Investigating the spatial and temporal components of speech production

Greig de Zubicaray[1] & Vitória Piai[2,3]

[1] *Faculty of Health, and Institute of Health and Biomedical Innovation, Queensland University of Technology (QUT), Brisbane, Australia.*

[2] *Donders Institute for Brain, Cognition and Behaviour, Donders Centre for Cognition, Radboud University, Nijmegen, The Netherlands.*
[3] *Department of Medical Psychology, Radboud Universiteit Medical Center, Nijmegen, the Netherlands*

Correspondence: Greig de Zubicaray, Faculty of Health, and Institute of Health and Biomedical Innovation, Queensland University of Technology (QUT), Brisbane, QLD 4059, Australia. Ph. +617 31389639, Email: greig.dezubicaray@qut.edu.au

Abstract

The cognitive processes and neural mechanisms supporting language production have received considerably less research attention than those involved in language comprehension. This situation is partly attributable to the methodological challenges involved in acquiring electrophysiological and neuroimaging data during overt speech. However, our knowledge is increasing with the addition of lesion-symptom mapping studies in patients, and the application of novel brain stimulation technologies. In this chapter, we review findings from a range of studies investigating the spatial and temporal components of speech production in patients and healthy participants, with a particular emphasis on those employing psycholinguistic paradigms to identify and characterize core processes and components of the network involved in retrieving words from the mental lexicon.

Keywords: speech production, picture naming, fMRI, EEG, MEG, lesion symptom mapping, brain stimulation

Introduction

To date, two major behavioural approaches have been utilized in speech production research. The first involves characterizing speech errors or dysfluencies produced either spontaneously, experimentally, and/or as a consequence of brain injury (i.e., aphasia). The second entails experimental investigations of the time-course of speech production in different contexts by means of response time analyses. These psycholinguistic approaches have been complemented recently by a third, involving the application of electrophysiological, neuroimaging, and brain stimulation technologies, following from the realization that we require a better understanding of the neural mechanisms underlying the behavioural data. Theoretical accounts of speech production have also begun to incorporate neural mechanisms in their feature set, although most remain limited to explaining specific effects observed during performance of particular paradigms. However, application of these technologies involves surmounting methodological challenges not typically encountered in studies of language comprehension, largely due to speech-related artefacts introduced by articulation and breathing.

In this chapter, we provide an overview of relevant spatial and temporal components of language production. We then review findings from a range of studies investigating language production in brain-lesion patients and healthy participants, with an emphasis on evidence from context manipulations in picture naming paradigms designed to identify and characterize core processes and components of the network involved in retrieving words from the mental lexicon. Boxes 1 to 4 address the methodological challenges that need to be surmounted by speech production researchers in order to provide valid and reliable data.

Box 1. Methodological challenges for investigations of speech production

From a methodological perspective, neurolinguistic investigations of speech production are inherently more complicated to conduct than those targeting non-language processes and/or comprehension. Two of the most widely employed technologies, functional magnetic resonance imaging (fMRI; see Chapter by Heim and Specht, this volume) and neurophysiological recordings via electroencephalography and magnetoencephalography (EEG and MEG, respectively; see Chapters by Leckey & Federmeier, and Salmelin, Kujala, & Liljeström, this volume) are especially sensitive to speech-related signal artefacts that compromise data quality and confound interpretations. Consequently, the overwhelming majority of studies using these technologies have examined the production of single words or short phrases. The nature of these speech-related artefacts and the appropriate data acquisition and analysis techniques to ameliorate them are briefly described in Boxes 2 and 3.

The earliest, and perhaps most intuitive approach to dealing with speech production artefacts in fMRI and electrophysiological studies was to avoid them entirely. In the first fMRI studies conducted in the 1990s, participants were typically instructed to respond *covertly*, i.e., use *inner* speech. Aside from the obvious problem of having no behavioural data to confirm the presence of a significant psycholinguistic effect, later studies comparing covert and overt speech within the same paradigm revealed engagement of different cerebral regions. However, the relative contributions of those areas are difficult to interpret (for example, greater engagement of regions during overt production might represent syllabification and articulatory-motor processes, while covert production might additionally engage domain general processes such as response inhibition and self-monitoring while reducing activity in production related regions; see the review by Indefrey & Levelt, 2004). Thus, it became clear that avoiding speech-related artefacts by requesting covert responses from participants was not the optimal strategy.

Box 2. Methodological challenges for neuroimaging investigations

In fMRI, the physical act of producing speech introduces blood oxygen level dependent (BOLD) signal changes during continuous imaging, necessitating a different approach to data acquisition. These task-correlated, speech-induced signal changes are not a consequence of neural processes, although they tend to mimic or mask them as they occur over the perisylvian cortex, and are not limited to the speech act itself – they contaminate subsequent images acquired in a functional time series. They also do *not* merely reflect whole-head rigid body motion, and persist following the application of standard image realignment algorithms. However, it is not uncommon to read papers in which authors claim to have inspected their realigned data and confirmed it was free from speech related artefacts. Rather, signal changes from producing even a single word during continuous imaging reflect movement-by-magnetic-susceptibility *interactions*. Areas of magnetic-susceptibility-related signal variation are located in inferior temporal and frontal cortices where air-bone-tissue interfaces exist, and are problematic for both structural and functional MRI. Producing speech worsens these susceptibility-related artefacts by increasing geometric distortions, losing *and* introducing signal - the latter from regions within and immediately outside the field of view linked to muscle movement. A relatively reliable finding is that articulatory muscle movement produces a rapid signal decrease in and around Broca's area and its right hemisphere homologue, compromising any task-related signal changes (see Mehta et al., 2006). Consequently, the BOLD signal in continuous fMRI acquisitions is unavoidably compromised during speech production, and the simultaneous gradient noise often impedes recording and measuring of accuracy and speech onset latencies. It is therefore essential to adopt well-validated, alternative approaches. As we will note later, some of the inconsistencies in reported fMRI findings for speech production paradigms can be directly attributed to inappropriate acquisition methods.

In the early 2000s, a concerted international research program demonstrated the utility of a different image acquisition method for speech production fMRI studies, now broadly referred to as *sparse temporal sampling*. In brief, the technique involves acquiring only a single fMRI volume *after* each utterance, timed to capture the estimated peak BOLD signal response. As speech is produced during relative silence (i.e., without gradient noise associated with image acquisition), both accuracy and latency are able to be measured reliably (see de Zubicaray et al., 2001). Independent validation studies subsequently confirmed the technique is able to acquire artefact-free images during spoken word production (Gracco, Tremblay & Pike, 2005), and showed that neural activity patterns during non-speech paradigms are comparable for sparse and continuous imaging acquisitions (Nebel et al., 2005). Note that the technique requires relatively long inter-trial intervals, and BOLD signal time course data is not available. However, haemodynamic time course information is rarely analysed even in fMRI studies using continuous imaging whose primary purpose is to provide spatial information.

Another recently introduced fMRI acquisition method that shows promise for investigating speech production is arterial spin labelling (ASL), which detects increases in cerebral perfusion associated with task performance. Continuous perfusion imaging acquisitions are demonstrably less sensitive to speech-related susceptibility artefacts than BOLD fMRI (Kemeny, Ye, Birn, & Braun, 2005), and so may prove useful for sentence production paradigms. Although the sensitivity and temporal resolution of ASL are typically lower than BOLD acquisitions, there is less inter-individual variability in perfusion signal changes compared to the BOLD signal, resulting in increased sensitivity to group-level effects (Detre et al., 2012).

Box 3. Methodological challenges for electrophysiological investigations

Until recently, the dominant approach in electrophysiological (here collectively defined as EEG and MEG; see Chapters by Leckey & Federmeier, and Salmelin, Kujala, & Liljeström, this volume) studies of speech production has been to analyse data time-locked to stimulus onset (e.g., a picture to be named) up to an arbitrary point before articulation (e.g., 400 ms; see Ganushchak, Christoffels, & Schiller, 2011). Muscle activity associated with mouth, head and eye movements are well-known sources of artefacts that confound EEG and MEG recordings. Yet, myogenic and speech-related breathing artefacts can in fact precede production by as much as 500-600 ms (e.g., Brooker & Donald, 1980; Galgano & Froud, 2008). A number of approaches to measuring and correcting for speech-related artefacts have since been proposed and adopted, including using electromyogenic (EMG) recordings from the mouth/lip (e.g., Ouyang et al., 2016; Porcaro, Meaglia & Krott, 2015; De Vos et al., 2010). In addition to stimulus-locked analyses, recent studies have also begun analyzing ERPs locked to articulation onset (e.g., Riès et al., 2013). However, it is important to note that response-locked ERPs are sensitive to onset differences due to phonetic features, e.g., /b/ and /p/ onset words vary by about 40 ms (Fargier et al., 2017), so balanced designs are essential.

Additional sources of variability are encountered in intracranial EEG studies of speech production (for a review, see Llorens et al., 2011). These studies are typically conducted in candidates for epilepsy surgery who have intractable seizures. This necessitates comprehensive language assessments for each patient to characterize their cognitive function, and investigations to be performed in medication- and seizure-free periods. Due to clinical constraints, electrode placement is often inconsistent across patients within a study, a problem referred to as *sparse sampling* (not to be confused with the fMRI acquisition technique).

