Neural Correlates of Naturally Occurring Speech Errors during Picture Naming in Healthy Participants

Angelique Volfart1, Katie L. McMahon1,2, David Howard3, and Greig I. de Zubicaray1

Abstract

Most of our knowledge about the neuroanatomy of speech errors comes from lesion-symptom mapping studies in people with aphasia and laboratory paradigms designed to elicit primarily phonological errors in healthy adults, with comparatively little evidence from naturally occurring speech errors. In this study, we analyzed perfusion fMRI data from 24 healthy participants during a picture naming task, classifying their responses into correct and different speech error types (e.g., semantic, phonological, omission errors). Total speech errors engaged a wide set of left-lateralized frontal, parietal, and temporal regions that were almost identical to those involved during the production of correct responses. We observed significant perfusion signal decreases in the left posterior middle temporal gyrus and inferior parietal lobule (angular gyrus) for semantic errors compared to correct trials matched on various psycholinguistic variables. In addition, the left dorsal caudate nucleus showed a significant perfusion signal decrease for omission (i.e., anomic) errors compared with matched correct trials. Surprisingly, we did not observe any significant perfusion signal changes in brain regions proposed to be associated with monitoring mechanisms during speech production (e.g., ACC, superior temporal gyrus). Overall, our findings provide evidence for distinct neural correlates of semantic and omission error types, with anomic speech errors likely resulting from failures to initiate articulatory-motor processes rather than semantic knowledge impairments as often reported for people with aphasia.

INTRODUCTION

Neurotypical humans are usually very efficient at producing a contextually correct word, being both fast and accurate at this task. In a naturalistic setting, this means that very few speech errors are committed: A frequently cited estimate is one or two in every thousand words produced (Levelt, 1999). Most of our knowledge about the brain correlates of various speech error types has therefore come from lesion-symptom mapping (LSM) studies in patients with chronic or acute stroke aphasia as they err more frequently, particularly during picture naming (e.g., Meier et al., 2022; Bruffaerts et al., 2020; Halai, Woollams, & Lambon Ralph, 2018; Stark et al., 2019; Tochadse, Halai, Lambon Ralph, & Abel, 2018; Schwartz et al., 2011; Walker et al., 2011). For example, semantic errors (e.g., producing “cat” when presented with the picture of a dog) have been primarily associated with lesions of the left anterior temporal lobe (ATL; Chen, Middleton, & Mirman, 2019; Mirman, Chen et al., 2015; Mirman, Zhang, Wang, Coslett, & Schwartz, 2015; Walker et al., 2011; Schwartz et al., 2009), consistent with the proposed involvement of this region in semantic processing (de Zubicaray, Rose, & McMahon, 2011; see also Halai et al., 2018; Tochadse et al., 2018; Lambon Ralph, Jefferies, Patterson, & Rogers, 2017; Lambon Ralph, Sage, Jones, & Mayberry, 2010). Several studies have also reported the involvement of more posterior regions (e.g., posterior middle and superior temporal gyrus [pMTG and pSTG, respectively], angular gyrus [AG] and occipital cortex; Meier et al., 2022; Stark et al., 2019; Fridriksson et al., 2018; Cloutman et al., 2009; see also Bruffaerts et al., 2020, for the distribution of this error type in patients with primary progressive aphasia). Phonological paraphasias on the other hand (e.g., producing “tog” instead of dog) have been consistently related to lesions in left inferior parietal and frontal regions (e.g., supramarginal gyrus [SMG], precentral and postcentral gyrus, premotor cortex; Halai et al., 2018; Mirman, Chen et al., 2015; Mirman, Zhang, et al., 2015; Schwartz, Faseyitan, Kim, & Coslett, 2012). Lesions have also been the primary source of information for associating neuroanatomical regions with semantic and phonological parameters (i.e., weights) in computational simulations of speech errors (e.g., Tochadse et al., 2018; Dell, Schwartz, Nozari, Faseyitan, & Branch Coslett, 2013).

