact that activation in pre-SMA does not correlate with any of our language measures supports the claim that activation in the pre-SMA is not tied specifically to language, but rather to a domain-general process. In keeping with previous findings, the present results thus suggest that the left pre-SMA is involved in selecting a response in the context of sentence production. Further, it appears that despite increased complexity, response selection in the context of sentence production engages similar mechanisms to response selection for isolated words and oro-facial gestures.

In addition to task-related activation in the pre-SMA, we also found task-related deactivation in the rostral cingulate gyrus during sentence generation that was not present during sentence repetition. The rostral cingulate area is known to be part of a putative default mode network (DMN), which was first identified through a meta-analysis of positron emission tomography studies (Shulman et al., 1997). In addition to the anterior cingulate, the DMN also includes the medial frontal cortex, the posterior cingulate cortex, precuneus, inferior parietal cortex, and the amygdala/hippocampus. It is now recognized that parts of the DMN are differentially engaged depending on task (e.g., Hasson et al., 2009; Newton et al., 2011), and it is postulated that these deactivations are the consequence of either increased or reduced task-related effort (Lin et al., 2011). In a recent study, it was shown that a cortical region including both rostral cingulate region and anterior medial frontal cortex was deactivated during a working memory task, and further, that deactivation in this region was positively correlated with working memory performance (Hampson et al., 2006), suggesting increased working memory demands for the sentence generation condition relative to the sentence repetition condition.

LATERAL PREMOTOR AREAS IN MOTOR RESPONSE SELECTION

In the present study, we examined three anatomically distinct parts of the lateral premotor area: the pars opercularis of the IFG, the rostral PMv corresponding to the ventral precentral sulcus, and the caudal PMv, corresponding to the ventral precentral gyrus. In the left hemisphere all three areas were significantly active in both sentence repetition and sentence generation, while in the right hemisphere, only the caudal PMv was significantly active (for both tasks). The left rostral and caudal parts of PMv both exhibited a significant task-related modulation, extending previous findings of a modulation of PMv activation during single word selection under different selection modes (Tremblay and Gracco, 2010). While this pattern of activation suggests a role in response selection, the finding that activation magnitude in both regions is negatively correlated with number of words produced during the sentence generation tasks seems counterintuitive. Indeed, if a linear relationship exists between these two factors, one would predict that the more words are produced (hence the more motor programs compete for selection), the more activation there should be in a region involved in response selection; this pattern was not

Table 4 | Family-wise error-corrected group-level ($N = 20$), cortical surface results for the contrast of Sentence generation against Sentence repetition.

Anatomical description	Hemi	Max *t*-value	*p*	*x*	*y*	*z*	Cluster size in nodes
Transverse temporal gyrus and sulcus	Left	8.308	0.00000	−49	−23	2	5457
Ventral precentral gyrus and central sulcus, extending caudally into the postcentral gyrus		6.711	0.00000	−46	−6	47	2142
Body of the calcarine fissure and cuneus		5.186	0.00004	−7	−82	12	1359
Inferior frontal gyrus, including both pars opercularis and triangularis		6.097	0.00001	−40	29	24	1503
Medial frontal gyrus (pre-SMA)		6.089	0.00001	−8	7	60	696
Collateral sulcus		5.494	0.00002	−42	−47	−14	286
Caudal cuneus and calcarine sulcus		5.265	0.00004	−9	−103	10	189
Parietal operculum		4.448	0.00025	−55	−16	19	241
Precentral sulcus		6.107	0.00001	−36	1	38	229
Calcarine sulcus (rostral and body), extending dorsally into the cuneus, and ventrally into the lingual gyrus	Right	6.088	0.00001	28	−60	8	2131
Posterior superior temporal sulcus		6.509	0.00000	53	−33	6	1982
Ventral precentral gyrus and central sulcus, extending caudally into the postcentral gyrus		5.112	0.00005	51	−3	26	1577
Transverse temporal gyrus and sulcus		7.678	0.00000	58	−8	−1	1265
Cingulate sulcus		4.526	0.00021	13	18	36	289

All coordinates are in Talairach space and represent the peak surface node for each of the cluster (minimum cluster size: 168 contiguous surface nodes, each significant at $p < 0.005$).

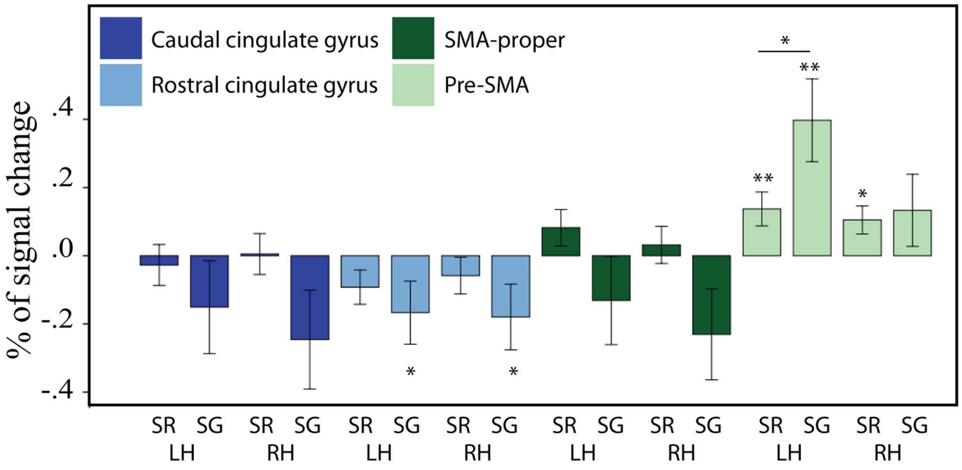


FIGURE 4 | Brain activity (expressed as a percentage of signal change) for sentence repetition (SR) and sentence generation (SG), for the medial frontal ROI group. LH = left hemisphere, RH = right hemisphere. From left to right: caudal cingulate gyrus (left, right), rostral cingulate gyrus, SMA-proper, and pre-SMA. Asterisks indicate significance level: double asterisk indicate that the statistics survive an FDR correction ($q = 0.05$); single asterisk indicate uncorrected significance. The error bars represent the SE of the mean.

found. Additional studies are required to examine further the contribution of the left PMv in response selection. Nevertheless, the present results clearly demonstrate that no lateral premotor area is more strongly involved in stimulus-driven actions than the medial regions, which challenges the "medio-lateral gradient of control" hypothesis of Goldberg (1985). In this seminal article, Goldberg described two separate systems (medial and lateral) for the control of voluntary actions. The medial system was organized around the SMA/pre-SMA, sensitive to internal events, and operated in an anticipatory mode, being primarily concerned with "volitional" actions. In contrast, the lateral was organized around the lateral premotor cortex sensitive to the external world, and operated in a responsive, interactive manner rather than being focused on internal events. The present results do not support the idea of a dual system for the control of actions. Instead, we suggest that motor response selection (whether it is volitional or stimulus-driven) is

Table 5 | Brain/behavior correlations.

ROI	Hemi	Accuracy	Words	Syllables	Fluency	Reading span
Rostral cingulate	Left	−0.05	**−0.35**	−0.31	−0.09	−0.03
	Right	−0.06	−0.22	−0.21	−0.04	0.00
Caudal cingulate	Left	−0.10	**−0.32**	−0.28	0.15	0.03
	Right	−0.10	−0.25	−0.21	0.15	−0.06
Pre-SMA	Left	0.26	−0.23	−0.13	0.00	0.09
	Right	0.18	−0.12	0.08	0.10	−0.01
SMA-proper	Left	−0.05	−0.21	−0.14	0.14	−0.21
	Right	−0.22	−0.07	−0.04	0.20	−0.16
Caudal PMv	Left	0.24	**−0.32**	−0.26	−0.03	0.14
	Right	0.03	−0.12	−0.05	0.18	0.09
Rostral PMv	Left	0.26	**−0.36**	−0.29	0.06	0.14
	Right	0.29	**−0.31**	−0.23	0.08	0.07
Opercularis	Left	0.17	**−0.38**	**−0.34**	0.09	0.18
	Right	0.23	−0.11	−0.03	**0.30**	−0.06
Triangularis	Left	0.17	**−0.33**	−0.23	0.20	0.21
	Right	0.14	−0.12	−0.05	**0.40**	0.11

Bold values indicate medium or large correlation coefficients.

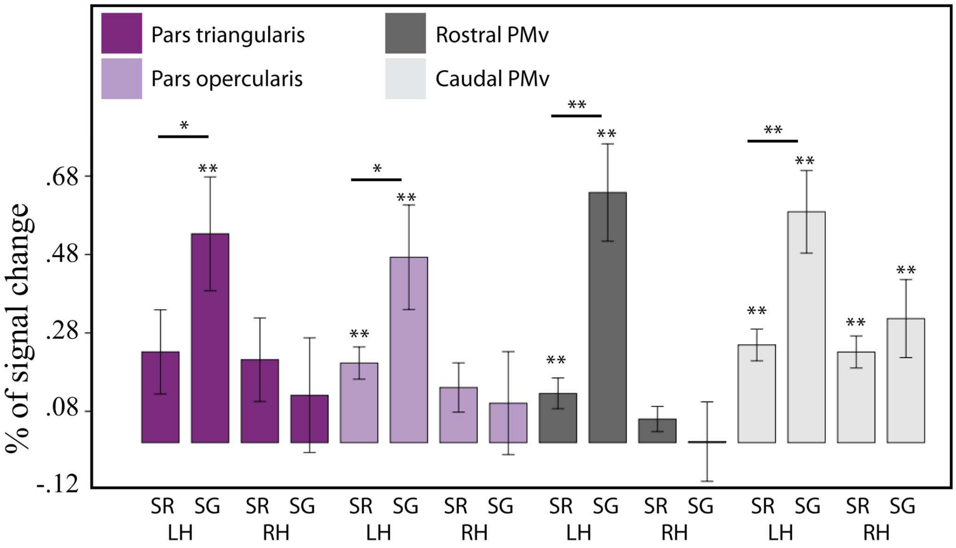


FIGURE 5 | Brain activity (expressed as a percentage of signal change) for sentence repetition (SR) and sentence generation (SG), for the lateral frontal ROI group. LH = left hemisphere, RH = right hemisphere. From left to right: IFG pars triangularis (left, right), IFG pars opercularis, rostral PMv, and caudal PMv. Asterisks indicate significance level: double asterisk indicate that the statistics survive an FDR correction ($q = 0.05$); single asterisk indicate uncorrected significance. The error bars represent the SE of the mean.

accomplished within a single system involving both the pre-SMA and the rostral and caudal parts of PMv.