As Piai, Riès, and Knight (2015) noted in their commentary, a major issue in the electrophysiological literature on speech production has been variability in reporting analysis steps and suboptimal statistical testing. In particular, a number of studies appear to lack procedures for correcting speech-related artefacts. Moreover, electrophysiological data is multidimensional (i.e., many channels can be analysed at multiple time points), which requires analyses strategies that take into account this multiple-comparisons problem. However, a number of studies have adopted inappropriate statistical methods for dealing with this problem (see Piai, Dahlslätt, & Maris, 2015). These factors have likely contributed to the lack of consistency among some of the reported findings we discuss later.

Box 4. Methodological challenges for brain stimulation techniques

Brain stimulation methods have also seen increasing application to investigate speech production. Non-invasive techniques include transcranial magnetic stimulation (TMS; see Chapter by Schuhmann, this volume) and transcranial direct current stimulation (tDCS; for review, see Hartwigsen, 2014), while direct electrical stimulation (DES) is invasive and most frequently used intraoperatively for surgical planning (see Chapter by Duffau, this volume). Brain stimulation methods are useful for studies of language production because researchers can modulate brain activity, and even temporarily disrupt the function of a targeted region (see Chapter by Schuhmann, this volume), and examine what impact this has on task performance. Brain stimulation studies therefore provide an additional dimension of information regarding the necessity of a brain region for task performance. For both tDCS and TMS, effective sham approaches are able to blind participants to the stimulation conditions. Like TMS, online and offline stimulation protocols are possible with tDCS. However, it has been proposed that they differ in terms of their effects on neural activity: the former is thought to modulate a specific network involved in task performance, while the latter likely produces modifications to a broader network of neural activity that lasts beyond the period of stimulation (Miniussi, Harris, & Ruzzoli, 2013). The latter approach may therefore be less relevant to the study of psycholinguistic effects. The two modes of stimulation present a challenge in terms of reporting and interpreting effects in neurolinguistic studies. In particular, findings from offline and online studies should not be discussed synonymously in terms of replications. Ideally, a comparison of offline and online protocols in the one study would be most informative.

Of the three techniques, tDCS is less well-established from a methodological perspective, and the validation of specific protocols in both healthy participants and aphasic patients requires further research attention. Inconsistencies in reported findings appear attributable to variability in protocols applied across studies, including departures from standard experimental paradigms (e.g., Westwood et al., 2017; see Gauvin, Meinzer & de Zubicaray, 2017 for comment). The tDCS technique involves applying a weak electrical current between scalp-affixed electrode pads to modulate resting membrane potentials on the cortical surface. These pads are typically 5 x 7 cm, although high density electrode sets are now being marketed by some manufacturers. Localisation of effects in behavioural studies is therefore inexact, dependent on electrode placement. Combined tDCS-fMRI studies show promise in terms of localizing effects beneath electrodes (e.g., Meinzer et al., 2012). Unlike TMS, tDCS is not able to provide information about the timing of functionally relevant brain regions. Likewise, in the case of DES, the clinical setting does not easily afford timing investigations given that the stimulation cannot be applied systematically at different time points during stimulus presentation.

Spatiotemporal components relating to processing stages in speech production models

Based upon speech error and response time data, models of spoken word production have stipulated that the processing stages involved in producing words include *perceptual and conceptual preparation, lexical selection, word-form encoding*, and *articulation*. Indefrey and Levelt's (2000, 2004, henceforth I&L) influential review and meta-analysis represented the first serious attempt to relate neuroimaging and electrophysiological data to these processing stages. In 2004, the EEG and MEG literature on picture naming and word production was scant compared to neuroimaging studies (fMRI and PET), and the available data reflected a mix of covert and overt production. In 2011, Indefrey reviewed the literature for new studies and provided an updated version of the time course of these processing stages and the corresponding brain regions. According to this updated version, and assuming a response time of 600 ms, perceptual and conceptual preparation reliably involve the ventral surface of temporal-occipital cortex. Perceptual preparation (given a visual or auditory stimulus) is completed around 100-150 ms after stimulus onset, whereas conceptual preparation is completed around 200 ms. The mid-section of the lateral middle temporal cortex is reliably involved in lexical-semantic (*lemma*) processing and the posterior superior and middle temporal cortex (Wernicke's area) in phonological word form (*lexeme*) retrieval. These processing stages span between 200-290 ms and 290-370 ms, respectively. Correlates of post-lexical stages of processing, including syllabification at 355-475 ms, phonetic encoding at 455-600 ms, followed by motor articulation were found to encompass both left inferior frontal gyrus (IFG) and premotor cortices (for a review of these latter mechanisms, see the Chapter by Tremblay, Deschamps & Dick, this volume). Finally, I&L (2004) ascribed the monitoring of both internal and external speech to bilateral superior temporal gyrus, noting the prolonged activation observed in these structures. However, Indefrey (2011) later acknowledged self-monitoring processes likely engage a more comprehensive network involving the anterior cingulate cortex (ACC) and supplementary motor area (SMA), perhaps reflecting more domain general mechanisms, an issue we will return to later.

This account of the relative timing of brain regions engaged during production has been challenged by recent reviews. For example, Llorens et al. (2011) reviewed intracranial EEG studies, noting the findings were not always consistent with the earlier scalp recorded EEG and MEG findings. In particular, they did not find consistent evidence for the involvement of the mid portion of the middle temporal gyrus (MTG) linked with the time window for lexical-semantic processing, and noted much earlier engagement of motor cortex and IFG (from 200 ms), challenging an interpretation of the latter regions' roles in terms of only post-lexical processes. Recently, Munding, Dubarry, and Alario (2016) reviewed 20 years of MEG research on speech production, and concluded that while the data broadly supported the conclusions of I&L's meta-analyses, the evidence was not consistent with a serial processing architecture, but rather suggested parallel activation. In addition, they presented evidence of very early activation of IFG and motor cortex in some studies, which they attributed to top-down control and selection processes. This review provoked considerable debate (see responses by Piai, 2016; Riès, 2016; Strijkers, 2016). A relatively consistent criticism was that Munding and colleagues reviewed studies with different experimental manipulations that elicited much longer production latencies than the 600 ms assumed by I&L, blurring the temporal estimates for the processes in question. Therefore, the study by Munding et al. (2016) is inconclusive with respect to the question of serial versus parallel processing in spoken word production. Another recent critique by Strijkers and Costa (2016) likewise advocated parallel over serial processing and invoked top-down selection and control processes within the framework of a neural assembly rather than chronometric model (but see the response by Indefrey, 2016).

Despite the emphasis on electrophysiological recordings to inform temporal components of speech production, research with single and multiple pulse TMS has afforded additional information about the *critical* or *necessary* timing of the brain regions in the production network. A number of studies have consistently targeted left IFG/Broca's area and employed picture naming (e.g., Chouinard et al., 2009; Schuhmann et al., 2009, 2012; Wheat et al., 2013). Across these studies, TMS reliably slowed naming latencies when applied to the left IFG 300-400 ms following picture presentation, consistent with Indefrey's (2011) updated account. Critically, when TMS was applied to left IFG at 100-200 ms following picture onset (i.e., the time window of early activation reported in some MEG studies; see Munding et al., 2016; Strijkers & Costa, 2016), naming latencies were *not* significantly affected (Schuhmann et al., 2009, 2012; Wheat et al., 2013). Therefore, if early activation of IFG is attributable to top-down selection and control mechanisms as some authors propose, then the consistent finding from TMS studies does not support the necessary early involvement of these processes in simple picture naming. The same can be said for the early involvement of motor cortex based upon some MEG results (e.g., Munding et al., 2016; Strijkers & Costa, 2016), with two TMS studies consistently failing to observe effects of stimulation of the motor cortex on naming latencies at 0, 100 or 300 ms after picture onset (Mottaghy, Sparing & Töpper, 2006; Töpper et al., 1998).

To our knowledge, three TMS studies of picture naming have targeted Wernicke's area at the 200-300 ms time period identified by Indefrey (2011) as corresponding to phonological word form retrieval (Mottaghy et al., 2006; Schuhmann et al., 2012; Töpper et al., 1998). Surprisingly, *none* reported a significant effect on naming latencies, although one observed an effect at 400 ms post picture onset, interpreted as being consistent with self-monitoring

(Schuhmann et al., 2012). Note that these studies, while problematic for I&L's account, likewise showed that earlier activation of posterior STG is not necessary for successful picture naming, as TMS did not produce an effect at 0, 100, or 200 ms (cf., Munding et al., 2016; Strijkers & Costa, 2016). Although it is tempting to conclude that word form retrieval might occur later than estimated by both chronometric and electrophysiological methods, it is important to note that all three studies applied TMS solely to the left posterior STG (Brodmann area 22), whereas I&L's meta-analysis attributed word form processing to cortical tissue encompassing the posterior sections of *both* left MTG and STG. Schuhmann et al.'s (2012) study was the only one to apply TMS to the mid portion of the MTG, finding an effect at 225 ms post picture onset, consistent with I&L's attribution of lexical-semantic processing. Consequently, mid and posterior sections of MTG should be targets for future TMS investigations concerning the time course of production.