Despite being relatively frequent in people with aphasia (PWA; Halai et al., 2018; see Lambon Ralph, Moriarty, & Sage, 2002), omission errors (i.e., failure to produce a complete object name) have been reported in only a handful of LSM studies. They have been recently linked to lesions in the left mid-anterior to posterior temporal cortex (Meier et al., 2022; Chen et al., 2019; Halai et al., 2018; Tochadse et al., 2018) and left frontal cortex (Meier et al., 2022; Chen et al., 2019) and attributed to deficits in lexical selection and semantic cognition processes. However, a recent LSM study in patients with primary progressive
aphasia reported an association with atrophy in the bilateral medial ATL and right AG (Bruffaerts et al., 2020). Hence, caution is needed when attempting to generalize results from lesion data as they may be subject to variations and specificities depending on the clinical population under scrutiny (e.g., lesion extent, hemispheric focus, sudden vs. progressive onset, time since stroke onset; see Breining et al., 2022).

Finally, another clinical source of information about the neural correlates of speech error types during picture naming comes from direct electrical stimulation (DES) studies in patients undergoing surgery for epilepsy and brain tumors. Semantic paraphasias have been consistently elicited after DES in the left SMG and inferior to superior temporal regions (Perrone-Bertolotti et al., 2020; Miozzo, Williams, McKhann, & Hamberger, 2017; Tate, Herbet, Moritz-Gasser, Tate, & Duffau, 2014; Corina et al., 2010; Duffau et al., 2005). However, another study focusing on the ventral temporal cortex did not report any occurrence of this error type (Bédos Ulvin et al., 2017). Phonological paraphasias have been reported upon stimulating the left inferior frontal gyrus (IFG), dorsolateral prefrontal cortex, and inferior to superior temporal regions (Perrone-Bertolotti et al., 2020; Bédos Ulvin et al., 2017; Miozzo et al., 2017; Tate et al., 2014). Across studies, DES most often led to a complete absence of response (i.e., anomia) and was associated with a wider set of bilateral but mostly left-lateralized frontal (IFG and precentral gyrus), inferior parietal, and temporal (including fusiform and inferior, middle, and superior temporal gyri) regions than the LSM studies (Perrone-Bertolotti et al., 2020; Bédos Ulvin et al., 2017; Miozzo et al., 2017; Tate et al., 2014; Corina et al., 2010).

**Evidence from Neurotypical Participants**

Most of our knowledge about the brain mechanisms of speech errors in neurotypical participants comes from neuroimaging studies employing experimental paradigms designed primarily to induce phonological/phonetic speech errors that occur relatively rarely in picture naming. These include “tongue twisters” (e.g., “How can a clam cram in a clean cream can?”; Okada, Matchin, & Hickok, 2018; Gauvin, De Baene, Brass, & Hartsuiker, 2016), priming of spoonerisms with word pairs (e.g., “tap coast” for the target “cap toast”; Runnqvist et al., 2021), and manipulation of auditory feedback during production of syllables and words (see Meekings & Scott, 2021, for a review). Other studies have examined lexical errors with stop-signal paradigms in which participants are required to withhold a naming response when presented with infrequent auditory signals (Hansen, McMahon, & de Zubircaray, 2019; Xue, Aron, & Poldrack, 2008) or with connected speech paradigms using special constraints (e.g., Grande et al., 2012). Although informative, these paradigms have typically elicited activation in a broad bilateral network of cortical and subcortical regions, likely reflecting the additional engagement of mechanisms specific to the manipulation employed (see Meekings & Scott, 2021). Within this network, authors have emphasized the involvement of extra-perisylvian language regions such as the cerebellum and ACC in internal and external monitoring/error detection (e.g., Runnqvist et al., 2021; Gauvin et al., 2016) in addition to STG.

TMS studies of picture naming have yielded mixed results in healthy participants, with semantic errors elicited by stimulation of the left SMG in one study (Lioumis et al., 2012) but showing no preferential locus in three others (Krieg et al., 2016; Hernandez-Pavon, Mäkelä, Lehtinen, Lioumis, & Mäkelä, 2014; Rösler et al., 2014), precluding any conclusions about the brain regions associated with this error type. Phonological errors also occurred too rarely across studies to be reliably associated with any specific brain region (but see Sakreida et al., 2018, for an anterior–posterior distinction within the IFG for semantic and phonological errors, respectively). However, omissions (also referred to as anomia) have been consistently associated with stimulation over several cortical regions, including the left IFG, precentral and postcentral gyri, SMG, and STG (Krieg et al., 2016; Hernandez-Pavon et al., 2014; Rösler et al., 2014; Lioumis et al., 2012).