THE CASE OF PARS TRIANGULARIS

The role of Broca's area in language has been a central theme in language neuroscience since the nineteenth century. Multiple functions have been proposed to account for the complex and seemingly multifold contribution of this cortical area to language, including domain-specific functions such as syntactic processes (e.g., Grodzinsky and Friederici, 2006), and more general functions such as action understanding (e.g., Fadiga et al., 2009) and information integration (Hagoort, 2005). Of particular interest in the context of the current framework is the hypothesis that the anterior sector of Broca's area, the pars triangularis, is involved in a domain-general, response selection process (Thompson-Schill et al., 1997, 1998; Robinson et al., 2005). In the present study, the left pars triangularis was significantly active in sentence generation, a task that is contingent upon semantic processing, but not in sentence repetition, a task with a limited reliance on semantic processes. This finding challenges the hypothesis of a general role for this area in response selection. As noted above, our hypothesis was that regions involved in response selection should be modulated by selection mode (generation > repetition), but also, we expected these regions to be active in both sentence production tasks since both require selection and sequencing of motor programs. Admittedly, it is possible that selection mode (volitional, externally constrained) does not affect response selection. If that were the case, one would still expect a region involved in motor response selection to be active both during sentence generation and sentence repetition, a pattern that was not found in pars triangularis.

One possible interpretation of these results is that the left pars triangularis is involved in response selection by helping resolve response competition, consistent with Thompson-Schill et al. (1997, 1998) but only when competition occurs in the linguistic/semantic domain. In line with this interpretation, previous results have shown that pars triangularis is not active for selecting single word and single oral communicative gestures when selection is not dependent upon semantic or linguistic processes (Nagel et al., 2008; Tremblay and Gracco, 2010). Moreover, evidence for a role of pars triangularis in semantic/linguistic processing abounds (e.g., Poldrack et al., 1999; Wagner et al., 2001; Devlin et al., 2003; Amunts et al., 2004; Costafreda et al., 2006). For example, results of a combined fMRI/rTMS study show that the left pars triangularis is involved in the process of making semantic decisions about words presented visually, and further shows that rTMS over the pars triangularis interferes with a semantic decision task (Devlin et al., 2003), thereby demonstrating the importance of this region for semantic processing. Taken together, these results suggest that one way in which pars triangularis contributes to language production is by helping resolve response competition when competition occurs in the semantic domain. At a more general level, the entirety of the IFG is likely to participate in a large number of neural networks that act upon language input for a variety of context-dependent purposes.

MOTOR VS. LEXICAL SELECTION IN SPOKEN LANGUAGE PRODUCTION

It could be argued that the patterns of response that were found in the left pre-SMA and PMv (generation > repetition) in the present study reflect lexical rather than motor response selection. Indeed, from sentence repetition to sentence generation, the demands on lexical selection processes increase because different lexical entries compete for expressing a given meaning. However, another explanation (that we favor) is that during spoken language production, competition for selection occurs simultaneously at multiple levels of representation (lexical, motor). Although inconsistent with serial cognitive models of spoken language production such as that of Levelt (1999), such an interpretation is in line with cascaded models of spoken language production, such as those of Peterson and Savoy (1998) and Morsella and Miozzo (2002), both of which postulate that activation spreads (cascades) from lexico-semantic representations to phonological-motor representations during the preparation for speech production, until a selection is made.

Neurobiological models also support the existence of multiple simultaneous processes. Previous biological studies suggest that lexical and motor competition/selection rely on (at least partially) distinct neural circuits (pre-SMA and PMv for motor selection, left middle temporal gyrus for lexical selection). For instance, based on a comprehensive meta-analysis of the literature on spoken language production, Indefrey and Levelt (2000) and Indefrey and Levelt (2004) identified one region that appears to be critical for lexical selection: the central portion of the left middle temporal gyrus. In contrast, selection of non-speech oro-facial actions (which does not involve lexical selection) activates the pre-SMA and PMv, but not the central portion of the left middle temporal gyrus (Braun et al., 2001; Tremblay and Gracco, 2010). Moreover, studies on finger/hand response selection have shown that motor response selection occurs at the level of the pre-SMA and PMv (e.g., Deiber et al., 1996; Van Oostende et al., 1997; Hadland et al., 2001; Ullsperger and Von Cramon, 2001; Weeks et al., 2001; Lau et al., 2004, 2006). Finally, imaging studies in which a primed picture-naming paradigm was used to elicit over verbal responses support the claim of parallel processing through anatomically segregated circuits (de Zubicaray et al., 2006). In this study, semantically primed pictures were compared to unprimed pictures and activation was found in regions involved in both phonological retrieval and lexical–conceptual processing during picture-naming, as well as in the pre-SMA, suggesting multiple levels of competition during lexical access in spoken language production. In sum, both cognitive models and neurobiological data support the claim that selection occurs simultaneously at multiple levels during spoken language production.

Admittedly, the current study was not specifically designed to disentangle the possible levels of competition. It is therefore possible, although the evidence presented here suggests otherwise, that the activation patterns found in pre-SMA and PMv reflect lexical rather than motor competition. It is also possible, though unlikely, that lexical and motor competition processes are not dissociable anatomically. Additional studies are needed to characterize further the neural underpinnings of competition and selection during spoken language production, and to the extent possible, to disentangle the simultaneous competition mechanisms and the neural networks that implement them.

CONCLUSION

In general, results of the present study help clarify the contribution of the pre-SMA, cingulate areas, PMv, and pars triangularis to the process of selecting motor responses in the context of sentence production. Further, the present results suggest that motor response selection during sentence production engages neural resources similar to those engaged in the selection of isolated words and oral motor gestures, focusing on the left pre-SMA as well as the left rostral and caudal parts of PMv.

ACKNOWLEDGMENTS

We thank T. Talavage, E. Mok, R. Siugzdaite, M. Flynn, B. Buchholz, J. Sheffield, and A. S. Dick for their help collecting and/or analyzing the data. Thanks also to U. Hasson and A. S. Dick for helpful comments on the manuscript. This study was performed at The University of Chicago, and we acknowledge the support of the Departments of Neurology and Psychology, and of the Pritzker School of Medicine during the time this study was conducted. The study was supported by a research grant from the NIDCD (R01 DC003378) to S. L. Small, and by a postdoctoral fellowship from CIHR (Canadian Institute of Health Research) to P. Tremblay. Their support is gratefully acknowledged.

REFERENCES

Alario, F. X., Chainay, H., Lehericy, S., and Cohen, L. (2006). The role of the supplementary motor area (SMA) in word production. *Brain Res.* 1076, 129–143.

Alcock, K. (2006). The development of oral motor control and language. *Downs Syndr. Res. Pract.* 11, 1–8.

Alcock, K. J., Passingham, R. E., Watkins, K. E., and Vargha-Khadem, F. (2000). Oral dyspraxia in inherited speech and language impairment and acquired dysphasia. *Brain Lang.* 75, 17–33.

Amunts, K., Weiss, P. H., Mohlberg, H., Pieperhoff, P., Eickhoff, S., Gurd, J. M., Marshall, J. C., Shah, N. J., Fink, G. R., and Zilles, K. (2004). Analysis of neural mechanisms underlying verbal fluency in cytoarchitectonically defined stereotaxic space – the roles of Brodmann areas 44 and 45. *Neuroimage* 22, 42–56.

Argall, B. D., Saad, Z. S., and Beauchamp, M. S. (2006). Simplified intersubject averaging on the cortical surface using SUMA. *Hum. Brain Mapp.* 27, 14–27.

Bates, E., D'amico, S., Jacobsen, T., Szekely, A., Andonova, E., Devescovi, A., Herron, D., Lu, C. C., Pechmann, T., Pleh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., and Tzeng, O. (2003). Timed picture naming in seven languages. *Psychon. Bull. Rev.* 10, 344–380.

Bock, J. K. (1986). Syntactic persistence in language production. *Cogn. Psychol.* 18, 355–387.

Bock, K. J. (1990). Structure in language: creating form in talk. *Am. Psychol.* 45, 1221–1236.

Braun, A. R., Guillemin, A., Hosey, L., and Varga, M. (2001). The neural organization of discourse: an H2 15O-PET study of narrative production in English and American sign language. *Brain* 124, 2028–2044.

Costafreda, S. G., Fu, C. H., Lee, L., Everitt, B., Brammer, M. J., and David, A. S. (2006). A systematic review and quantitative appraisal of fMRI studies of verbal fluency: role of the left inferior frontal gyrus. *Hum. Brain Mapp.* 27, 799–810.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173.

Crosson, B., Sadek, J. R., Maron, L., Gökçay, D., Mohr, C., Auerbach, E. J., Freeman, A. J., Leonard, C. M., and Briggs, R. W. (2001). Relative shift in activity from medial to lateral frontal cortex during internally versus externally guided word generation. *J. Cogn. Neurosci.* 13, 272–283.

Dale, A. M., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage* 9, 179–194.

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466.

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.

Deiber, M. P., Ibanez, V., Sadato, N., and Hallett, M. (1996). Cerebral structures participating in motor preparation in humans: a positron emission tomography study. *J. Neurophysiol.* 75, 233–247.

Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., and Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980.

Devlin, J. T., Matthews, P. M., and Rushworth, M. F. (2003). Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *J. Cogn. Neurosci.* 15, 71–84.

Dirnberger, G., Fickel, U., Lindinger, G., Lang, W., and Jahanshahi, M. (1998). The mode of movement selection. Movement-related cortical potentials prior to freely selected and repetitive movements. *Exp. Brain Res.* 120, 263–272.

Duvernoy, H. M. (1991). *The Human Brain: Structure, Three-Dimensional Sectional Anatomy and MRI.* New York: Springer-Verlag.

Fadiga, L., Craighero, L., and D'ausilio, A. (2009). Broca's area in language, action, and music. *Ann. N. Y. Acad. Sci.* 1169, 448–458.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., and Dale, A. M. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33, 341–355.

Fischl, B., Sereno, M. I., and Dale, A. M. (1999). Cortical surface-based analysis: II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9, 195–207.

Fischl, B., Van Der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., and Dale, A. M. (2004). Automatically parcellating the human cerebral cortex. *Cereb. Cortex* 14, 11–22.

Gentilucci, M. (2003). Grasp observation influences speech production. *Eur. J. Neurosci.* 17, 179–184.

Gentilucci, M., Benuzzi, F., Gangitano, M., and Grimaldi, S. (2001). Grasp with hand and mouth: a kinematic study on healthy subjects. *J. Neurophysiol.* 86, 1685–1699.

Goldberg, G. (1985). Supplementary motor area structure and function: review and hypothesis. *Behav. Brain Sci.* 8, 567–616.

Grodzinsky, Y., and Friederici, A. D. (2006). Neuroimaging of syntax and syntactic processing. *Curr. Opin. Neurobiol.* 16, 240–246.

Guenther, F. H., Ghosh, S. S., and Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang.* 96, 280–301.

Hadland, K. A., Rushworth, M. F., Passingham, R. E., Jahanshahi, M., and Rothwell, J. C. (2001). Interference with performance of a response selection task that has no working memory component: an rTMS comparison of the dorsolateral prefrontal and medial frontal cortex. *J. Cogn. Neurosci.* 13, 1097–1108.

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

Haller, S., Radue, E. W., Erb, M., Grodd, W., and Kircher, T. (2005). Overt sentence production in event-related fMRI. *Neuropsychologia* 43, 807–814.

Hampson, M., Driesen, N. R., Skudlarski, P., Gore, J. C., and Constable, R. T. (2006). Brain connectivity related to working memory performance. *J. Neurosci.* 26, 13338–13343.

Hasson, U., Nusbaum, H. C., and Small, S. L. (2009). Task-dependent organization of brain regions active during rest. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10841–10846.

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.

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

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

Kent, R. D. (2000). Research on speech motor control and its disorders: a review and prospective. *J. Commun. Disord.* 33, 391–427.

Lau, H., Rogers, R. D., and Passingham, R. E. (2006). Dissociating response selection and conflict in the medial frontal surface. *Neuroimage* 29, 446–451.

Lau, H. C., Rogers, R. D., Ramnani, N., and Passingham, R. E. (2004). Willed action and attention to the selection of action. *Neuroimage* 21, 1407–1415.

Levelt, W. J. (1999). Models of word production. *Trends Cogn. Sci. (Regul. Ed.)* 3, 223–232.

Lin, P., Hasson, U., Jovicich, J., and Robinson, S. (2011). A neuronal basis for task-negative responses in the human brain. *Cereb. Cortex.* 21, 821–830.

Lu, M. T., Preston, J. B., and Strick, P. L. (1994). Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *J. Comp. Neurol.* 341, 375–392.

Luppino, G., Matelli, M., Camarda, R., and Rizzolatti, G. (1993). Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. *J. Comp. Neurol.* 338, 114–140.

Luppino, G., and Rizzolatti, G. (2000). The organization of the frontal motor cortex. *News Physiol. Sci.* 15, 219–224.

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

Mushiake, H., Inase, M., and Tanji, J. (1991). Neuronal activity in the primate premotor, supplementary, and precentral motor cortex during visually guided and internally determined sequential movements. *J. Neurophysiol.* 66, 705–718.

Nagel, I. E., Schumacher, E. H., Goebel, R., and D'esposito, M. (2008). Functional MRI investigation of verbal selection mechanisms in lateral prefrontal cortex. *Neuroimage* 43, 801–807.

Newton, A. T., Morgan, V. L., Rogers, B. P., and Gore, J. C. (2011). Modulation of steady state functional connectivity in the default mode and working memory networks by cognitive load. *Hum. Brain Mapp.* 32, 1649–1659.

Nichols, T., Brett, M., Andersson, J., Wager, T., and Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage* 25, 653–660.

Nichols, T., and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15, 1–25.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.

Penfield, W., and Welch, K. (1951). The supplementary motor area of the cerebral cortex; a clinical and experimental study. *AMA Arch. Neurol. Psychiatry* 66, 289–317.

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

Pickering, M. J., and Branigan, H. P. (1999). Syntactic priming in language production. *Trends Cogn. Sci. (Regul. Ed.)* 3, 136–141.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., and Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* 10, 15–35.

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.

Rizzolatti, G., Luppino, G., and Matelli, M. (1998). The organization of the cortical motor system: new concepts. *Electroencephalogr. Clin. Neurophysiol.* 106, 283–296.

Robinson, G., Shallice, T., and Cipolotti, L. (2005). A failure of high level verbal response selection in progressive dynamic aphasia. *Cogn. Neuropsychol.* 22, 661–694.

Saad, Z., Reynolds, R., Argall, B., Japee, S., and Cox, R. (2004). SUMA: an interface for surface-based intra-and inter-subject analysis with AFNI. *IEEE Int. Symp. Biomed. Imaging* 1510–1513.

Saad, Z. S., Glen, D. R., Chen, G., Beauchamp, M. S., Desai, R., and Cox, R. W. (2009). A new method for improving functional-to-structural MRI alignment using local Pearson correlation. *Neuroimage* 44, 839–848.

Shulman, G. L., Corbetta, M., Fiez, J. A., Buckner, R. L., Miezin, F. M., Raichle, M. E., and Petersen, S. E. (1997). Searching for activations that generalize over tasks. *Hum. Brain Mapp.* 5, 317–322.

Szekely, A., D'amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., and Bates, E. (2003). Timed picture naming: extended norms and validation against previous studies. *Behav. Res. Methods Instrum. Comput.* 35, 621–633.

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

Thompson-Schill, S. L., Swick, D., Farah, M. J., D'esposito, M., Kan, I. P., and Knight, R. T. (1998). Verb generation in patients with focal frontal lesions: a neuropsychological test of neuroimaging findings. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15855–15860.

Tremblay, P., and Gracco, V. L. (2006). Contribution of the frontal lobe to externally and internally specified verbal responses: fMRI evidence. *Neuroimage* 33, 947–957.

Tremblay, P., and Gracco, V. L. (2009). Contribution of the pre-SMA to the production of words and non-speech oral motor gestures, as revealed by repetitive transcranial magnetic stimulation (rTMS). *Brain Res.* 1268, 112–124.

Tremblay, P., and Gracco, V. L. (2010). On the selection of words and oral motor responses: evidence of a response-independent fronto-parietal network. *Cortex* 46, 15–28.

Tremblay, P., Shiller, D. M., and Gracco, V. L. (2008). On the time-course and frequency selectivity of the EEG for different modes of response selection: evidence from speech production and keyboard pressing. *Clin. Neurophysiol.* 119, 88–99.

Tremblay, P., and Small, S. L. (2011). From language comprehension to action understanding and back again. *Cereb. Cortex.* 21, 1166–1177.

Ullsperger, M., and Von Cramon, D. Y. (2001). Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *Neuroimage* 6, 1387–1401.

Van Oostende, S., Van Hecke, P., Sunaert, S., Nuttin, B., and Marchal, G. (1997). FMRI studies of the supplementary motor area and the premotor cortex. *Neuroimage* 6, 181–190.

Vul, E., and Kanwisher, N. (2009). "Begging the question: the non-independence error in fMRI data analysis," in *Foundational Issues of Human Brain Mapping,* eds S. B. Hanson and M. Bunzl (Cambridge: MIT Press), 71–92.

Wagner, A. D., Pare-Blagoev, E. J., Clark, J., and Poldrack, R. A. (2001). Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron* 31, 329–338.

Wang, Y., Isoda, M., Matsuzaka, Y., Shima, K., and Tanji, J. (2005). Prefrontal cortical cells projecting to the supplementary eye field and presupplementary motor area in the monkey. *Neurosci. Res.* 53, 1–7.

Weeks, R. A., Honda, M., Catalan, M. J., and Hallett, M. (2001). Comparison of auditory, somatosensory, and visually instructed and internally generated finger movements: a PET study. *Neuroimage* 14, 219–230.

Woolsey, C. N., Settlage, P. H., Meyer, D. R., Sencer, W., Pinto Hamuy, T., and Travis, A. M. (1952). Patterns of localization in precentral and "supplementary" motor areas and their relation to the concept of a premotor area. *Res. Publ. Assoc. Res. Nerv. Ment. Dis.* 30, 238–264.

Zhang, J. X., Feng, C. M., Fox, P. T., Gao, J. H., and Tan, L. H. (2004). Is left inferior frontal gyrus a general mechanism for selection? *Neuroimage* 23, 596–603.

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

Received: 03 April 2011; accepted: 13 September 2011; published online: 30 September 2011.