Alternatively, I&L's attribution of posterior MTG/STG might be wrong, and a different region could subserve phonological word form retrieval. For example, using lesion-symptom mapping (LSM; see Chapter by Wilson, this volume), Schwartz et al. (2012) and Dell et al. (2013) analysed speech error data from large cohorts of aphasic patients and implicated a region *adjacent* to Wernicke's area, the left inferior parietal lobe (IPL, comprising supramarginal and angular gyri, and planum temporale), in phonological/phonemic errors during picture naming. They interpreted their findings as consistent with a dorsal stream, articulatory-based account of phonological processing in production (see also Chapter by Hickok, this volume). However, a reanalysis of Schwartz et al.'s (2012) data examining phonological neighborhood density effects (as lexical phonological processing involves activating not only the target word but also phonologically related words) implicated both posterior STG and IPL (Mirman & Graziano, 2013). Both regions were also reported in a sparse design fMRI study of phonological neighbourhood effects in production (Peramunage et al., 2010). Of note, other lesion symptom mapping studies have reported left posterior MTG/STG rather than IPL, consistent with I&L's account (e.g., Butler, Lambon Ralph, & Woollams, 2014). Future investigations using more fine-grained analyses of aphasics' speech errors differentiating lexical versus post-lexical phonological representations are needed (see Goldrick & Rapp, 2007; e.g., Schwartz et al., 2012 analysed *all* nonword errors, regardless of their phonological overlap with the target).

Investigating the spatiotemporal components of speech production models with psycholinguistic paradigms

In this next section, we summarise and discuss some key findings from neurolinguistic investigations using context manipulations in picture naming paradigms to characterise processing stages in speech production. As the spatiotemporal components corresponding to representational stages in production models are relatively well-delineated (see above), neurolinguistic investigations are informative for testing rival accounts that attribute manipulations of an experimental factor to different processing stages (see de Zubicaray, 2012). However, context manipulations influencing the time course of production in chronometric studies likewise influence the timing of evoked responses in electrophysiological recordings, so should not be used for estimating temporal signatures of specific processing stages during simple production tasks (see Piai et al., 2015; cf. Munding et al., 2016). Rather, they are useful for demonstrating the correlates of the cognitive effect of interest. Here, we constrain our review to paradigms and context effects for which results from multiple sources of evidence are available, i.e., across functional neuroimaging, electrophysiological, lesion-mapping and/or brain simulation methods. We interpret only reliably reported findings across multiple studies.

Picture word interference

Over 4 decades, the picture word interference (PWI) paradigm has been used to test hypotheses about processing stages involved in spoken word production (Rosinski, Golinkoff, & Kukish, 1975). In brief, the PWI paradigm involves participants naming a set of target pictures in context with written or auditorily presented distractor words. The most commonly employed manipulations in neurolinguistic studies involve distractors either semantically or phonologically related to the target picture name, compared with unrelated words. The *semantic interference* effect refers to the finding that naming latencies are significantly slower for pictures paired with categorically related versus unrelated words (e.g., a pictured COW with related distractor "*pig*" versus unrelated distractor "*pin*"). Conversely, phonologically related distractors (i.e., showing phoneme overlap with the picture name, e.g., a picture COW with related distractor "*couch*") have been shown to facilitate (speed) target naming compared to unrelated distractors (e.g., "*pin*"). Both effects occur reliably within a small range of distractor stimulus onset asynchronies (SOAs) around target picture presentation (-150 to 150 ms, i.e., the distractor word is presented up to 150 ms before the picture or after picture presentation), with simultaneous distractor-target presentation usually employed. Virtually all production accounts attribute the locus of the phonological facilitation effect to the word form retrieval stage of processing. However, the locus of the semantic interference effect is a matter of debate between rival accounts. According to the lexical-selection-by-competition account, lemmas compete for selection and semantically related distractors increase this competition, delaying the selection process, which surfaces as longer naming latencies for the related condition (e.g., Piai et al., 2014). Conversely, according to post-lexical accounts, the semantic interference effect emerges in later stages, closer to articulation onset (see Mahon et al., 2007). In addition, recent accounts have increasingly proposed involvement of top-down selection and control processes for resolving competition among lexical candidates.

Our review identified over 20 neurolinguistic studies of semantic interference and phonological facilitation using the PWI paradigm with EEG, MEG, fMRI, tDCS and lesion patients. Of note, there is considerable variability in experimental designs across studies, including departures from the standard PWI procedure. A critical design feature of the PWI paradigm is that the same set of target pictures is used in each condition, with each picture paired with a different distractor word, ensuring target processing is identical. Distractors are matched on a range of lexical variables, and the same words are often re-paired with different pictures to create the unrelated condition. Distractors may also be target picture names, i.e., members of the response set. Some studies include another condition, e.g., picture naming in the absence of a distractor, or the picture name itself, or a neutral distractor condition (e.g., a row of Xs, nonwords or symbols). Here, we restrict our review to canonical (i.e., related versus unrelated) contrasts for the semantic interference and phonological facilitation effects of interest. For both effects, we report differential activity associated with related compared to unrelated distractors, i.e., mean activity increases for related versus unrelated distractors, and vice versa.

Seven fMRI, two lesion, one tDCS and one MEG studies have provided spatial location/source information for the semantic interference effect, all except two of which employed written distractors (Abel et al., 2009, 2012; de Zubicaray & McMahon, 2009). Of the fMRI studies, four departed from the traditional PWI design by using unique sets of target pictures per distractor condition or a distractor SOA of 550 ms (i.e., distractor presented after picture onset), and all used continuous rather than optimal sparse imaging designs (Abel et al., 2009, 2012; Diaz et al., 2014; Rizio, Moyer, & Diaz, 2017; Spalek & Thompson-Schill, 2008). Diaz et al. (2014) failed to observe any significant activity for the contrast of related versus unrelated distractor conditions in either whole brain or ROI analyses. Abel et al.'s (2012) reanalysis of their 2009 study data reported significantly increased activity in left IFG and reduced activity in left lingual and bilateral precentral gyri, left ACC, posterior STG, parietal operculum and bilateral cuneus. Conversely, Rizio et al. observed only significantly increased activity in the bilateral middle frontal gyrus (MFG), bilateral MTG, and precuneus. Spalek and Thompson-Schill (2008) did not observe a significant semantic interference effect in naming latencies (likely due to the late SOA employed), and reported increased BOLD signal in the right posterior cerebellum and left fusiform and parahippocampal gyri. The inconsistent results across the four studies might therefore reflect the use of unique picture sets in each condition, written versus auditory distractors, a late distractor SOA, speech-related artefacts in continuous imaging, or a combination of all these issues.

More consistent findings are provided by the one sparse and one continuous event-related fMRI and one MEG studies using the standard PWI design with written distractors (de Zubicaray et al., 2013; Piai et al., 2013, 2014). A consistent finding across all studies is *reduced* activity in left mid-to-posterior MTG and STG for related versus unrelated distractors. This pattern was also found in another sparse design fMRI study using auditory distractors (de Zubicaray & McMahon, 2009). Piai and Knight (2017; see also Piai, Riès, & Swick, 2016) also reported increased semantic interference error rates for patients with left lateral temporal cortex lesions, but not for patients with IFG lesions. However, Henseler et al. (2014) failed to observe any modulation of semantic interference with tDCS applied to the left posterior MTG or IFG. Notably, across studies, semantic interference was not consistently associated with involvement of cortical regions attributed to post-lexical processes or mechanisms proposed for resolving competition among semantic competitors, e.g., premotor cortex or IFG. Thus, the available data support a lexical-level rather than post-lexical account of semantic interference in PWI (cf., Mahon et al., 2007).

Of the six EEG and one MEG studies attempting to characterise event-related responses (i.e., potentials, ERPs, and fields, ERFs) to semantic interference using stimulus-locked analyses, one departed from the typical PWI design by blocking (and cycling) category members (Aristei et al., 2011) while two others failed to observe any significant differences between related and unrelated ERPs (Hirschfeld et al., 2008; Piai, Roelofs, & van der Meij, 2012). Across the remaining four studies, semantic interference consistently modulated the event-related responses in the N400 time window (around 250-550 ms), showing a more negative-going waveform or smaller amplitude for related compared to unrelated distractors (i.e., N200-400; Blackford et al., 2012; Dell'Acqua et al., 2010; Piai et al., 2014; Zhu, Damian & Zhang, 2015). Given the relatively longer naming latencies observed in picture-word interference, the electrophysiological data are also consistent with a lexical-level account of the semantic interference effect.