Tip-of-the-tongue (ToT) phenomena are also informative for the neural correlates of different speech error types, as their behavioral manifestation is highly similar if not identical to omission or anomic errors. In all these instances, the participant is unable to retrieve or produce a target word, although the underlying cognitive processes for this failure likely differ. Most often studied with proper name retrieval tasks in older adults (but see Ubaldi, Rabini, & Fairhall, 2022; Maril, Wagner, & Schacter, 2001, for examples of general knowledge tasks), ToT states have been associated with bilateral frontal, insular, and cingulate regions (Ubaldi et al., 2022; Huijbers et al., 2017; Shafto, Stamatakis, Tam, & Tyler, 2010; Shafto, Burke, Stamatakis, Tam, & Tyler, 2007; Kikyo, Ohki, & Sekihara, 2001; Maril et al., 2001) that have been interpreted as reflecting error monitoring processes (Neta et al., 2015; for a review, see Ullsperger, Harsay, Wessel, & Riddervold, 2010). Interestingly, unlike the LSM studies of omission errors mentioned above, anterior temporal regions have only rarely been reported in relation with ToT states in healthy participants (Shafto et al., 2010). This might be because a ToT state is always accompanied by a report of preserved knowledge of the concept for which the word form cannot be retrieved, whereas anomic errors in patients can also reflect either loss or impaired retrieval of semantic knowledge (when assessed). Hence, differences in the neural correlates of speech errors in clinical and non-clinical populations will likely reflect the involvement of different mechanisms.

As Meekings and Scott (2021) noted, paradigms that do not involve external manipulation are likely to come
closest to inducing natural speech errors as conceptualized by models of spoken production. To our knowledge, only one fMRI study to date (Abel et al., 2009) has reported brain regions activated by naturally occurring speech errors during a picture naming task in a sample of left- and right-handed healthy adults. Abel and colleagues compared total errors versus matched correct responses in 22 participants who had committed seven or more errors. This contrast identified significantly greater activation for errors in right SMA and middle frontal gyrus (MFG) and left insula. Abel and colleagues also noted that most of their speech errors were semantic, followed by omissions and mixed semantic/phonological errors. However, they did not investigate fMRI activation associated with these specific error types.

**The Present Study**

To achieve a better understanding of the neural correlates of different speech error types elicited in a naturalistic setting in neurotypical adults (i.e., without any external task manipulation or brain stimulation), we conducted novel analyses of data from a perfusion fMRI study of picture naming in healthy, neurotypical participants (de Zubicaray, McMahon, & Howard, 2015). This data set allowed us to explore the neural correlates of the two most frequent error types in our healthy sample, namely, semantic and omission errors. In addition, we investigated the influence of various psycholinguistic variables, including lexical frequency and age of acquisition (AoA), that have been reported to influence error production in both healthy participants and PWA (e.g., Brysbaert & Ellis, 2016; Abel et al., 2009). On the basis of the literature from both healthy and clinical populations we reviewed above, we hypothesized errors would be associated with perfusion signal changes primarily in left-hemisphere perisylvian language regions, especially IFG, ATL, MTG, STG, AG, and SMG, as well as regions implicated in error detection/monitoring (e.g., ACC, STG), with semantic and omission errors involving distinguishable subsets of these regions.

**METHODS**

**Participants**

Participants were 27 healthy right-handed native English speakers (11 men, mean age = 22.74 years) recruited among students and staff of the University of Queensland. They all had normal or corrected-to-normal vision, no history of neurological or psychiatric disorder, no substance dependence, and no hearing deficits. All participants gave their informed consent and were paid AUD$30 to participate in the study that was approved by the Medical Research Ethics Committee of the University of Queensland. As in de Zubicaray and colleagues (2015), the analysis was carried out on the fMRI data of 24 participants (10 men, mean age = 23.10 years) because three participants were excluded due to excessive head movement (>3-mm translation; n = 1), use of a deteminer before naming every trial (e.g., “a bear”; n = 1), or use of superordinate category for all items within the category (e.g., “fruit” rather than “apple”, “pear”; n = 1). This sample size exceeds the recommendation from Thirion and colleagues (2007) for fMRI studies, ensuring sufficient power.