Citation: Tremblay P and Small SL (2011) Motor response selection in overt sentence production: a functional MRI study. Front. Psychology ***2****:253. doi: 10.3389/fpsyg.2011.00253*

This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 08 February 2012
doi: 10.3389/fpsyg.2012.00026

Neural specificity for grammatical operations is revealed by content-independent fMR adaptation

Kevin A. Shapiro[1,2,3], Lauren R. Moo[3] and Alfonso Caramazza[1,4]*

[1] Department of Psychology, Harvard University, Cambridge, MA, USA
[2] Department of Medicine, Children's Hospital Boston, Boston, MA, USA
[3] Department of Neurology, Massachusetts General Hospital, Boston, MA, USA
[4] Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy

Edited by:
Kristof Strijkers, University of Barcelona, Spain

Reviewed by:
Ian G. Roberts, University of Cambridge, UK
Laura Menenti, Max Planck Institute for Psycholinguistics, Netherlands

***Correspondence:**
Alfonso Caramazza, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA.
e-mail: caram@wjh.harvard.edu

The ability to generate novel sentences depends on cognitive operations that specify the syntactic function of nouns, verbs, and other words retrieved from the mental lexicon. Although neuropsychological studies suggest that such operations rely on neural circuits distinct from those encoding word form and meaning, it has not been possible to characterize this distinction definitively with neuroimaging. We used functional magnetic resonance imaging (fMRI) to show that a brain area engaged in a given grammatical operation can be identified uniquely by a monotonic decrease in activation as that operation is repeated. We applied this methodology to identify areas involved selectively in the operation of inflection of nouns or verbs. By contrast, areas involved in processing word meaning do not show this monotonic adaptation across stimuli. These results are the first to demonstrate adaptation in the fMR signal evoked not by specific stimuli, but by well-defined cognitive linguistic operations.

Keywords: nouns, verbs, grammatical class, inflection, fMRI, adaptation

INTRODUCTION

Several reports in the neuropsychological literature have described patients whose patterns of brain damage have resulted in disproportionate difficulty using either nouns or verbs in speech (Miceli et al., 1984; McCarthy and Warrington, 1985; Caramazza and Hillis, 1991; Damasio and Tranel, 1993; Daniele et al., 1994). In many of these cases, the apparent grammatical deficit may reflect the destruction of neurons involved in representing features of entities or actions, the prototypical referents of nouns and verbs, respectively. Thus, patients with lesions affecting higher-order visual association areas in the temporal lobe tend to have more difficulty with object words (produced as nouns), while those with lesions in premotor areas fare more poorly in naming action words (verbs) (Damasio and Tranel, 1993; Daniele et al., 1994; Bak et al., 2001; Luzzatti et al., 2006).

Interestingly, there are patients whose impairment seems specifically to affect the use of words in a particular grammatical context, while access to word meaning is relatively unaffected (Shapiro et al., 2000; Berndt et al., 2002; Shapiro and Caramazza, 2003; Laiacona and Caramazza, 2004; Crepaldi et al., 2006). Some patients are unable to produce even nonsense words marked inflectionally as nouns (e.g., *the toves*), although they are able to produce the same strings as verbs (*he toves*) (Shapiro et al., 2000); other patients exhibit the mirror dissociation (Shapiro and Caramazza, 2003; Laiacona and Caramazza, 2004). These cases support the idea that there are neuroanatomically dissociable regions involved in the processing of nouns and verbs as grammatical objects.

Limited anatomical comparisons suggest that neural regions in the left prefrontal cortex may be crucial to support the distinction between nouns and verbs in grammatical processing (Shapiro et al., 2001; Shapiro and Caramazza, 2003; Cappelletti et al., 2008). However, category-selective deficits are present in only a subset of patients with left prefrontal lesions. Most patients with lesions affecting the left frontal operculum (Broca's area) and underlying structures, as well as the striate nuclei of the basal ganglia (Longworth et al., 2005), do have some degree of impairment in producing morphological affixes (Goodglass, 1973; Dronkers et al., 2000). These impairments often do not respect grammatical category distinctions (Miceli et al., 1989), but may selectively affect certain kinds of morphological operations for words of all categories, like the production of regular inflectional affixes (Marslen-Wilson and Tyler, 1997; Ullman et al., 1997; Faroqi-Shah and Thompson, 2004). Such patterns suggest that distinct cortical regions are engaged in processing morphological information within and across grammatical categories.

Studies using repetitive transcranial magnetic stimulation (rTMS) in unimpaired subjects have supported this notion, and have helped to identify anatomically dissociable circuits within the left prefrontal cortex that may be involved in category-specific and category-general morphological operations. Two studies have demonstrated that grammatical operations involving verbs, but not nouns, are disrupted selectively by stimulation to a circumscribed portion of the left mid-frontal cortex that is superior to Broca's area and well anterior to motor planning areas (Shapiro et al., 2001; Cappelletti et al., 2008). By contrast, stimulation to an adjacent portion of Broca's area affects noun and verb production equally (Cappelletti et al., 2008). One interpretation of this pattern is that the left mid-frontal region is important for handling abstract grammatical (morphosyntactic) information that pertains to verbs (Shapiro et al., 2001; Shapiro and Caramazza, 2003; Cappelletti et al., 2008), while Broca's area is important for category-general morphological processes, and perhaps specifically for the phonological realization of morphological features (Kean, 1978; Obler et al., 1999).

Functional neuroimaging methods have also been applied in an attempt to elucidate the neural circuits involved in the grammatical processing of nouns and verbs. Interestingly, these experiments for the most part have not provided evidence to corroborate the anatomical distinction that is strongly implied by neuropsychological and TMS studies. Imaging studies have shown some differences in activation evoked by nouns and verbs in the left inferior prefrontal cortex, but these differences tend to correlate with the linguistic complexity of inflected noun and verb stimuli, rather than with categorical grammatical operations as such (Perani et al., 1999; Tyler et al., 2004; Sahin et al., 2006; Longe et al., 2007). When noun and verb stimuli are well matched, as in the case of homophones, no differences between categories are observed in Broca's area (Tyler et al., 2008; Burton et al., 2009).

Several studies have shown anatomical dissociations between nouns and verbs in other regions, particularly in the left middle temporal lobe (Perani et al., 1999; Shapiro et al., 2006; Yokoyama et al., 2006; Palti et al., 2007; Bedny et al., 2008; Liljeström et al., 2008; Burton et al., 2009; Khader et al., 2010). These studies have employed a wide variety of tasks involving both production and comprehension, including phrase completion (Shapiro et al., 2006; Khader et al., 2010), picture naming (Liljeström et al., 2008), lexical decision (Perani et al., 1999; Yokoyama et al., 2006), semantic judgment (Palti et al., 2007; Bedny et al., 2008), and grammatical judgment (Palti et al., 2007; Burton et al., 2009). Regardless of task, the observed dissociations in areas outside the left prefrontal cortex are most plausibly related to differences in word meaning – even when subjects perform tasks that do not involve semantic processing, or specifically emphasize the manipulation of grammatical information.

The mixed results of functional imaging studies on the question of noun–verb differences may be attributable in part to the inherent difficulty of constructing an imaging experiment to isolate differences in a single processing component among many that are engaged during even a simple inflection task. One recent review has gone so far as to suggest that the neural circuits underlying noun and verb processing, if they exist, are irresoluble using current imaging methods (Crepaldi et al., 2011). A potential means of circumventing this problem may lie in the properties of the dependent variable in imaging studies – in the case of fMRI studies, the blood oxygen level dependent (BOLD) signal. Standard event-related fMRI contrast analyses assume that the amplitude of the BOLD signal evoked by a given task remains constant over time. Recently, however, numerous investigators have shown that when a particular stimulus is repeated, the activation associated with processing that stimulus is reduced (Henson et al., 2000; Grill-Spector et al., 2006). This effect, known as fMR adaptation or repetition suppression (Grill-Spector et al., 2006), has been exploited productively as a means of making inferences about the specificity of brain regions for processing different kinds of representational content in various domains, including object naming (van Turrennout et al., 2003), semantic and phonological processing (Wagner et al., 2000; Gold et al., 2005), and lexical and syntactic processing (Menenti et al., 2011).

Here we are interested in identifying not brain areas that process a particular kind of content, but areas specialized for performing particular sets of operations, namely, the context-dependent morphosyntactic marking of nouns and verbs. Generalizing from earlier studies demonstrating *stimulus*-specific fMR adaptation, we reasoned that neural circuits engaged in particular cognitive operations of this sort may show *operation*-specific continuous adaptation during repetition of a task regardless of the content of individual stimuli. There is some evidence that brain regions may show a decrement in activity when a task is repeated using novel stimuli (Martin et al., 1997; Toni et al., 1998). Toni et al. (1998), for example, observed that there was a decrease in the magnitude of the BOLD signal over time in parts of the prefrontal, premotor, and parietal cortex as subjects learned simple motor sequences. Our hypothesis can be viewed as an extension of this finding.

We therefore designed an fMRI experiment in which subjects were required to complete simple phrases with grammatically appropriate forms of nouns and verbs and looked for brain regions demonstrating a linear decrement in activity for morphologic operations involving nouns or verbs. In the first session of the experiment, the stimuli consisted of meaningless pseudowords whose grammatical category was indicated solely by phrasal context (*he zeeks, they…; many glushes, one*…). In the second session, the stimuli were real nouns and verbs (*one loss, many…; he sings, they*…). The use of pseudowords underscores the emphasis on identifying target words according to their syntactic roles, rather than on retrieval of lexical–semantic information.

Half of the trials in each session required subjects to change the morphological form of the stimulus word; in the other half, subjects were required to produce the stimulus word in the same form. Thus, words were produced in both their inflected (*zeeks*) and unmarked forms (*zeek*), and the target form was not predictable from the stimulus phrase. Moreover, for half of the real and pseudowords in each category, the inflected form differed in syllabic structure from the unmarked form (e.g., *glush/glushes*). These manipulations were intended to ensure that contrasts in activation evoked by nouns and verbs could be interpreted as reflecting primarily grammatical operations, rather than processes at the level of word meaning or phonological form.