Phonological facilitation in PWI has been investigated by five fMRI studies. Two of these used the standard paradigm, each with written and auditory distractors in sparse-designs (de Zubicaray et al., 2002; de Zubicaray & McMahon, 2009). Three further fMRI studies departed from the standard PWI design and used continuous imaging. Surprisingly, all three of these latter studies failed to observe significant differences in naming latencies, i.e., the experimental manipulation did not produce the phonological facilitation effect of interest (Abel et al., 2009, 2012; Diaz et al., 2014; Rizio et al., 2017). This might reflect the use of unique picture sets or gradient noise associated with continuous imaging that prevents accurate measurements of response latency even with noise-cancelling techniques (although Abel and colleagues also failed to observe a significant difference in an identical behavioural experiment conducted outside the scanner). Diaz et al. (2014) reported increased activation for phonologically related versus unrelated distractors in bilateral supramarginal and angular gyri, whereas Rizio et al. (2017) reported signal increases in the right angular gyrus and left superior parietal cortex, and Abel et al. reported increased activity in left supramarginal gyrus and inferior parietal lobule (BA 40) accompanied by signal reductions in bilateral visual cortex, left ACC and parahippocampal gyrus. Ignoring the absence of evidence for a significant phonological effect in the behavioural data across studies, the fMRI data could be interpreted as reflecting relatively consistent involvement of the left IPL, which might be considered to support a dorsal stream, articulatory-based account of phonological processing in production (e.g., Dell et al., 2013; Schwartz et al., 2012; but see below).

Across the two fMRI studies using sparse event related designs, BOLD signal decreases were observed consistently in the left posterior MTG/STG (de Zubicaray et al., 2002; de Zubicaray & McMahon, 2009). Additional evidence for left posterior temporal cortex involvement is provided by Pisoni and colleagues (2017) using anodal tDCS that significantly reduced the magnitude of the facilitation effect. The evidence for left IFG involvement is equivocal. Piai et al. (2016) reported an increased facilitation effect in left IFG lesion patients compared to healthy controls. Significantly reduced BOLD signal responses in left IFG were reported in one auditory PWI study (de Zubicaray & McMahon, 2009). However, another fMRI study did not observe differential left IFG activity for the phonological effect (de Zubicaray et al., 2002), and Pisoni et al. (2017) reported tDCS applied to left IFG did not influence the facilitation effect. Of the four EEG studies to investigate event-related responses to phonological facilitation in PWI, two failed to observe any significant differences in stimulus-locked waveforms (Blackford et al., 2012; Bürki, 2017), and the remaining two reported different results, ranging from a less negative-going waveform between 250 and 450 ms (Dell'Acqua et al., 2010) to a more positive-going waveform between 450-600 ms (Zhu et al., 2015; in Chinese).

Figure 1 provides a summary of the spatiotemporal components associated with the semantic (left) and phonological (right) effects in picture-word interference. Both significant and non-significant effects are shown for studies using the standard paradigm. Each study is represented by a circle, which is colour-coded according to the method employed. The most consistent pattern for the semantic effect seems to be decreased brain activity in the left posterior temporal lobe (i.e., related < unrelated), in a window ranging around 250-500 ms. For the phonological effect, decreased brain activity in the posterior temporal lobe seems to be the only reproducible pattern, although not as consistent as for the semantic effect.

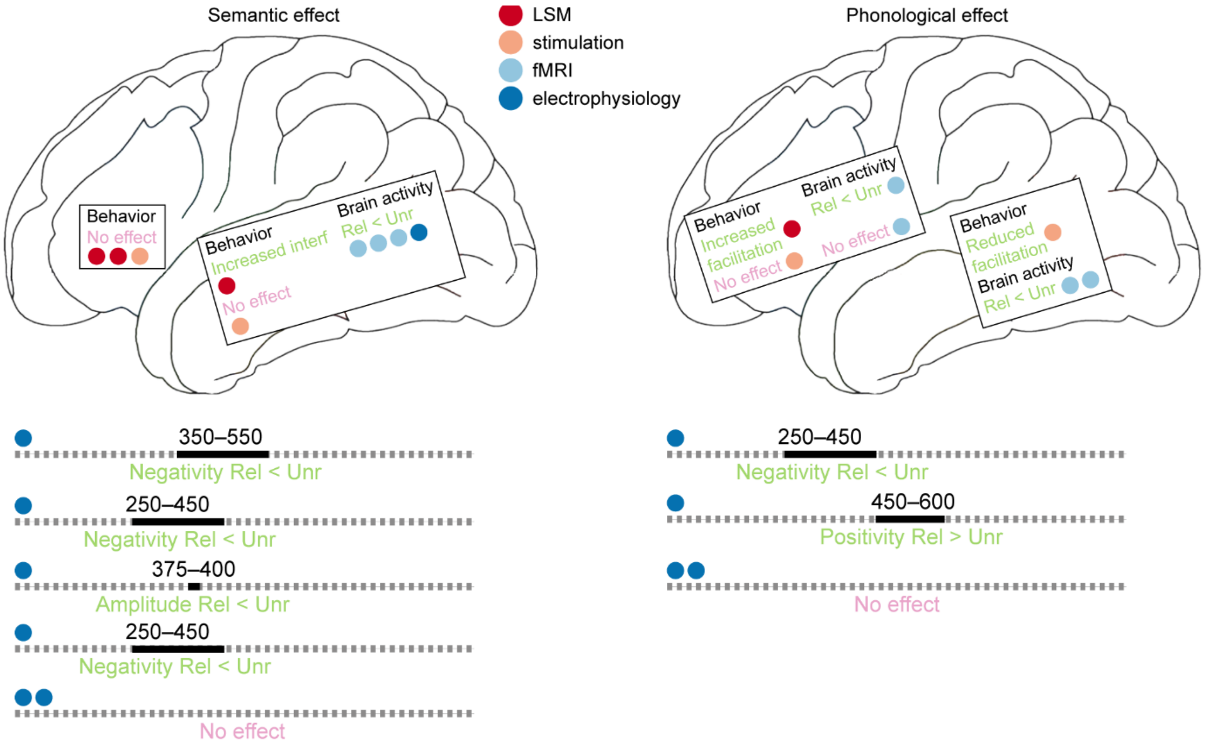

Figure 1. Schematic view of the evidence on the spatial (for the left mid-to-posterior superior and middle temporal gyri and left inferior frontal gyrus) and temporal components of semantic and phonological context effects in picture-word interference. Only studies using the standard paradigm are shown. Each method is colour coded (orange: fMRI; yellow: electrophysiology; purple: lesion-symptom mapping, LSM; blue: non-invasive brain stimulation). Each coloured circle represents one study. Interf = interference; rel = related; unr = unrelated.

Blocked cyclic naming

Another production paradigm with a relatively large neurolinguistic evidence-base is blocked cyclic naming. The most frequently implemented version of the paradigm entails a small set of pictures that participants name repeatedly over several cycles, alternating between categorically related (e.g., animals) versus unrelated contexts (animals, musical instruments, vehicles, vegetables). Like the PWI paradigm, categorically related contexts elicit a semantic interference effect, typically observed only from the second cycle onward (for a review, see Belke & Stielow, 2013). Theoretical accounts have been devised to specifically explain semantic interference in this paradigm, with several incorporating evidence from lesion, neuroimaging and non-invasive brain stimulation studies. All of these accounts propose roles for both the left pMTG/STG and left IFG, in lexical-semantic processing and various control/selection-biasing mechanisms, respectively (e.g., Belke & Stielow, 2013; Oppenheim et al., 2010; Schnur et al., 2009). An incremental learning mechanism has also been proposed to be responsible for the persistence of interference across cycles but has received less research attention in terms of its neural correlates (Damian & Als, 2005; Oppenheim et al., 2010).

Neuroimaging, electrophysiological, brain stimulation and LSM studies of semantic interference in blocked cyclic naming mostly provide evidence for involvement of left pMTG/STG. However, there is little consistency with respect to the direction of the effect across studies. Two early fMRI studies involved non-standard design and analysis manipulations, making results difficult to interpret. For example, Hocking et al. (2009) manipulated visual feature overlap across blocks and were unable to record naming latencies during perfusion fMRI, while Schnur et al. (2009) reported findings for a contrast of semantic interference versus phonological facilitation (i.e., [semantic > unrelated] > (phonological > unrelated]). Both studies reported signal increases. One perfusion fMRI experiment using the standard design reported signal reductions (de Zubicaray et al., 2014). Two tDCS studies reported opposite effects for online and offline stimulation protocols (Meinzer et al., 2016; Pisoni et al., 2012). One LSM study reported a significant increase in error rates (Harvey & Schnur, 2015). Interestingly the peak spatial coordinates reported by Harvey and Schnur's (2015) LSM and de Zubicaray et al.'s (2014) perfusion fMRI studies in posterior MTG/STG were virtually identical. An MEG study likewise reported differential MTG/STG activity (Maess et al., 2002; note the direction of the effect could not be interpreted given the principal component analysis approach). An intracranial EEG study reported decreases in the evoked responses in the MTG and STG, but increases in the inferior temporal lobe (Riès et al., 2017). Finally, a TMS study by Krieger-Redwood and Jefferies (2014) reported an effect of stimulation over pMTG/STG only in the first cycle of naming, i.e., prior to the emergence of the interference effect.