**Stimuli and Procedure**

Stimuli were the same as those presented in Howard, Nickels, Coltheart, and Cole-Virtue (2006) and de Zubicaray and colleagues (2015). Participants were presented with 165 color pictures, including 120 items belonging to 24 semantic categories (five exemplars in each) and 45 filler items, and instructed to name each with a single word as rapidly and accurately as possible. Each participant was presented with a unique experimental list, in which category exemplars were separated by different lags (two, four, six, or eight items) in different orders. For each trial, a fixation cross was first displayed for 500 msec, followed by a blank screen for 250 msec. Each picture was presented for 2000 msec on the screen and was followed by a blank screen for 3750 msec.

**Apparatus**

The experiment was run on a PC with Microsoft VisualBasic and ExacTicks (Ryle Design) software. Pictures were displayed centrally on a white background and back-projected using a BenQ MW512 projector onto a screen that the participants viewed through a mirror mounted on the head coil. Each picture subtended about 10° of visual angle. A 30-dB attenuating headset was used to reduce gradient noise.

Naming responses were recorded on digital audio files (sampling rate = 11 kHz) using a custom positioned fiber-optic dual-channel noise-canceling microphone attached to the head coil (FOMRI-III, Optoacoustics Ltd.: www.optoacoustics.com). Naming latencies were automatically detected through custom voice-key software and manually checked through Audacity software (sourceforge.net/projects/audacity/) per de Zubicaray and colleagues (2015).

**Analysis of Naming Responses**

Both experimental and filler items were included, unlike de Zubicaray and colleagues (2015) who reported only correct trials from the series of semantic category exemplars. Following Howard and colleagues (2006), we also included “acceptable alternative” responses for picture names (see their Appendix for examples).

Naming responses were classified following Dell’s classification (Abel et al., 2009; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997), as correct, semantic error,
phonological error, mixed semantic/phonological error, unrelated error, nonword error, and omission error, with an additional dysfluency category (e.g., stuttering: “heah-headphones”; hesitations: “ah…paint”).

Various psycholinguistic parameters (SUBTLEX word frequency, number of phonemes, number of syllables, AoA) were assembled for each item through eLexicon (elexicon.wustl.edu/; Balota et al., 2007). Visual descriptors were also collected for each picture item: Visual complexity of the image was assessed through its JPEG file size (Donderi, 2006; Székely & Bates, 2000), and mean luminosity of the image was measured through Photoshop Histogram’s tool (Adobe Photoshop v. 23.1.1, www.adobe.com/sea/).

We compared correct responses and error types according to these psycholinguistic and visual parameters using independent t tests (two-tailed) in SPSS (v. 28.0.1.0, www.ibm.com/analytics/spss-statistics-software), checking for homogeneity of variance with Levene’s tests and adjusting for unequal variances where required. p Values were corrected for multiple comparisons using Bonferroni FWE correction.

We also conducted a logistic regression analysis in JASP (v. 0.16.1, jasp-stats.org/) to investigate the influence of these psycholinguistic and visual variables on speech error rate. Variables that contributed significantly to the error rate (see Results section below) were used to build sets of matched correct trials for total errors, semantic errors, and omission errors, in an analogous way to Abel and colleagues (2009), that is, for each participant with errors comprising 5% or more of the stimulus set (i.e., ≥8 errors). This resulted in 24 participants meeting that criterion for total errors, 23 for semantic errors, and nine for omission errors. Correct items were selected to match the characteristics of speech errors (on word frequency, number of phonemes, number of syllables, and AoA) using Match software (www.mrc-cbu.cam.ac.uk/people/maarten-van-casteren/mixandmatch/).

MRI Acquisition and Image Analysis

Participants were scanned on a Bruker Medspec 4-T system equipped with a transverse electromagnetic head coil (Vaughan et al., 2002). The pulsed arterial spin labeling (ASL) was acquired using FAIREST (Wang et al., 2002). The pulsed arterial spin labeling (ASL) was acquired using FAIREST (Wang et al., 2002). The saturation slab was applied inferior to the imaging slab, with a 10-mm margin at each edge, to ensure optimum inversion. A single run of 446 images (alternating control and tag) was acquired: Inversion Time 1 = 800 msec, Inversion Time 2 = 1800 msec, repetition time (RT)/echo time = 2500/11 msec, matrix = 64 × 64, voxel in-plane resolution = 3.5 × 3.5 mm², and flip angle = 90°. Each volume had 12 slices, 6 mm thick with a 1.5-mm gap, and oriented to optimize coverage of the temporal and inferior frontal lobes. The first four volumes were discarded to allow tissue magnetization to achieve steady state. A 3-D T1-weighted magnetization prepared rapid gradient echo structural image was acquired: TR/echo time = 2300/2.95 msec, inversion time = 900 msec, 256 × 256 × 176 matrix, voxel in-plane resolution = 0.94 × 0.94 × 0.9 mm, and pulse angle = 9°. The total scan time was approximately 50 min (de Zubicaray et al., 2015).