We constructed two factorial models for random effects analysis of the functional data. In the first model (the "activation model"), we used statistical parametric maps corresponding to conventional event-related contrasts, based on standard assumption that the amplitude of the hemodynamic response to an event (a noun or verb trial) does not vary over time. In the second model (the "adaptation model"), we used maps which were estimated assuming a linear decrease in the amplitude of sequential trials of a given event type (i.e., grammatical operation), no matter the form of the stimulus. Again, this differs from prior fMRI adaptation experiments, which have generally examined stable differences in activation between repeated and novel stimuli. Our choice of a linear function, as opposed to a higher-order function, was arbitrary. It was our intention simply to determine whether modeling a monotonic decrease in signal could capture important information about the neural correlates of grammatical processing.

In this context, three strong predictions follow from the hypothesis that neural adaptation occurs for the cognitive operations involved in noun and verb inflection. The first is that adaptation for both noun and verb trials should be observed in brain regions known to be important for processing regular morphological inflection irrespective of grammatical category. These regions may include the left inferior prefrontal cortex and the striate nuclei of the basal ganglia (Ullman et al., 1997; Longworth et al., 2005), areas that are thought to constitute a circuit subserving regular inflectional affixation.

Although activation has been observed in these areas in neuroimaging experiments that have examined the processing of inflected words, such experiments have also shown activation in areas *not* specifically implicated in inflection in the aphasiological literature, including premotor areas that may be engaged in articulation (Sahin et al., 2006) and orbital prefrontal and temporal areas that may support semantic processes (Tyler et al., 2005; Sahin et al., 2006). These prior studies, like many other neuroimaging studies, have not been able to distinguish areas that are activated in a given contrast from areas that are specifically dedicated to or crucial for the task of interest. By contrast, we predict that grammatical operation dependent adaptation should not be expected in regions that handle information that varies between trials, including information about the phonological or semantic properties of individual stimulus items.

A final and crucial prediction is that grammatical category-specific adaptation for verb trials should be observed in the portion of the left middle frontal gyrus whose disruption with rTMS selectively interferes with the grammatical processing of verbs. (An analogous region engaged in grammatical processing of nouns has not yet been identified.) Such a finding would suggest that content-independent task or operation based adaptation effects can be used not only as a means of discriminating between areas activated by a task and areas dedicated to a task, but also as a sensitive means of identifying areas critical for the performance of well-defined cognitive operations.

MATERIALS AND METHODS

SUBJECTS

Twenty-two right-handed native English speakers (11 male) participated in the experiment, which was approved by institutional review boards at Harvard University and Massachusetts General Hospital. Informed consent was obtained from all subjects, and the experiment was conducted according to the principles expressed in the Declaration of Helsinki.

MATERIALS

The stimuli consisted of 104 pseudowords in session 1 and 104 monosyllabic real words (52 nouns and 52 verbs) in session 2. The real words were matched across categories for frequency (Francis and Kucera, 1982), length in phonemes, and length in letters. Nouns and verbs were also matched for the number of consonants in the codas of the words; moreover, half of the words in each category ended in sibilant consonants, while half did not. The latter two criteria were intended to ensure that noun and verb stimuli were matched with respect to the phonological complexity of morphologically inflected forms. Pseudowords were generated by modifying the initial phonemes of the real words (e.g., *seek/zeek*).

The behavioral paradigm was identical to that used in an earlier experiment (Shapiro et al., 2006). In brief, subjects were presented with a written stimulus phrase (e.g., *he zeeks*) for 1 s, followed for 1 s by a cue phrase (*they*) indicating the form in which the subject should produce the word or pseudoword that appeared in the stimulus phrase. In the preceding example, the subject would have been required to say "zeek." Trials were separated by the appearance of a fixation cross for intervals of varying duration between 2 and 6 s. We have previously demonstrated that this task does not produce differences in reaction time by category (noun vs. verb) or lexical status (real word vs. pseudoword) for stimuli matched in the manner described above (Shapiro et al., 2006). In the current experiment, each stimulus word appeared in two trials (only one of which required the word to be produced in the inflected form), so that there were 208 unique trials per session.

Prior to the fMRI session, subjects completed a brief training session on a laptop computer. During the training session the subjects were instructed to respond aloud, but without moving their jaws, as if they were practicing ventriloquy. This method has minimizes head motion artifacts that result from the production of spoken responses inside the scanner. The training session used a set of pseudowords that did not occur in the experiment. In the scanner, the stimuli were projected using a collimating lens onto a mirror affixed to the head coil at a comfortable viewing distance. Responses were monitored by intercom, but were not recorded.

DATA ACQUISITION AND PRE-PROCESSING

Whole brain functional images were acquired on a 3-T Siemens MR scanner using a T2*-weighted echoplanar imaging sequence (repetition time (TR) = 2 s, echo time (TE) = 30 ms, flip angle = 90°, 64 × 64 acquisition matrix, 33 slices, no gap, 4 mm^3 isotropic resolution). Pre-processing of the functional images in SPM5 included slice-time correction, realignment and unwarping, normalization to a standard template for echoplanar imaging sequences, resampling at 2 mm^3, spatial smoothing with an 8-mm^3 full width at half maximum Gaussian kernel, and application of a temporal high-pass filter.

For all subjects, we also acquired high-resolution T1-weighted anatomical images (TR = 30 ms, TE = 3.3 ms, 128 slices, no gap, 1 mm × 1 mm × 1.33 mm resolution). The anatomical images were spatially coregistered using SPM5 and the automated segmentation algorithm included in this software package was used to generate maps of gray matter voxels, including cortical, subcortical, and cerebellar gray matter while excluding white matter and cerebrospinal fluid spaces. These maps were then averaged to generate a mean gray matter map, which was used for masking in analysis of the functional data (see below) and for purposes of display (figures show areas of functional activation projected on a surface rendering derived from the mean gray matter map).

STATISTICAL MODELING

Using the pre-processed functional images, we constructed first-level statistical models for each participant in which square waves representing onsets and durations of trials of each event type were convolved with a canonical hemodynamic response function, generating a basis function for each event type. There were four event types (real nouns, real verbs, pseudonouns, and pseudoverbs), in addition to fixation. The model also included two covariates to account for length in letters of the visual stimuli and the phonological complexity of the form being manipulated (sibilant or non-sibilant ending).

Crucially, for each event type we included an "adaptation" parameter modeling a linear decrement in the amplitude of the square wave function over time. This parameter simply made use of the option provided in SPM5 for first-order time modulation of an event type in a first-level model, predicting greater positive deviation from the average hemodynamic signal on earlier trials, and smaller positive deviation or a greater negative deviation (representing a relative suppression of activation) on later trials. In other words, it models a decrement in activation for a given trial type over time. The decrement is determined over a sequence of trials of the same type, irrespective of the position of a given trial within a series or block (see **Figure 1** for an illustration). This is similar to the type of analysis described by Toni et al. (1998), except that in that paper, the authors modeled changes over time with polynomial functions up to the third order. We also did not remove low-frequency changes over time that are common to conditions, as we directly compared experimental conditions to each other – in which case common effects would be expected to cancel out – rather than to a baseline.

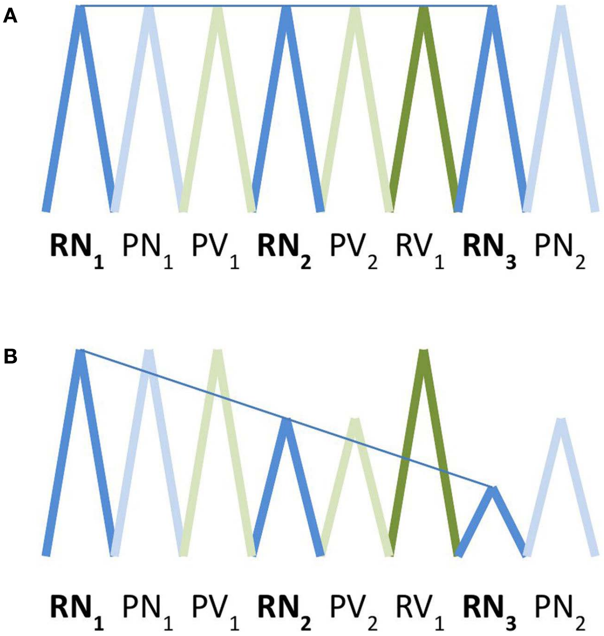


FIGURE 1 | Schematic depiction of basis functions used for modeling of activation (A) and adaptation (B). Square wave functions convolved with the standardized hemodynamic response function are represented here as triangle waves for ease of illustration. In the activation model, the amplitude of activation within a voxel for a given trial type was expected to remain constant over time. In the activation model, the amplitude of activation for a given trial type was expected to decrease in a linear fashion over successive trials. RN, real noun; RV, real verb; PN, pseudonoun; PV, pseudoverb.

We acknowledge that choosing a first-order function is a simplistic assumption both computationally and physiologically; we have no compelling reason to suppose that operation-specific adaptation is best modeled by a linear function as opposed to some other function, such as exponential decay. Indeed it is possible that the shape of the adaptation function differs in different brain regions, as in the study by Toni et al. (1998). In this sense the choice of a linear function can be regarded as a strong test of an underconstrained hypothesis. We return to this point in the Discussion.

After the first-level models including the basis functions and adaptation parameters for each event type were completed, we selected statistical parametric maps corresponding to β-weights of correlation between the predicted and observed hemodynamic response for nouns, verbs, pseudonouns, and pseudoverbs. We entered these into second-level ANCOVA models. Separate models were constructed using maps estimated using the standard event-related function and maps estimated using the linear adaptation parameter. In estimating the models, we applied a mask representing the mean distribution of gray matter voxels across subjects, as described in the Methods (Data Acquisition and Pre-Processing).