The evidence from a similar range of studies is mostly consistent with left IFG involvement, although there are some notable exceptions. Whereas only one of three fMRI studies reported significant IFG activity using a non-standard comparison (Schnur et al., 2009; cf. de Zubicaray et al., 2014; Hocking et al., 2009), two of three studies of aphasics with IFG lesions noted significant effects in error rates (Riès et al., 2015; Schnur et al., 2009; cf. Harvey & Schnur, 2015). Two of three tDCS studies reported an effect of stimulation to left IFG reducing the magnitude of the interference effect using offline (Pisoni et al., 2012) and online stimulation protocols (Meinzer et al., 2016; cf., Westwood et al., 2017), although one

TMS study observed an effect only in the first cycle of naming (i.e., prior to semantic interference occurring; Krieger-Redwood & Jefferies, 2014). An additional tDCS reported a significant reduction in the magnitude of semantic interference following online but not offline stimulation over the dorsal frontal cortex (Wirth et al., 2011).

Unfortunately, EEG and MEG estimates of stimulus-locked event-related responses for the semantic interference effect vary considerably both in terms of timing and polarity of waveforms: from 150-225 ms (Maess et al., 2002), 200-500 ms (Wang, Shao, Chen & Schiller, 2017; smaller negativity; in Chinese), 220-450 ms (Janssen, Carreiras & Barber, 2011; smaller negativity), 270-315 ms (Python, Farghier, & Laganaro, 2017; smaller positivity), to 500-750 ms (Janssen et al., 2014; smaller positivity). Two other electrophysiological studies either used a non-standard design combining PWI and blocking (Aristei et al., 2011) or failed to observe any significant differences in event-related responses (Llorens et al., 2014). The considerable variability in analysis techniques across studies might explain the inconsistent findings, as might the authors' choice of interpretations for event-related responses in particular time windows. With respect to the latter, some studies reported more than one time window for event-related responses, permitting some flexibility in interpretation. For example, Maess et al. (2002) reported a second evoked response around 450-475 ms interpreted as self-monitoring, whereas Janssen et al. (2014) reported an effect in an earlier 250-400 ms time window they interpreted in terms of conceptual processing.

Substantial evidence in animals and humans implicates the hippocampus in both implicit (unconscious) and explicit retrieval of relational information (see Duss et al., 2014; for review see Henke, 2010), making it a plausible candidate for the implicit incremental learning mechanism proposed in psycholinguistic accounts of semantic interference in blocked cyclic naming (e.g., Damian & Als, 2005; Oppenheim et al., 2010). One perfusion fMRI and one intracranial EEG studies explicitly targeted hippocampal activity using the standard design, interpreting their findings as reflecting the operation of an incremental learning mechanism (de Zubicaray et al., 2014; Llorens et al., 2016). The perfusion fMRI study revealed a reduction in activity during the related context. Using intracranial electrodes implanted directly in bilateral hippocampus, Llorens et al. (2016) reported that the amplitude of the event-related responses (a negativity peaking around 600 ms) in the related blocks was smaller than in unrelated blocks. Crucially, this negative peak emerged progressively from the second cycle of naming onwards.

Figure 2 provides a summary of the spatiotemporal components associated with the semantic effect in blocked cyclic naming. Both significant and non-significant effects are shown for the studies using the standard paradigm. Each study is represented by a circle, colour-coded according to the method employed. The most consistent pattern for the semantic effect seems to be an impact on the magnitude of the interference depending on whether the LIFG is stimulated non-invasively or damaged. Decreased brain activity in the left posterior temporal lobe (i.e., related < unrelated) is also relatively consistent, and again brain stimulation or damage in that area impact the magnitude of the interference effect. Regarding the temporal component, modulations in the 200-450 ms range seem to be the most consistent pattern, with the homogeneous condition showing decreased amplitude relative to the heterogeneous condition, although as noted above, authors tend to disagree on the interpretation of that modulation.

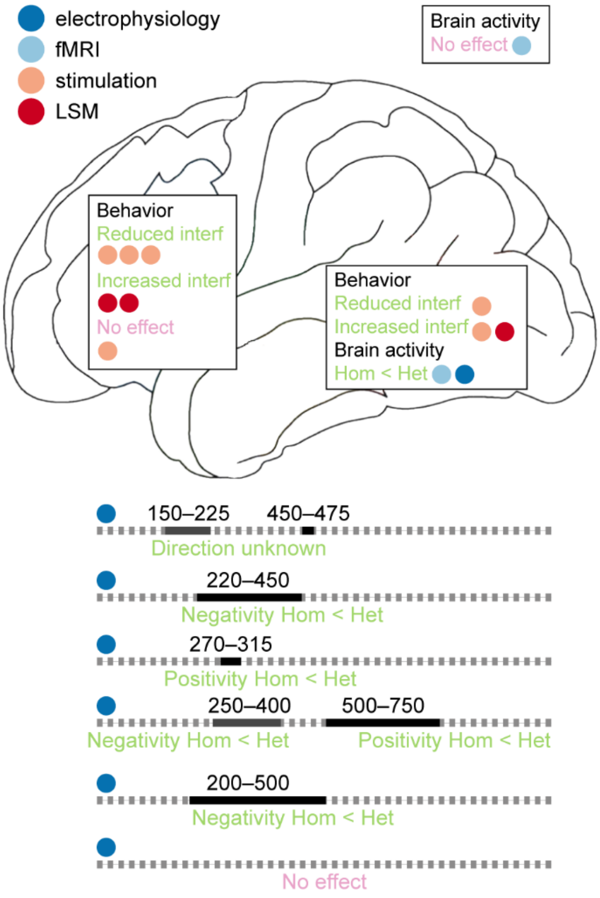

Figure 2. Schematic view of the evidence on the spatial (for the left superior and middle temporal gyri and left inferior frontal gyrus) and temporal components of the semantic context effect in blocked cyclic naming. Only studies using the standard paradigm are shown. Each method is colour-coded (orange: fMRI; yellow: electrophysiology; purple: lesion-symptom mapping, LSM; blue: non-invasive brain stimulation). Each coloured circle represents one study. Het = heterogeneous; hom = homogeneous; interf = interference.

Continuous naming

The continuous naming paradigm introduced by Howard et al. (2006) requires participants to name a pseudorandom series of pictures, and likewise elicits a semantic interference effect. Within the series, exemplars from a range of semantic categories are interspersed with filler items. The interval or lag between each consecutive category exemplar is also varied. Semantic interference in this paradigm manifests from the second ordinal position within a category and accumulates linearly at ~30 ms for each successive categorically related picture. As with the semantic interference effects in PWI and blocked cyclic naming tasks, the left lateral temporal lobe is proposed to play a prominent role in lexical-semantic retrieval in continuous naming (e.g., Belke, 2013). These accounts also tentatively ascribe a role for the left IFG in either top-down, selection-biasing or activation boosting mechanisms (Belke, 2013; Canini et al., 2016; Oppenheim et al., 2010). Further, at least one account presumes semantic interference in the continuous and blocked cyclic naming arises due to a common mechanism(s) (Oppenheim et al., 2010), leading to the expectation that the spatiotemporal mechanisms will bear at least a strong resemblance across paradigms. A cumulative interference effect for phonologically related words has also been reported in continuous reading aloud (Mulatti et al., 2012), but is yet to be subjected to neurolinguistic investigation.

Despite being introduced only a decade earlier, the continuous naming paradigm has been the subject of three fMRI studies, as many EEG studies, one lesion and one tDCS investigation. Of the three fMRI studies, two employed continuous BOLD acquisitions (Canini et al., 2016; Wilson et al., 2009) while the third used perfusion imaging (de Zubicaray et al., 2015). Wilson et al. (2009) employed Howard et al.'s (2006) stimuli, yet were unable to detect any significant BOLD signal correlates of cumulative interference. Canini et al.'s (2016) design departed from Howard et al.'s by presenting participants with two different experimental lists, averaging the ordinal position data. They were unable to report naming latency data due to the gradient noise accompanying continuous imaging, and did not find evidence for cumulative interference in error rates. Their parametric fMRI analysis revealed a linear increase in BOLD signal in the left IFG and caudate. However, this analysis included the first ordinal position data whereas the cumulative effect is calculated from the second ordinal position onward. Using perfusion imaging with the original Howard et al. experimental lists, de Zubicaray et al. (2015) reported a significant linear increase in left mid-MTG and perirhinal cortex activity from the second ordinal position onward. However, Westwood et al. (2017) reported that tDCS to left posterior MTG did not modulate the interference effect compared to sham (but see Gauvin et al., 2017 for a critique of Westwood et al.'s methods). Thus, the evidence for left lateral temporal lobe involvement in cumulative semantic interference is best described as inconsistent and in need of further investigation.

The findings from lesion, EEG and tDCS studies are more consistent with respect to the left IFG. Riès et al. (2015) failed to observe an effect of left IFG lesions on cumulative interference in either naming latencies or error rates. Westwood et al. (2017) reported that tDCS to left IFG did not modulate the interference effect compared to sham (but see Gauvin et al., 2017). Using a non-standard design omitting the lag manipulation, one study failed to observe any differences associated with cumulative interference (Llorens et al., 2016). Both Costa et al. (2009) and Rose and Abdel Rahman (2016) reported a linear modulation of positive waveforms at around 200-400 ms post picture onset over only posterior electrodes, in addition to significant correlations with naming latencies between 208 and 388 ms, and 268 and 413 ms, respectively, noting their findings were broadly consistent with the time window suggested for lexical selection (see spatial and temporal components section above).