Preprocessing was carried out using the ASL toolbox (ASLtbx; Wang et al., 2008) implemented in SPM8 (Wellcome Department of Imaging Neuroscience, UCL Queen Square Institute of Neurology; www.fil.ion.ucl.ac.uk/spm/software/spm8/) identically to de Zubicaray and colleagues (2015). Motion correction was applied on ASL image series using INRAlign (Freire, Roche, & Mangin, 2002), and the realigned series were smoothed with an 8-mm FWHM isotropic Gaussian kernel. The segmented T1-weighted image was used to create an intracranial mask to exclude extracranial voxels for cerebral blood flow calculation. Perfusion imaging time series were constructed by a pairwise simple subtraction between label (tagged) and control acquisitions, resulting in image volumes with an effective TR of 5 sec (Liu & Wong, 2005). A mean image was created from the perfusion time series and coregistered to the T1-weighted structural image. The structural image was spatially normalized to Montreal Neurological Institute atlas space, and the transformation was applied to the coregistered perfusion imaging time series, and volumes were resliced to 2-mm³ voxels.

Analyses were carried out in SPM12 (Wellcome Department of Imaging Neuroscience, UCL Queen Square Institute of Neurology; www.fil.ion.ucl.ac.uk/spm/software/spm12/). At the participant level, event types corresponding to error types, their matched correct items, and remaining correct items were modeled as effects of interest with delta functions representing each picture onset. We conducted two-stage, mixed-effect statistical analyses. Each participant’s contrasts were then entered in a group-level analysis.

Our primary analysis comprised one t contrast and one 1-way repeated-measures ANOVA to identify the regional perfusion signal changes associated with total errors relative to baseline (fixation cross) and to compare perfusion signal changes associated with total errors relative to matched correct trials, respectively. Secondary analyses followed the same principle to compare semantic and omission error types with baseline and with their corresponding matched correct trials. Semantic and omission errors were not directly compared to each other because of different sample sizes (23 participants for semantic error analyses, nine for omission error analyses).

We conducted both unrestricted whole-brain and ROI analyses targeting specific brain regions identified by our literature review as being associated with speech error production and monitoring across patient and healthy participant studies. ROIs were selected from the Brainnetome Atlas parcellation (atlas.brainnetome.org/brainnetome.html; Fan et al., 2016; Figure 1) and included the left anterior, middle, and posterior sections of the middle
temporal gyrus (antMTG, midMTG, and pMTG, respectively; Chen et al., 2019; Corina et al., 2010; see, e.g., Stark et al., 2019; Tate et al., 2014), the pSTG (e.g., Stark et al., 2019; Tate et al., 2014), the AG (e.g., Fridriksson et al., 2018; Cloutman et al., 2009), the SMG (e.g., Lioumis et al., 2012; Cloutman et al., 2009; also previously reported in Abel et al., 2009, as associated with error rate), the orbitofrontal gyrus (OFG; e.g., Meier et al., 2022; Duffau et al., 2005; also correlated with the number of errors in Abel et al., 2009), the IFG (anterior [aIFG] and posterior [pIFG] parts corresponding to Brodmann’s areas [BA] 44 and 45, respectively; e.g., Perrone-Bertolotti et al., 2020; Tate et al., 2014; Schwartz et al., 2009), and the bilateral ACC (e.g., Bruffaerts et al., 2020; Huijbers et al., 2017; Shafto et al., 2007). A height threshold of $p < .001$ was adopted in conjunction with spatial cluster extent thresholds of $p < .05$ (FWE-corrected).