To obtain a general view of brain areas engaged by the task, we examined average effects using a voxel-wise type I error threshold of $\alpha = 0.0005$, corrected for family-wise error. We applied an arbitrary cluster extent threshold of $k = 80$ resampled voxels, equivalent in spatial extent to 10 original (non-resampled) voxels. In the activation model, the resulting statistical parametric map shows areas in which activity was correlated with the behavioral task, regardless of trial type, or position within the series. In the adaptation model, the corresponding map shows areas in which activity is correlated with the task and decreases for sequential trials of the same type, over all trial types.

For main effects of grammatical category in each model, we adopted a voxel-wise type I error threshold of $\alpha = 0.005$ and used the cluster extent method to correct for multiple comparisons (Slotnick et al., 2003). Areas surpassing a corrected cluster-wise type I error threshold of $\alpha = 0.05$ ($k > 56$ voxels) were selected for further analysis to determine the directionality of category-specific main effects and to test for interactions. For this *post hoc* analysis, we extracted subject wise first eigenvariate values for each significant cluster as an estimate of a region's hemodynamic response to processing nouns and verbs. These values were then entered into ANOVA models with lexical status and grammatical category as within-subjects repeated measures. A Bonferroni correction ($\alpha = 0.05$) was applied to account for comparisons across multiple areas. The values of the first eigenvariates corresponding to activation and adaptation for trials of each category within each significant cluster are plotted in **Figure 2**.

RESULTS

Using both the activation and adaptation fMRI analysis models, we aimed to identify areas with task-related effects irrespective of grammatical category, as well as areas showing selective effects for one grammatical operation.

TASK GENERAL EFFECTS

Using the activation model, we found that grammatical category-general activation was most robust in a cortical network

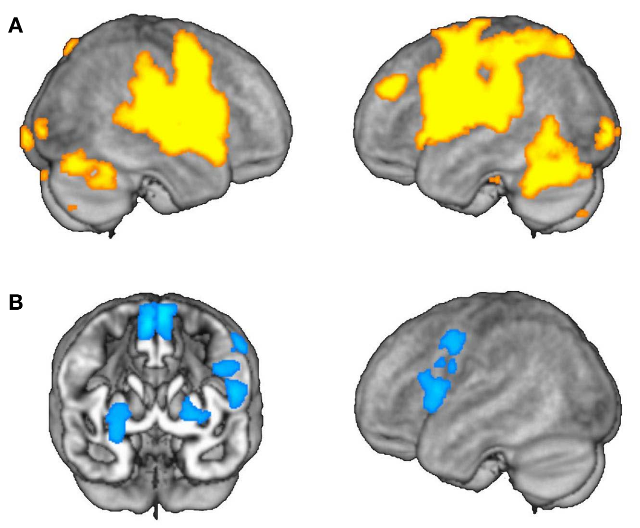


FIGURE 2 | Average effects of activation and adaptation. (A) Areas showing average effects of task-related activation, irrespective of grammatical category. **(B)** Areas showing average effects of adaptation, irrespective of grammatical category.

Table 1 | Areas activated for nouns and verbs in the morphological transformation task (*k* = cluster size).

	k	**MNI coordinates**		
		x	***y***	***z***
NOUN ∪ VERB				
Left superior temporal gyrus	10146	−58	−16	4
Left precentral gyrus		−46	−16	36
Right precentral gyrus	10296	64	−2	14
Right superior temporal gyrus		60	−16	−2
Left middle frontal gyrus	445	36	−52	−18
Left occipital pole	382	−18	−96	−4
Left fusiform gyrus		−30	−78	−10
Right fusiform gyrus	661	36	−52	−18
Right occipital pole		18	−94	2
Left anterior cingulate gyrus	2067	−6	19	30
Right anterior cingulate gyrus		6	22	38
Right posterior cingulate gyrus	261	2	−52	16
Right inferior parietal lobule	186	40	−58	54
Right middle occipital gyrus	93	32	−82	20
Right caudate	224	−12	0	2
Right putamen	112	18	4	16
Right cerebellum	5907	18	−62	−26
Left cerebellum		−18	−62	24
VERB > NOUN				
Left inferior parietal lobule	204	−44	−40	50
Left middle/superior temporal gyrus	272	−54	−40	4
Left middle frontal gyrus	131	−50	8	38
NOUN > VERB				
Left fusiform gyrus	65	−24	−28	−28

Table 2 | Areas showing adaptation for nouns, verbs, or words of both categories in the morphological transformation task (*k* = cluster size).

	k	**MNI coordinates**		
		x	***y***	***z***
NOUN ∪ VERB				
Left medial/superior frontal gyrus	911	−2	0	64
Left inferior frontal gyrus, pars opercularis	385	−52	8	10
Left middle frontal gyrus	239	−52	2	42
Left putamen	97	−22	0	6
Right putamen	249	26	4	−12
VERB > NOUN				
Left middle/inferior frontal gyrus	60	−52	24	24

encompassing bilateral perisylvian and precentral regions and the anterior parts of the cingulate gyri, as well as the cerebellar hemispheres and bilateral subcortical nuclei (**Table 1**; **Figure 2A**). On the other hand, the adaptation model identified linear decreases in operation-related activation for both nouns and verbs primarily in the left inferior and posterior frontal cortex (including Broca's area) and in the lentiform nuclei and supplementary motor areas (SMAs) bilaterally (**Table 2**; **Figure 2B**).

GRAMMATICAL CATEGORY-SPECIFIC EFFECTS

The activation model revealed a significant main effect of grammatical category in seven clusters of contiguous voxels, including areas in the left posterior middle frontal gyrus, left posterior superior temporal sulcus/left middle temporal gyrus, left ventral temporal cortex, bilateral inferior parietal lobules, and bilateral occipital poles (**Table 1**; **Figure 3**).

We then interrogated these areas to determine whether they were preferentially activated for nouns or verbs, and whether these effects were modulated by lexical status, as described above. We found that three areas (left posterior middle frontal gyrus, left middle temporal gyrus, and left inferior parietal lobule) were activated more for verb trials than for noun trials. A fourth area (right inferior parietal lobule) also showed relatively greater activation for verb trials, but this difference did not survive correction for multiple comparisons in the *post hoc* analysis; we therefore did not consider this area to show significant category selectivity. One area (left ventral temporal cortex) showed greater activation for noun trials. In the remaining two areas (occipital poles), the apparent category effect was due to a significant *decrease* in activation for verb trials, while there was no change in the baseline for noun trials (**Figure 4A**).

In two areas (left inferior parietal lobule and left posterior middle frontal gyrus), the *post hoc* analysis also identified main effects of lexical status (real or pseudoword), with relatively greater activation for pseudoword trials than for real word trials. However, no significant interaction between lexical status and grammatical category was observed in any area. Overall, these findings are concordant with the results of an earlier study using the same paradigm (Shapiro et al., 2006), except for the occipital effect, which was not previously noted.

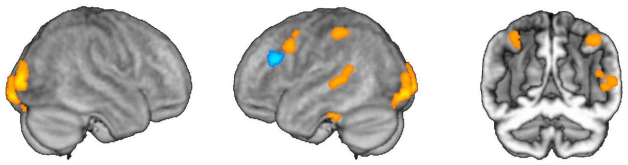


FIGURE 3 | Selective effects of activation and adaptation. Regions depicted in orange showed main effects of grammatical category in event-related activation. Regions depicted in blue showed main effects corresponding to adaptation for category-specific operations.

We then conducted an identical analysis for the adaptation model. Here, the results corresponding to the main effect of grammatical category were confined to an area near the junction of the left inferior and middle frontal gyri, which showed greater adaptation for verbs. There was no effect of lexical status, and no interaction between lexical status and grammatical category. No areas showed greater adaptation for nouns at the statistical thresholds we employed (**Table 2**; **Figures 3** and **4B**).

In a final analysis, we selected the four areas in which we observed grammatical category-selective activation in the activation model and adopted them as volumes of interest (VOIs) in the adaptation model, to determine whether these regions might show sub-threshold adaptation effects. (We did not include the areas in which categorical effects were driven exclusively by decreased activity.) No significant differences in "operation adaptation" across categories were observed within the regions that showed category specificity in the activation analysis (**Figure 4B**).

DISCUSSION

The combination of activation and operation-specific fMR adaptation analyses proved to be a potentially useful technique for identification areas selectively involved in the processing of nouns and verbs. Using a conventional activation model, we found that there was category-selective activation for verb production in the left posterior middle frontal gyrus, the left medio-lateral temporal cortex, and the left inferior parietal lobule. By contrast, noun production was associated with greater activation in the left fusiform gyrus. These results replicate the observations of an earlier fMRI study using a similar paradigm (Shapiro et al., 2006), and overlap in part with findings from other studies which employed a variety of production and comprehension tasks (Perani et al., 1999; Shapiro et al., 2005; Yokoyama et al., 2006; Palti et al., 2007; Bedny et al., 2008; Liljeström et al., 2008; Burton et al., 2009; Khader et al., 2010).

In addition, we found one brain area in which the amount of activation decreases over time for morphological transformations involving only one category of words, an effect we have called operation-specific adaptation. Interestingly, the operation-specific adaptation for verbs was observed in a more anterior portion of the middle frontal gyrus than the area in which activation for verbs was observed in this study and previous studies. However, previous work with rTMS has demonstrated that targeted suppression of this area results in selective interference for verbs in the same behavioral task used here (Shapiro et al., 2001; Cappelletti et al., 2008).

These findings have numerous limitations, not the least of which is that they rely on a specific and admittedly arbitrary assumption about the rate of adaptation. A more rigorous approach might involve analyzing the raw data on a subject-by-subject basis in order to derive empirically a function that describes changes in activation levels over time. On the other hand, assuming a simple linear adaptation parameter, as we have here, allows us to test hypotheses about the relationship between cognitive operations and neural activity using a technique that is easily implemented and not very computationally intensive.

With this caveat, we contend that the data presented here provide interesting insights into the neuroanatomical substrates of cognitive processes involved in language production. In general, the areas of category-general activation correspond to cortical regions engaged in various stages of word production, including the retrieval of phonological word forms, phonetic coding and articulation (Indefrey and Levelt, 2000), and lexical and semantic processing (Menenti et al., 2011). These processes are common to all tasks requiring spoken output.