Figure 3 provides a summary of the spatiotemporal components associated with the cumulative semantic effect in continuous naming. Both significant and non-significant effects are shown for the studies using the standard paradigm. Each study is represented by a circle, colour-coded according to the method employed. The paucity of studies contributes to an impression of a lack of consistency in findings. In the temporal domain, a linear modulation of positive-going waveforms in the 250-400 range seems to be the consistent finding over two studies.

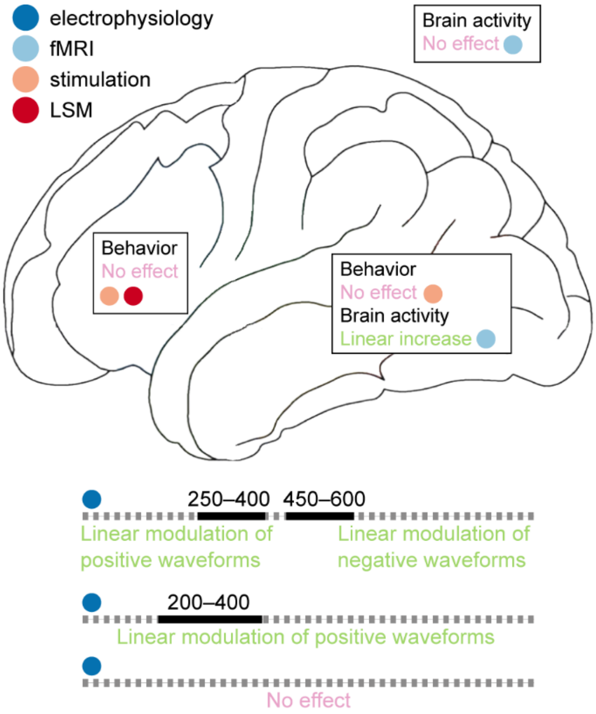

Figure 3. Schematic view of the evidence on the spatial (for the left middle temporal gyrus and left inferior frontal gyrus) and temporal components of the cumulative semantic effect in continuous naming. Only studies using the standard paradigm are shown. Each method is colour-coded (orange: fMRI; yellow: electrophysiology; purple: lesion-symptom mapping, LSM; blue: non-invasive brain stimulation). Each coloured circle represents one study.

Are top-down control/selection biasing mechanisms really required to resolve semantic interference?

Our review of findings from the PWI, blocked cyclic and continuous naming paradigms is informative both from a neurolinguistic perspective and for constraining cognitive accounts of production. All three paradigms were designed to promote activation of multiple lexical candidates, and observations of behavioural semantic interference effects across paradigms are typically interpreted in terms of common mechanisms. Recent psycholinguistic accounts of all three paradigms have begun to incorporate neural data, with a particular emphasis on intervention by domain general, top-down cognitive control mechanisms to resolve competition during production. For the most part, these accounts have proposed a prominent role for the left IFG in resolving competition among lexical-semantic competitors. Indeed, Belke and Stielow (2013) concluded “It appears that any future model of word production unavoidably faces the challenge of specifying how left frontal mechanisms of domain-general cognitive control interact with paradigmatic interference during lexical-semantic encoding.” (p. 23). Based upon the evidence we reviewed above, this conclusion appears to be a significant over-statement.

We could find no reliable evidence for a role of left IFG in semantic interference in either PWI or continuous naming paradigms. Although absence of evidence might not be evidence of absence, the fact that left IFG involvement was observed only semi-reliably in blocked cyclic naming studies of semantic interference suggests the need for a re-assessment of proposals concerning the ubiquitous involvement of domain general, top-down mechanisms in biasing or resolving competition during spoken word production (subserved by the LIFG). At the very least, it demonstrates that semantic interference across the three naming paradigms does not necessarily reflect identical mechanisms (cf. Oppenheim et al., 2010). Of the three paradigms, blocked cyclic naming is the least akin to naturalistic speech, involving the massed repetition/cycling of a small set of responses. The prominence afforded this paradigm in speech production accounts is therefore questionable. It is also worth emphasising that semantic interference effects may not necessarily reflect lexical-level processes either. In other naming paradigms, semantic interference has been attributed to prelexical, conceptual processes, e.g., postcue naming (Dean et al., 2001; Hocking, McMahon, & de Zubicaray, 2010) and negative priming (de Zubicaray et al., 2006; Tipper, 1985).

As we noted, proposals for top-down involvement have now extended to simple production tasks such as basic picture naming (e.g., Munding et al., 2016; Strijkers & Costa, 2016). Such proposals require theoretical motivation beyond the mere observation of neurophysiological responses. For example, when speakers can produce 2 to 4 words per second, and produce errors no more than 1 to 2 times every 1,000 words during every day speaking (Levelt et al., 1999), it would be useful to explain precisely why top-down intervention in selecting lexical candidates for production is so essential.

This is not to say that there is no evidence of left IFG or other control-related mechanisms in the studies we reviewed. In neuroimaging studies of the PWI paradigm, lateral and medial frontal (ACC or SMA) engagement was reported relatively consistently for contrasts of related versus identity or neutral conditions (de Zubicaray et al., 2001; Piai et al., 2013, 2014) and patients with left IFG lesions similarly showed increased interference for the contrast of a lexical distractor versus a neutral condition when compared to healthy controls (Piai et al., 2016). This suggests a role for left IFG in resolving competition introduced by competing *linguistic* information, rather than selecting among semantic competitors per se. This may reflect the operation of an early attention blocking mechanism as Piai et al. (2016) suggested.

A view to the future

In this chapter, we examined a number of issues facing researchers when investigating the spatiotemporal components of speech production. A relatively clear impression from our review is that sub-optimal methods have compromised a significant proportion of studies. Inconsistencies in BOLD fMRI study results were a regular outcome with standard continuous imaging acquisitions. We recommend continuous BOLD acquisitions be avoided in future fMRI studies of speech production. As neuroimaging is an expensive enterprise, adherence to methodological best practice is both a scientific and economic imperative. Similarly, findings from electrophysiological recordings showed such variability that they were uninterpretable for some psycholinguistic effects. A consensus approach to design and analysis with these techniques is sorely needed. Non-standard experimental designs were also frequently a source of problems for interpretation. Throughout our review, we relied on converging evidence from multiple sources of data to support our interpretations. Perhaps the only relatively consistent finding across paradigms was for the 250-450 ms time window and posterior temporal lobe (MTG/STG) involvement. Notably, many arguments that engendered debate in the literature emphasised data from a single modality; proposals concerning evidence for parallel rather than serial activation of processing stages being a prominent example.

We reviewed findings from context manipulations in picture naming paradigms as these reflect the best developed evidence base for investigating the spatial and temporal components of processing stages in spoken word production, and in particular the stage of retrieving words from the mental lexicon. The neurolinguistic literature currently reflects only a fraction of the context manipulations conducted in psycholinguistic studies, and is strongly biased toward semantics and monolingual production. This imbalance needs to be addressed. Outside of Stroop-like colour naming paradigms that have been the topic of many reviews, there is a paucity of neurolinguistic evidence concerning context manipulations in reading aloud. Recent sparse and perfusion fMRI studies have successfully mapped the networks involved in sentence production (e.g., Geranmayeh et al., 2014; Kemeny et al., 2005; Tremblay & Small, 2011), paving the way for more sophisticated manipulations. Finally, compared to the psycholinguistic literature, neurolinguistic studies of self-monitoring mechanisms are also scarce, and there is a clear tension between domain general and speech perception based accounts that should prove a fruitful area for enquiry.

References

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

Abel, S., Dressel, K., Weiller, C., & Huber, W. (2012). Enhancement and suppression in a lexical interference fMRI-paradigm. *Brain and Behavior*, 2, 109–127.

Aristei, S., Melinger, A., & Abdel Rahman, R. (2011). Electrophysiological chronometry of semantic context effects in language production. *Journal of Cognitive Neuroscience*, 23, 1567-1586.

Belke, E. (2013). Long-lasting inhibitory semantic context effects on object naming are necessarily conceptually mediated: implications for models of lexical-semantic encoding. *Journal of Memory and Language*, 69, 228-256.

Belke, E., & Stielow, A. (2013). Cumulative and non-cumulative semantic interference in object naming: evidence from blocked and continuous manipulations of semantic context. *Quarterly Journal of Experimental Psychology*, 66(11), 2135-2160.

Blackford, T., Holcomb, P. J., Grainger, J., & Kuperberg, G. R. (2012). A funny thing happened on the way to articulation: N400 attenuation despite behavioral interference in picture naming. *Cognition*, 123(1), 84–99.

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

Bürki, A. (2017). Electrophysiological characterization of facilitation and interference in the picture-word interference paradigm. *Psychophysiology*, 54(9), 1370-1392.

Butler, R. A., Ralph, M. A. L., & Woollams, A. M. (2014). Capturing multidimensionality in stroke aphasia: Mapping principal behavioural components to neural structures. *Brain*, 137(12), 3248-3266.