RESULTS

Classification of Picture Naming Errors

A summary of behavioral responses (number and percentage of correct responses and errors at naming pictures) is presented in Table 1. On average, participants produced 23.71 errors on the whole set of 165 items (14.37%, ranging from 6.06% to 27.27%). Most errors were semantic (and taxonomic in nature, e.g., “amulet” for earring, “chair” for stool), followed by omissions (anomia) and unrelated errors (e.g., “statue” for waterfall, “frog” for biscuit).

As the original Howard and colleagues (2006) experimental stimuli included exemplars from 24 semantic categories, we conducted a separate chi-square analysis to compare the frequency of semantic errors across categories. There was a significant difference between categories: $\chi^2 = 149.876, p < .001$. Post hoc tests showed that there were less errors than predicted by the null hypothesis for the body parts and fruits categories (FWE-corrected $p$ values at $p = .010$ and .034, respectively) and more errors than predicted for the clothes and shellfish categories (FWE-corrected $p$ values at $p < .001$ and $p = .002$, respectively), suggesting that some categories were more susceptible to semantic errors than others (Figure 2).

We compared psycholinguistic and visual parameters for items that were consistently named correctly (correct responses in 100% of participants; 48 items) and items that were incorrectly named by at least one participant (independently of the error type; 117 items; Table 2). Two-tailed independent $t$ tests showed only a significant difference in AoA between consistently correct and erroneous items (mean difference = $-1.30, t = -5.36$, FWE-corrected $p < .001$), highlighting that the names of consistently correct items were acquired earlier in life. Other $t$ tests did not reveal any significant difference between the two response sets ($t$ range: $-1.57$ to $0.09$, all FWE-corrected $p > .05$).

Comparisons were also carried out between consistently correct items and items eliciting at least one semantic error, one omission error or one unrelated error.

| Response Type       | Number of Responses (Percentage) |
|---------------------|----------------------------------|
|                     | All Participants $(n = 24)$       | Participants with $\geq 8$ Semantic Errors $(n = 23)$ | Participants with $\geq 8$ Omission Errors $(n = 9)$ |
| Correct            | 3391 (85.63)                     | 3236 (85.27)                                      | 1213 (81.68)                                      |
| Error              | 569 (14.37)                      | 559 (14.73)                                      | 272 (18.32)                                       |
| Semantic           | 337 (8.51)                       | 330 (8.69)                                       | 130 (8.75)                                        |
| Phonological       | 0 (0)                            | 0 (0)                                            | 0 (0)                                             |
| Mixed              | 6 (0.15)                         | 6 (0.16)                                         | 1 (0.07)                                          |
| Unrelated          | 45 (1.14)                        | 44 (1.16)                                        | 21 (1.41)                                         |
| Nonword            | 171 (4.32)                       | 169 (4.45)                                       | 118 (7.95)                                        |
| Omission           | 9 (0.23)                         | 9 (0.24)                                         | 2 (0.14)                                          |
| Dysfluency         |                                  |                                                 |                                                  |
| Total              | 3960 (100)                       | 3795 (100)                                       | 1485 (100)                                        |
Comparisons with other error types were not carried out as there were less than 10 different items involved. Two-tailed independent t tests comparing correct responses and semantic errors showed again a significant difference in AoA (mean difference = -1.43, t = -5.07, FWE-corrected p < .001) but no difference on other variables (t range: -1.49 to 1.54, all FWE-corrected p > .05). The same pattern was observed when comparing correct responses with omission errors (mean difference = -1.70 to -0.16, all FWE-corrected p > .05 for other variables) and unrelated errors (mean difference = -1.86, t = -4.67, FWE-corrected p < .001 for AoA; t range: -1.28 to 0.13, all FWE-corrected p > .05 for other variables).

Overall, these results indicate that errors, independently of type, occurred more frequently on items whose name was acquired later in life.

Logistic regression showed that several psycholinguistic variables were predictive of the error rate. The log odds of a correct response were significantly higher (1.3) for words with a larger number of phonemes (ß coefficient = 0.26, SE = 0.05, Wald z = 5.24, p < .001) and (1.004) more frequent words (ß coefficient = 0.004, SE = 0.001, Wald z = 3.05, p = .002). Conversely, the log odds of a correct response were significantly lower for words acquired later in life (0.8; ß coefficient = -0.22, SE = 0.03, Wald z = -8.34, p < .001) and for words of more than one syllable (0.57; ß coefficient = -0.57, SE = .11, Wald z = -5.27, p < .001). Neither luminosity (log odds = 0.997; ß coefficient = -0.003, SE = .003, Wald z = -1.18, p = .24) nor visual complexity (log odds = 0.999; ß coefficient = -0.001, SE = .002, Wald z = -0.82, p = .41) of the depicted objects significantly influenced the probability of a correct response.