By contrast, some of the areas in which strong operation adaptation was observed – namely, the left inferior frontal gyrus (Caplan, 2001; Friederici and Kotz, 2003; Menenti et al., 2011) and the striatum (Ullman et al., 1997; Friederici and Kotz, 2003; Longworth et al., 2005) – have been implicated specifically in the processing of phrase structure and grammatical inflection. Indeed, suppression of this region of the left inferior frontal gyrus with rTMS interferes with performance of this task for nouns and verbs equally (Cappelletti et al., 2008). The finding that these areas adapt to a task that emphasizes regular morphological affixation confirms the first prediction made in the Introduction, which was that operation-specific adaptation across categories should be found in parts of the brain that are generally involved in syntactically driven processes, including (but not necessarily limited to) morphological inflection. We did not observe adaptation for the morphological transformation task in brain areas thought to support other aspects of language production, such as the left temporal lobe.

Operation adaptation in the SMA was not predicted, but emerged quite robustly in the adaptation analysis. It is possible that this represents a general effect of cognitive adaptation, perhaps reflecting the reallocation of attentional resources as subjects became more practiced and familiar with the task (Rushworth et al., 2007). Aphasic patients with lesions in the SMA are reported to have difficulty initiating speech, but preserved fluency in repetition and in answering questions (Ziegler et al., 1997; Pai, 1999), an observation consistent with the idea that this region may indeed be less heavily recruited when a task is repeated or richly supported by the pragmatic context. On the other hand, there is some evidence that the medial SMA may be particularly important for the production of morphologically inflected forms (Sahin et al., 2006), and it may be that some portion of the area plays a critical role in selecting and encoding morphemes, including inflectional affixes (Alario et al., 2006).

The data also confirm our second prediction: the left mid-frontal area found here to exhibit selective operation adaptation for verb trials corresponds to the area whose targeted suppression by TMS has been shown to result in a relative disruption in verb

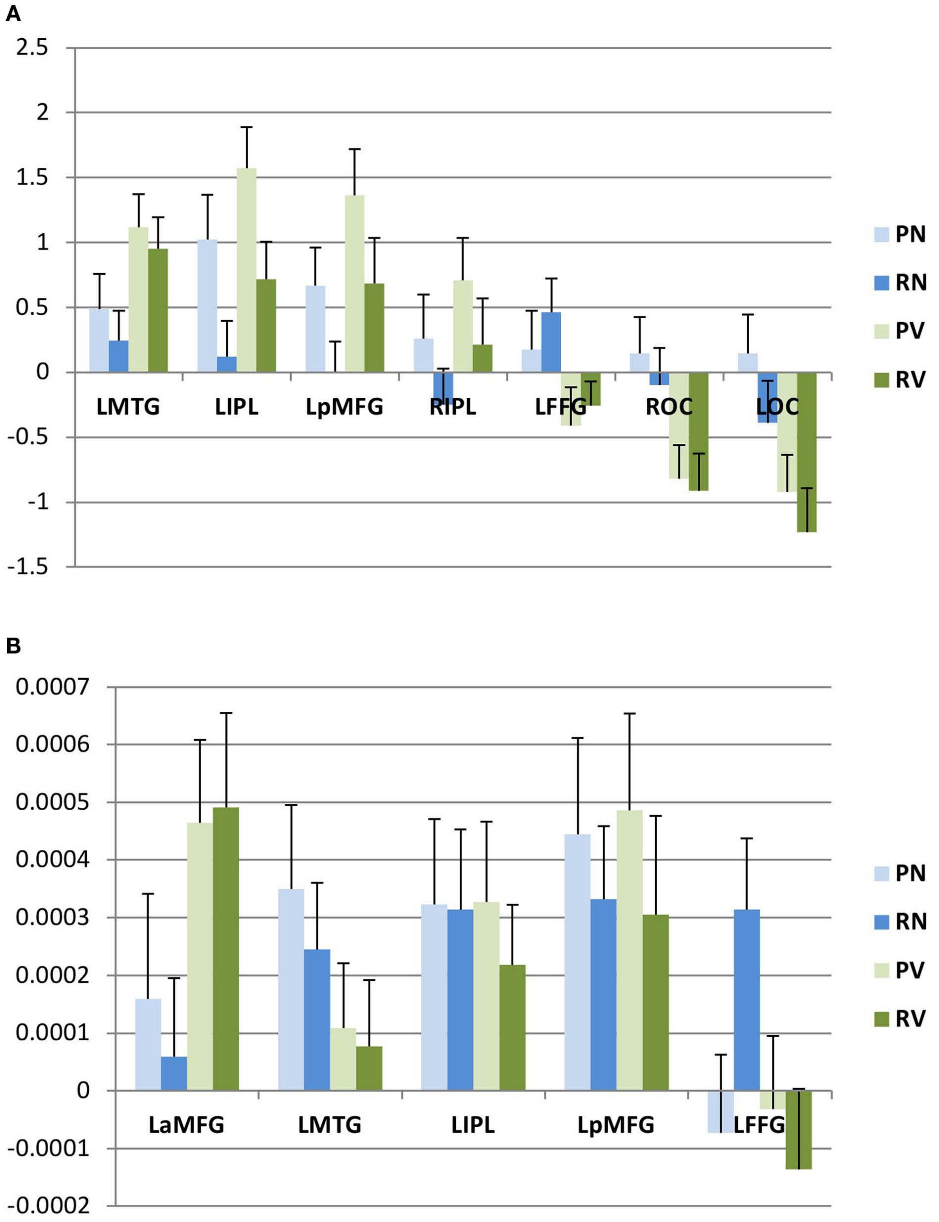


FIGURE 4 | *Post hoc* analysis. Within areas showing category-selective effects, we evaluated first eigenvariates for activation **(A)** and adaptation **(B)** by subject. Mean values (and standard errors of the mean) for these eigenvariates are plotted here. L, left; R, right; B, bilateral; a, anterior; p, posterior; MFG, middle frontal gyrus; IPL, inferior parietal lobule; MTG, middle temporal gyrus; FFG, fusiform gyrus; OC, occipital cortex.

processing (Shapiro et al., 2001; Cappelletti et al., 2008). By contrast, TMS produces no category-specific effects when applied to the more posterior left frontal area that emerged in the conventional analysis (Cappelletti et al., 2008). This finding supports our hypothesis that brain areas crucial for the performance of a cognitive operation should exhibit hemodynamic adaptation as the task is repeated. The adaptation model appears to be both more sensitive and more specific than the (conventional) activation model for the identification of areas engaged in grammatical processing.

Here, we have defined the task of interest as selecting contextually appropriate inflectional forms of nouns and verbs. For verb trials, this depended on the computation of subject–verb agreement, which may be the function performed by the left anterior mid-frontal gyrus. We do not, however, claim that this area is necessarily specific for a particular morphosyntactic operation; another possibility is that this area is more generally engaged in the processing of verbs as defined by their syntactic context, which may include computation of agreement, tense, mood, and other properties that distinguish verbs from nouns in a given language (c.f. Shapiro et al., 2001). While these alternatives cannot be differentiated using the data at hand, it is clear that the region in question shows adaptation for a grammatical operation that applies to words independently of their meaning or phonological form.

Conversely, areas whose event-related response functions are similar across different cognitive operations may be engaged in processing features of stimuli that differ across trials independent of the specific operation performed in each trial. In the case at hand, this might include unique phonological, lexical, or semantic properties of words to be produced. For example, selective activation for nouns or verbs in areas putatively linked to semantic processing, such as the left medio-lateral temporal cortex, could reflect attempts to retrieve specific features of meaning within the broad domains of knowledge related to entities and events – arguably an automatic aspect of natural speech, though not the one this task was designed specifically to probe. Thus, more generally, different trends in event-related response functions may correspond to distinct levels of cognitive processing. If this is correct, the application of operation adaptation alongside content adaptation approaches and conventional event-related analyses of fMRI data may help to reveal hitherto obscure distinctions in the organization of neural networks that support human cognition.

ACKNOWLEDGMENTS

We thank T. Daley for assistance in collecting data; S. Slotnick for critical discussions; and R. Buckner, D. Caplan, D. Holtzman, and J. Schwarzbach for review of the manuscript. The authors are funded by the National Institute on Deafness and Other Communication Disorders (Alfonso Caramazza, Lauren R. Moo) and by the National Institute of Neurological Disorders and Stroke (Kevin A. Shapiro).

REFERENCES

Alario, F. X., Chainay, H., Lehericy, S., and Cohen, L. (2006). The role of the supplementary motor area (SMA) in word production. *Brain Res.* 1076, 129–143.

Bak, T. H., O'Donovan, D. G., Xuereb, J. H., Boniface, S., and Hodges, J. R. (2001). Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the motor neurone disease-dementia-aphasia syndrome. *Brain* 124, 103–120.

Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., and Saxe, R. (2008). Concepts are more than percepts: the case of action verbs. *J. Neurosci.* 28, 11347–11353.

Berndt, R. S., Haendiges, A. N., and Burton, M. W. (2002). Grammatical class and imageability in aphasic word production: their effects are independent. *J. Neurolinguistics* 15, 353–371.

Burton, M. W., Krebs-Noble, D., Gullapalli, R. P., and Berndt, R. S. (2009). Functional neuroimaging of grammatical class: ambiguous and unambigious nouns and verbs. *Cogn. Neuropsychol.* 26, 148–171.

Caplan, D. (2001). Functional neuroimaging studies of syntactic processing. *J. Psycholinguist. Res.* 30, 297–320.

Cappelletti, M., Fregni, F., Shapiro, K., Pascual-Leone, A., and Caramazza, A. (2008). Processing nouns and verbs in the left frontal cortex: a TMS study. *J. Cogn. Neurosci.* 20, 707–720.