Canini, M., Della Rosa P. A., Catricalà E., Strijkers K., Branzi F. M., Costa A., et al. (2016). Semantic interference and its control: a functional neuroimaging and connectivity study. *Human Brain Mapping*, 37(11), 4179-4196.

Chouinard, P. A., Whitwell, R. L., & Goodale, M. A. (2009) The lateral-occipital and the inferior-frontal cortex play different roles during the naming of visually-presented objects. *Human Brain Mapping*, 30, 3851-3864.

Costa, A., Strijkers, K., Martin, C. D., & Thierry, G. (2009). The time-course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences USA*, 106, 21442–21446.

Damian, M. F., & Als, L. C. (2005). Long-lasting semantic context effects in the spoken production of object names. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 1372–1384.

Dell, G. S., Schwartz, M. F., Nozari, N., Faseyitan, O., & Coslett, H. B. (2013). Voxel-based lesion-parameter mapping: Identifying the neural correlates of a computational model of word production in aphasia. *Cognition*, 128, 380-396.

Dell'Acqua, R., Sessa, P., Peressotti, F., Mulatti, C., Navarrete, E., & Grainger, J. (2010). ERP evidence for ultra-fast semantic processing in the picture-word interference paradigm. *Frontiers in Psychology*, 1, 177.

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

de Zubicaray, G. I., & McMahon, K. L. (2009). Auditory context effects in picture naming investigated with event-related fMRI. *Cognitive, Affective and Behavioral Neuroscience*, 9(3), 260-269.

de Zubicaray, G. I., Hansen, S., & McMahon, K. L. (2013). Differential processing of thematic and categorical conceptual relations in spoken word production. *Journal of Experimental Psychology: General*, 142, 131-142.

de Zubicaray, G. I., McMahon, K. L., Eastburn, M. M., & 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(4), 1084-1093.

de Zubicaray, G. I., McMahon, K., & Howard, D. (2015). Perfusion fMRI evidence for priming of shared feature-to lexical connections during cumulative semantic interference in spoken word production. *Language and Cognitive Neuroscience*, 30(3), 261-272.

de Zubicaray, G. I., Wilson, S. J., McMahon, K. L., & Muthiah, S. (2001). The semantic interference effect in the picture-word paradigm: an event-related fMRI study employing overt responses. *Human Brain Mapping*, 14(4), 218-227.

de Zubicaray, G., Fraser, D., Ramajoo, K., & McMahon, K. (2017). Interference from related actions in spoken word production: Behavioural and fMRI evidence. *Neuropsychologia*, 96, 78-88.

de Zubicaray, G., Johnson, K., Howard, D., & McMahon, K. (2014). A perfusion fMRI investigation of thematic and categorical context effects in the spoken production of object names. *Cortex*, 54, 135-149.

de Zubicaray, G., McMahon, K., Eastburn, M., Pringle, A., & Lorenz, L. (2006). Classic identity negative priming involves accessing semantic representations in the left anterior temporal cortex. *Neuroimage*, 33(1), 383-390.

de Zubicaray, G.I. (2012). Strong inference in functional neuroimaging. *Australian Journal of Psychology*, 64, 19-28.

Detre, J. A., Rao, H., Wang, D. J., Chen, Y. F., & Wang, Z. (2012). Applications of arterial spin labeled MRI in the brain. *Journal of Magnetic Resonance Imaging*, 35, 1026-1037.

Diaz, M. T., Hogstrom, L. J., Zhuang, J., Voyvodic, J. T., Johnson, M. J. & Camblin, C. C. (2014). The influence of written distractor words on brain activity during overt picture naming. *Frontiers in Human Neuroscience*. 8, 167.

Duss, S. B., Reber, T. P., Hänggi, J., Schwab, S., Wiest, R., Müri, R.M., et al. (2014). Unconscious relational encoding depends on hippocampus. *Brain*, 27, 3355–3370.

Fargier, R., Bürki, A., Pinet, S., Alario, F-X., & Laganaro, M. (2017). Word onset phonetic properties and motor artifacts in speech production EEG recordings. *Psychophysiology*. doi: 10.1111/psyp.12982.

Galgano, J., & Froud, K. (2008). Evidence of the voice-related cortical potential: an electroencephalographic study. *Neuroimage*, 41, 1313–1323.

Ganushchak, L. Y., Christoffels, I. K., & Schiller, N. O. (2011). The use of electroencephalography in language production research: A review. *Frontiers in Psychology*, 2, 208.

Gauvin, H., Meinzer, M., & de Zubicaray, G. (2017). tDCS effects on word production: limited by design? Comment on Westwood et al. (2017). *Cortex*, 96, 137-142.

Geranmayeh, F., Wise, R. J., Mehta, A., & Leech, R. (2014). Overlapping networks engaged during spoken language production and its cognitive control. *Journal of Neuroscience*, 34, 8728–8740.

Goldrick, M., & Rapp, B. (2007). Lexical and post-lexical phonological representations in spoken production. *Cognition*,102, 219-260

Gracco, V. L., Tremblay, P, & Pike, G. B. (2005). Imaging speech production using fMRI. *Neuroimage*, 26, 294-301.

Hartwigsen, G. (2014). The neurophysiology of language: Insights from non-invasive brain stimulation in the healthy human brain. *Brain & Language*, 148, 81–94.

Harvey, D., & Schnur, T. T. (2015). Distinct loci of lexical and semantic access deficits in aphasia: Evidence from voxel-based lesion-symptom mapping and diffusion tensor imaging. *Cortex*, 67, 37–58.

Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience*, 11(7), 523–532.

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

Henseler, O., Mädebach, A., Kotz, S. A., & Jescheniak, J. D. (2014). Modulating brain mechanisms resolving lexico-semantic interference during word production: a transcranial direct current stimulation study. *Journal of Cognitive Neuroscience*, 26 (7), 1403-1417.

Hocking, J., McMahon, K. L., & de Zubicaray, G. I. (2009). Semantic context and visual feature effects in object naming: an fMRI study using arterial spin labeling. *Journal of Cognitive Neuroscience*, 21(8), 1571-1583.

Hocking, J., McMahon, K. L., & de Zubicaray, G. I. (2010). Semantic interference in object naming: an fMRI study of the postcue naming paradigm. *Neuroimage*, 50(2), 796-801.

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

Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. *Frontiers in Psychology*, 2, 255.

Indefrey, P. (2016). On putative shortcomings and dangerous future avenues: response to Strijkers & Costa. *Language, Cognition & Neuroscience*, 31, 517-520.

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

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

Janssen, N., Hernández-Cabrera, J. A., van der Meij, M., & Barber, H. A. (2015). Tracking the time course of competition during word production: Evidence for a post-retrieval mechanism of conflict resolution. *Cerebral Cortex*, 25(9), 2960–2969.

Kemeny, S., Ye, F. Q., Birn, R., & Braun, A. R. (2005). Comparison of continuous overt speech fMRI using BOLD and arterial spin labeling. *Human Brain Mapping*, 24, 173-183.

Krieger-Redwood, K., & Jefferies, E. (2014). TMS interferes with lexical-semantic retrieval in left inferior frontal gyrus and posterior middle temporal gyrus: Evidence from cyclical picture naming. *Neuropsychologia*, 64, 24–32.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1–38.

Llorens, A., Trébuchon, A., Liégeois-Chauvel, C., & Alario F-X. (2011). Intracranial recordings of brain activity during language production. *Frontiers in Psychology*, 2, 375.

Llorens, A., Trebuchon, A., Riès, S., Alario, F.-X., & Liegeois-Chauvel, C. (2014). How familiarization and repetition modulate the picture naming network. *Brain & Language*, 133, 47-58.

Maess, B., Friederici, A. D., Damian, M., Meyer, A. S., & Levelt, W. J. (2002). Semantic category interference in overt picture naming: Sharpening current density localization by PCA. *Journal of Cognitive Neuroscience*, 14, 455–462.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture–word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 33, 503–535.

Mehta, S., Grabowski, T. J., Razavi, M., Eaton, B., & Bolinger, L. (2006). Analysis of speech-related variance in rapid event-related fMRI using a time-aware acquisition system. *Neuroimage*, 29, 1278–1293.

Meinzer, M., Antonenko, D., Lindenberg, R., Hetzer, S., Ulm, L., Avirame, K., ... Flöel, A. (2012). Electrical brain stimulation improves cognitive performance by modulating functional connectivity and task-specific activation. *Journal of Neuroscience*, 32(5), 1859–1866.

Meinzer, M., Yetim, O., McMahon, K., & de Zubicaray, G. (2016). Brain mechanisms of semantic interference in spoken word production: An anodal transcranial Direct Current Stimulation (atDCS) study. *Brain & Language*, 157-158, 72-80.

Miniussi, C., Harris, J. A., & Ruzzoli, M. (2013). Modelling non-invasive brain stimulation in cognitive neuroscience. *Neuroscience and Biobehavioral Reviews*, 37(8), 1702–1712.