Imaging Data: Total Errors

Unrestricted Whole-brain Analyses

The t-contrast total errors > baseline (fixation cross) revealed significant perfusion signal increases in the bilateral inferior temporal gyrus (ITG) or BA 37, the left precen- tral gyrus or BA 4 (tongue and larynx region), and the right postcentral gyrus or BAs 1–3 (upper limb; head and face region; Table 3). The reverse contrast showed significant perfusion signal decreases in the bilateral precuneus or BA 31 and in the thalamus.

The one-way ANOVA contrasting total errors with their matched correct trials did not reveal any significant cluster in either direction.

ROI Analyses

The t-contrast total errors > baseline revealed a significant perfusion signal increase in the left pSTG, the left IFG (both BAs 44 and 45), the left SMG, and the bilateral ACC (Table 3). The reverse contrast showed a significant perfusion signal decrease in the left antMTG and midMTG.

The one-way ANOVA contrasting total errors with their matched correct trials did not reveal any significant perfusion increase but showed a significant perfusion decrease in the left AG (Figure 3).

Imaging Data: Categorizing Errors into Semantic and Omission Subtypes

After the analysis at the level of total errors, we also analyzed the pattern of brain responses for the two most common error types in our sample, that is, semantic errors and omission errors.
Table 2. Psycholinguistic and Visual Parameters of Consistently Correct Items, All Items with at Least One Participant Failing to Correctly Name It, Items Eliciting at Least One Semantic Error, Items Eliciting at Least One Omission Error, and Items Eliciting at Least One Unrelated Error

| Response Type | n Items | Word Frequency | n Phonemes | n Syllables | AoA | Visual Complexity | Mean Luminosity |
|---------------|---------|----------------|------------|-------------|-----|-------------------|----------------|
| Correct       | 48      | 35.52 (7.5)[0.2–292.1] | 4.40 (0.2)[2–9] | 1.54 (0.1)[1–4] | 4.58 (0.2)[2.5–7.9] | 61.10 (3.8)[24–129] | 67.42 (2.5)[26–102.4] |
| Error         | 117     | 34.44 (7.4)[0.3–509.4] | 4.68 (0.2)[1–10] | 1.74 (0.1)[1–4] | 5.88 (0.2)[2.7–13.1] | 64.14 (2.7)[7–149] | 68.68 (1.6)[20.8–106.8] |
| Semantic      | 90      | 22.02 (5.0)[0.3–402.5] | 4.74 (0.2)[2–10] | 1.74 (0.1)[1–4] | 6.01 (0.2)[2.7–13.1] | 65.19 (3.1)[7–149] | 70.04 (1.8)[30.8–106.8] |
| Omission      | 69      | 37.81 (10.9)[0.3–509.4] | 4.72 (0.2)[1–10] | 1.78 (0.1)[1–4] | 6.00 (0.2)[2.7–10.3] | 63.13 (3.4)[7–125] | 68.04 (2.1)[20.8–106.8] |
| Unrelated     | 27      | 34.63 (18.8)[0.3–509.4] | 4.63 (0.3)[2–8] | 1.78 (0.1)[1–3] | 6.44 (0.4)[3.4–10.3] | 60.22 (6.3)[7–119] | 67.42 (3.1)[39.3–99.2] |

Mean values are reported, with standard error in parentheses and range in brackets. Other error types were not reported here as they did not occur on more than 10 different items. Independent t tests (two-tailed) compared correct responses with the different error categories.

* FWE-corrected \( p < .001 \).
Unrestricted Whole-brain Analyses

The t-contrast of semantic errors > baseline highlighted significant perfusion signal increases in the bilateral ITG or BA 37 and the bilateral postcentral gyrus or BA 1/2/3 (tongue and larynx region). The reverse contrast showed significant perfusion signal decreases in the right precuneus or BA 31 (Table 4).

The one-way ANOVA contrasting semantic errors with their matched correct trials did not reveal any significant perfusion signal increase but showed a perfusion signal decrease in the left posterior inferior parietal lobule (IPL) or caudal BA 39.