Caramazza, A., and Hillis, A. E. (1991). Lexical organization of nouns and verbs in the brain. *Nature* 349, 788–790.

Crepaldi, D., Aggujaro, S., Arduino, L. S., Zonca, G., Ghirardi, G., Inzaghi, M. G., Colombo, M., Chierchia, G., and Luzzatti, C. (2006). Noun-verb dissociation in aphasia: the role of imageability and functional locus of the lesion. *Neuropsychologia* 44, 73–89.

Crepaldi, D., Berlingeri, M., Paulesu, E., and Luzzatti, C. (2011). A place for nouns and a place for verbs? A critical review of neurocognitive data on grammatical-class effects. *Brain Lang.* 116, 33–49.

Damasio, A. R., and Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proc. Natl. Acad. Sci. U.S.A.* 90, 4957–4960.

Daniele, A., Giustolisi, L., Silveri, M. C., Colosimo, C., and Gainotti, G. (1994). Evidence for a possible neuroanatomical basis for lexical processing of nouns and verbs. *Neuropsychologia* 32, 1325–1341.

Dronkers, N. F., Pinker, S., and Damasio, A. R. (2000). "Language and the aphasias," in *Principles of Neural Science*, eds E. R. Kandel, J. H. Schwartz, and T. M. Jessell (New York, NY: McGraw-Hill), 1169–1187.

Faroqi-Shah, Y., and Thompson, C. K. (2004). Semantic, lexical, and phonological influences on the production of verb inflections in agrammatic aphasia. *Brain Lang.* 89, 484–498.

Francis, W. N., and Kucera, H. (1982). *Frequency Analysis of English Usage.* Boston: Houghton Mifflin Company.

Friederici, A. D., and Kotz, S. A. (2003). The brain basis of syntactic processes: functional imaging and lesion studies. *Neuroimage* 20, S8–S17.

Gold, B. T., Balota, D. A., Kirchhoff, B. A., and Buckner, R. L. (2005). Common and dissociable activation patterns associated with controlled semantic and phonological processing: evidence from fMRI adaptation. *Cereb. Cortex* 15, 1438–1450.

Goodglass, H. (1973). "Studies on the grammar of aphasics," in *Psycholinguistics and Aphasia*, eds H. Goodglass and S. Blumstein (Baltimore, MD: Johns Hopkins University Press), 183–215.

Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci. (Regul. Ed.)* 10, 14–23.

Henson, R., Shallice, T., and Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science* 287, 1269–1272.

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

Kean, M.-L. (1978). The linguistic interpretation of aphasic syndromes: agrammatism in Broca's aphasia, an example. *Cognition* 5, 9–46.

Khader, P. H., Jost, K., Mertens, M., Bien, S., and Rösler, F. (2010). Neural correlates of generating visual nouns and motor verbs in a minimal phrase context. *Brain Res.* 1318, 122–132.

Laiacona, M., and Caramazza, A. (2004). The noun/verb dissociation in language production: varieties of causes. *Cogn. Neuropsychol.* 21, 103–124.

Liljeström, M., Tarkiainen, A., Parviainen, T., Kujala, J., Numminen, J., Hiltunen, J., Laine, M., and Salmelin, R. (2008). Perceiving and naming actions and objects. *Neuroimage* 41, 1132–1141.

Longe, O., Randall, B., Stamatakis, E. A., and Tyler, L. K. (2007). Grammatical categories in the brain: the role of morphological structure. *Cereb. Cortex* 17, 1812–1820.

Longworth, C. E., Keenan, S. E., Barker, R. A., Marslen-Wilson, W. D., and Tyler, L. K. (2005). The basal ganglia and rule-governed language use: evidence from vascular and degenerative conditions. *Brain* 2005, 584–596.

Luzzatti, C., Aggujaro, S., and Crepaldi, D. (2006). Verb-noun double dissociation in aphasia: theoretical and neuroanatomical foundations. *Cortex* 42, 875–883.

Marslen-Wilson, W., and Tyler, L. K. (1997). Dissociating types of mental computation. *Nature* 387, 592–594.

Martin, A., Wiggs, C. L., and Weisberg, J. (1997). Modulation of human medial temporal lobe activity by form, meaning, and experience. *Hippocampus* 7, 587–593.

McCarthy, R. A., and Warrington, E. K. (1985). Category specificity in an agrammatic patient: the relative impairment of verb retrieval and comprehension. *Neuropsychologia* 23, 709–727.

Menenti, L., Gierhan, S. M. E., Segaert, K., and Hagoort, P. (2011). Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychol. Sci.* 22, 1173–1182.

Miceli, G., Silveri, M. C., Romani, C., and Caramazza, A. (1989). Variation in the pattern of omissions and substitutions of grammatical morphemes in the speech of so-called agrammatic patients. *Brain Lang.* 36, 447–492.

Miceli, G., Silveri, M. C., Villa, G., and Caramazza, A. (1984). On the basis for the agrammatic's difficulty in producing main verbs. *Cortex* 20, 207–220.

Obler, L. K., Harris, K., Meth, M., Centeno, J., and Mathews, P. (1999). The phonology-morphosyntax interface: affixed words in agrammatism. *Brain Lang.* 68, 233–240.

Pai, M. C. (1999). Supplementary motor area aphasia: a case report. *Clin. Neurol. Neurosurg.* 101, 29–32.

Palti, D., Ben Shachar, M., Hendler, T., and Hadar, U. (2007). Neural correlates of semantic and morphological processing of Hebrew nouns and verbs. *Hum. Brain Mapp.* 28, 303–314.

Perani, D., Cappa, S. F., Schnur, T., Tettamanti, M., Collina, S., Rosa, M. M., and Fazio, F. (1999). The neural correlates of noun and verb processing: a PET study. *Brain* 122, 2337–2344.

Rushworth, M. F., Buckley, M., Behrens, T., Walton, M., and Bannerman, D. (2007). Functional organization of the medial frontal cortex. *Curr. Opin. Neurobiol.* 17, 220–227.

Sahin, N. T., Pinker, S., and Halgren, E. (2006). Abstract grammatical processing of nouns and verbs in Broca's area: evidence from fMRI. *Cortex* 42, 540–562.

Shapiro, K., and Caramazza, A. (2003). Grammatical processing of nouns and verbs in left frontal cortex? *Neuropsychologia* 41, 1189–1198.

Shapiro, K., Shelton, J., and Caramazza, A. (2000). Grammatical class in lexical production and morphological processing: evidence from a case of fluent aphasia. *Cogn. Neuropsychol.* 17, 665–682.

Shapiro, K. A., Moo, L. R., and Caramazza, A. (2006). Cortical signatures of noun and verb production. *Proc. Natl. Acad. Sci. U.S.A.* 103, 1644–1649.

Shapiro, K. A., Mottaghy, F. M., Schiller, N. O., Poeppel, T. D., Flüss, M. O., Müller, H.-W., Caramazza, A., and Krause, B. J. (2005). Dissociating neural correlates for nouns and verbs. *Neuroimage* 24, 1058–1067.

Shapiro, K. A., Pascual-Leone, A., Mottaghy, F. M., Gangitano, M., and Caramazza, A. (2001). Grammatical distinctions in the left frontal cortex. *J. Cogn. Neurosci.* 13, 713–720.

Slotnick, S. D., Moo, L. R., Segal, J. B., and Hart, J. Jr. (2003). Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Brain Res. Cogn. Brain Res.* 17, 75–82.

Toni, I., Krams, M., Turner, R., and Passingham, R. E. (1998). The time course of motor sequence learning: a whole-brain fMRI study. *Neuroimage* 8, 50–61.

Tyler, L. K., Bright, P., Fletcher, P., and Stamatakis, E. A. (2004). Neural processing of nouns and verbs: the role of inflectional morphology. *Neuropsychologia* 42, 512–523.

Tyler, L. K., Randall, B., and Stamatakis, E. A. (2008). Cortical differentiation for nouns and verbs depends on grammatical markers. *J. Cogn. Neurosci.* 20, 1381–1389.

Tyler, L. K., Stamatakis, E. A., Post, B., Randall, B., and Marslen-Wilson, W. (2005). Temporal and frontal systems in speech comprehension: an fMRI study of past tense processing. *Neuropsychologia* 43, 1963–1974.

Ullman, M., Corkin, S., Coppola, M., Hickok, G., Growdon, J., Koroshetz, W., and Pinker, S. (1997). A neural dissociation within language: evidence that the mental dictionary is part of declarative memory, and that grammatical rules are part of the procedural system. *J. Cogn. Neurosci.* 9, 266–276.

van Turrennout, M., Bielamowicz, L., and Martin, A. (2003). Modulation of neural activity during object naming: effects of time and practice. *Cereb. Cortex* 13, 381–391.

Wagner, A. D., Koutstall, W., Maril, A., Schacter, D. L., and Buckner, R. L. (2000). Task-specific repetition priming in left inferior prefrontal cortex. *Cereb. Cortex* 10, 1176–1184.

Yokoyama, S., Miyamoto, T., Riera, J., Kim, J., Akitsuki, Y., Iwata, K., Yoshimoto, K., Horie, K., Sato, S., and Kawashima, R. (2006). Cortical mechanisms involved in the processing of verbs: an fMRI study. *J. Cogn. Neurosci.* 18, 1304–1313.

Ziegler, W., Kilian, B., and Deger, K. (1997). The role of the left mesial frontal cortex in fluent speech: evidence from a case of left supplementary motor area hemorrhage. *Neuropsychologia* 35, 1197–1208.

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

Received: 29 March 2011; accepted: 19 January 2012; published online: 08 February 2012.

*Citation: Shapiro KA, Moo LR and Caramazza A (2012) Neural specificity for grammatical operations is revealed by content-independent fMR adaptation. Front. Psychology **3**:26. doi: 10.3389/fpsyg.2012.00026*

This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Psychology.