Mirman, D. & Graziano, K. M. (2013). The neural basis of inhibitory effects of semantic and phonological neighbors in spoken word production. *Journal of Cognitive Neuroscience*, 25(9), 1504-1516.

Mottaghy, F. M., Sparing, R., & Topper, R. (2006). Enhancing picture naming with transcranial magnetic stimulation. *Behavioral Neurology*, 17(3-4), 177-186.

Mulatti, C., Peresotti, F., Job, R., Saunders, S. & Coltheart, M. (2012). Reading aloud: The cumulative lexical interference effect. *Psychonomic Bulletin & Review*, 19, 662-667.

Munding, D., Dubarry, A.-S., & Alario, F.-X. (2015). On the cortical dynamics of word production: a review of the MEG evidence. *Language, Cognition and Neuroscience*. doi:10.1080/23273798.2015.1071857

Nebel, K., Stude, P., Wiese, H., Müller, B., de Greiff, A., Forsteing, M., Diener, H.C., Keidel, M. (2005): Sparse imaging and continuous event-related fMRI in the visual domain: a systematic comparison. *Human Brain Mapping*, 24, 130–143.

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

Ouyang, G., Sommer, W., Zhou, C., Aristei, S., Pinkpank, T., & Abdel Rahman, R. (2016). Articulation artifacts during overt language production in event-related brain potentials: Description and correction. *Brain Topography*, 29(6), 791–813.

Peramunage, D., Blumstein, S. E., Myers, E., Goldrick, M., & Baese-Berk, M. (2011). Phonological neighborhood effects in spoken word production: An fMRI study. *Journal of Cognitive Neuroscience,* 23, 593-603.

Piai, V. (2016). The role of electrophysiology in informing theories of word production: a critical standpoint. *Language, Cognition and Neuroscience*, 31, 471-473.

Piai, V., & Knight, R.T. (2017). Lexical selection with competing distractors: Evidence from left temporal lobe lesions. *Psychonomic Bulletin & Review*. In press

Piai, V., Dahlslätt, K., & Maris, E. (2015). Statistically comparing EEG/MEG waveforms through successive significant univariate tests: How bad can it be? *Psychophysiology*, 52, 440-443.

Piai, V., Riès, S.K., & Knight, R.T. (2015). The electrophysiology of language production: what could be improved. *Frontiers in Psychology*, 5, 5160.

Piai, V., Riès, S.K., & Swick, D. (2016). Lesions to lateral prefrontal cortex impair interference control in word production. *Frontiers in Human Neuroscience*, 9, 721.

Piai, V., Roelofs, A., Acheson, D. J., & Takashima, A. (2013). Attention for speaking: domain-general control from anterior cingulate cortex in spoken word production. *Frontiers in Human Neuroscience*, 7, 832.

Piai, V., Roelofs, A., Jensen, O., Schoffelen, J.M., & Bonnefond, M. (2014). Distinct patterns of brain activity characterise lexical activation and competition in spoken word production. *PLoS ONE*, 9(2), e88674.

Piai, V., Roelofs, A., & Van der Meij, R. (2012). Event-related potentials and oscillatory brain responses associated with semantic and Stroop-like interference effects in overt naming. *Brain Research*, 1450, 87-101.

Pisoni, A., Cerciello, M., Cattaneo, Z., & Papagno, C. (2017). Phonological facilitation in picture naming: When and where? A tDCS study. *Neuroscience*, 352, 106-121.

Pisoni, A., Papagno, C., & Cattaneo, Z. (2012). Neural correlates of the semantic interference effect: New evidence from transcranial direct current stimulation. *Neuroscience*, 223, 56–67.

Porcaro, C., Medaglia, M. T., & Krott, A. (2015). Removing speech artifacts from electroencephalographic recordings during overt picture naming. *NeuroImage*, 105, 171–180.

Python, G., Fargier, R., & Laganaro, M. (2017). ERP evidence of distinct processes underlying semantic facilitation and interference in word production. *Cortex*, https://doi.org/10.1016/j.cortex.2017.09.008

Riès, S. (2015). Serial versus parallel neurobiological processes in language production: Comment on Munding, Dubarry, and Alario, 2015. *Language, Cognition, and Neuroscience*, doi:10.1080/23273798.2015.1117644

Riès, S. K., Greenhouse, I., Dronkers, N. F., Haaland, K. Y., & Knight, R. T. (2014). Double dissociation of the roles of the left and right prefrontal cortices in anticipatory regulation of action. *Neuropsychologia*, 63, 215–225.

Riès, S., Janssen, N., Burle, B., & Alario, F.-X. (2013). Response-locked brain dynamics of word production. *PLoS ONE*, 8(3): e58197.

Riès, S., Karzmark, C., Navarrete, E., Dronkers, N., & Knight, R.T. (2015). Specifying the role of the left prefrontal cortex in word selection. *Brain & Language*, 149:135-47.

Riès, S., Xie, K., Haaland, K., Dronkers, N., & Knight, R. T. (2013). Role of the lateral prefrontal cortex in speech monitoring. *Frontiers in Human Neuroscience*, 7, 703.

Riès, S.K., Dhillon, R.K., Clarke, A., King-Stephens, D., Laxer, K.D., Weber, P.B., ... & Lin, J.J. (2017). Spatiotemporal dynamics of word retrieval in speech production revealed by cortical high-frequency band activity. *Proceedings of the National Academy of Sciences USA*, 114, E4530-E4538.

Rizio, A.A., Moyer, K.J., & Diaz, M. T. (2017). Neural evidence for phonologically-based language production deficits in older adults: An fMRI investigation of age-related differences in picture-word interference. *Brain and Behavior*, 15:7(4), 1-19.

Roelofs A., & Piai, V. (2011). Attention demands of spoken word planning: A review. *Frontiers in Psychology*, 2, 307.

Rose, S. B., & Abdel Rahman, R. (2016). Semantic similarity promotes interference in the continuous naming paradigm: Behavioural and electrophysiological evidence. *Language, Cognition and Neuroscience*, 32, 55-68.

Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Development*, 46, 247-253.

Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of Broca’s area. *Proceedings of the National Academy of Sciences USA*, 106, 322–327.

Schuhmann, T., Schiller, N. O., Goebel, R., & Sack, A. T. (2009). The temporal characteristics of functional activation in Broca's area during overt picture naming. *Cortex,* 45(9), 1111-1116.

Schuhmann, T., Schiller, N. O., Goebel, R., & Sack, A. T. (2012). Speaking of which: Dissecting the neurocognitive network of language production in picture naming. *Cerebral Cortex*, 22, 701–709.

Schwartz, M. F., Faseyitan, O., Kim, J., & Coslett, H. B. (2012). The dorsal stream contribution to phonological retrieval in object naming. *Brain*, 135, 3799-3814.

Spalek, K., & Thompson-Schill, S. L. (2008). Task-dependent semantic interference in language production: An fMRI study. *Brain & Language*, 107, 220-228.

Strijkers K. (2016). Can hierarchical models display parallel cortical dynamics? A non-hierarchical alternative of brain language theory. *Language, Cognition and Neuroscience*, 31, 465-469.

Strijkers, K., & Costa, A. (2016). The cortical dynamics of speaking: present shortcomings and future avenues. *Language, Cognition and Neuroscience*, 31(4), 484–503.

Töpper, R., Mottaghy, F. M., Brugmann, M., Noth, J., & Huber, W. (1998). Facilitation of picture naming by focal transcranial magnetic stimulation of Wernicke's area. *Experimental Brain Research*, 121(4), 371-378.

Tremblay, P., & Small, S.L. (2011). Motor response selection in overt sentence production: a functional MRI study. *Frontiers in Psychology*, 2, 253.

Wang, M., Shao, Z., Chen, Y., & Schiller, N. O. (2017). Neural correlates of spoken word production in semantic and phonological blocked cyclic naming. *Language, Cognition and Neuroscience*, 33, in press. DOI: 10.1080/23273798.2017.1395467

Westwood, S. J., Olson, A., Miall, R. C., Nappo, R., & Romani, C. (2017). Limits to tDCS effects in language: Failures to modulate word production in healthy participants with frontal or temporal tDCS. *Cortex*, 86, 64-82.

Wheat, K. L., Cornelissen, P. L., Sack, A. T., Schuhmann, T., Goebel, R., & Blomert, L. (2013). Charting the functional relevance of Broca's area for visual word recognition and picture naming in Dutch using fMRI-guided TMS. *Brain & Language*, 125(2), 223-230.

Wilson, S.M., Isenberg, A.L., & Hickok, G. (2009). Neural correlates of word production stages delineated by parametric modulation of psycholinguistic variables. *Human Brain Mapping*, 30, 3596-608.

Wirth, M., Abdel Rahman, R., Kuenecke, J., Koeniga, T., Horn, H., Sommer, W., & Dierks, T. (2011). Effects of transcranial direct current stimulation (tDCS) on behaviour and electrophysiology of language production. *Neuropsychologia*, 49, 3989–3998.

Zhu, X., Zhang, Q., & Damian, M.F. (2016). Additivity of semantic and phonological effects: Evidence from speech production in Mandarin. *Quarterly Journal of Experimental Psychology*, 69, 2285-2304.