The t contrast of omission errors > baseline revealed a significant perfusion signal increase in the right lingual gyrus only (Table 5). The reverse contrast did not reveal any significant cluster.

---

Table 3. Cerebral Regions Showing Significant Signal Perfusion Changes During Speech Errors (All Types Combined)

| Errors Compared to Baseline | Peak MNI (x y z) | z Score | Cluster Size (n Voxels) |
|----------------------------|-----------------|--------|------------------------|
| Total errors > baseline    |                 |        |                        |
| Right ITG\(^a\)            | 56              | −58    | −18                    | 4.96 | 2659 |
| Left precentral gyrus\(^a\) | −62             | 4      | 6                      | 4.72 | 2734 |
| Right postcentral gyrus\(^a\) | 68             | −4     | 18                     | 4.61 | 1513 |
| Left ITG\(^a\)            | −42             | −74    | −16                    | 4.50 | 1093 |
| Left pSTG\(^b\)           | −64             | −30    | 4                      | 4.38 | 486  |
| Left SMG\(^b\)            | −68             | −22    | 24                     | 3.94 | 144  |
| Left IFG / BA 44\(^b\)     | −58             | 12     | 4                      | 4.10 | 81   |
| Left IFG / BA 45\(^b\)     | −52             | 22     | −6                     | 3.89 | 37   |
| Left ACC\(^b\)            | −2              | 24     | 22                     | 3.53 | 24   |
| Right ACC\(^b\)           | 2               | 26     | 20                     | 3.39 | 6    |
| Total errors < baseline    |                 |        |                        |
| Bilateral precuneus\(^a\) | −2              | −60    | 38                     | 4.38 | 1100 |
| Thalamus\(^a\)            | 2               | −22    | −4                     | 4.11 | 488  |
| Left antMTG\(^b\)         | −62             | −8     | −24                    | 4.03 | 134  |
| Left midMTG\(^b\)         | −66             | −20    | −20                    | 3.69 | 56   |
| Total errors > matched correct trials | None\(^a\)\(^b\) | | |
| Total errors < matched correct trials | Left AG\(^b\) | −46 | −62 | 32 | 3.37 | 45 |

Height threshold at p < .001 and pFWE at p < .05.

\(^a\) Whole brain corrected.

\(^b\) Small volume corrected.

---

**Figure 3.** Cerebral region showing significant perfusion signal changes for total errors < matched correct trials (n = 24). Height thresholded at p < .001 and cluster thresholded at five voxels for visualization purposes.
Table 4. Cerebral Regions Showing Significant Signal Perfusion Changes during Semantic Speech Errors ($n = 23$)

|                        | Peak MNI (x y z) | z Score | Cluster Size (n Voxels) |
|------------------------|-----------------|---------|-------------------------|
| Semantic errors > baseline |                 |         |                         |
| Right ITG$^a$          | 50              | -62     | -16                     | 4.87 | 2291 |
| Left postcentral gyrus$^a$ | -66             | -2      | 14                      | 4.48 | 2010 |
| Right postcentral gyrus$^a$ | 68              | -6      | 18                      | 4.24 | 846  |
| Left ITG$^a$           | -42             | -74     | -16                     | 3.95 | 491  |
| Left pSTG$^b$          | -62             | -20     | 10                      | 4.38 | 372  |
| Left IFG / BA 44$^b$   | -58             | 12      | 4                       | 4.06 | 85   |
| Left IFG / BA 45$^b$   | -52             | 22      | -6                      | 4.01 | 48   |
| Left OFG$^b$           | -50             | 24      | -10                     | 3.74 | 25   |

| Semantic errors < baseline |                 |         |                         |
| Right precuneus$^a$       | 2               | -56     | 38                      | 3.74 | 761  |
| Left antMTG$^b$           | -60             | -6      | -24                     | 4.15 | 128  |

| Semantic errors > matched correct trials |                 |         |                         |
| None$^{ab}$                  |                 |         |                         |

| Semantic errors < matched correct trials |                 |         |                         |
| Left posterior IPL$^a$        | -30             | -74     | 16                      | 3.80 | 663  |
| Left AG$^b$                   | -38             | -70     | 18                      | 3.66 | 300  |
| Left pMTG$^b$                 | -58             | -70     | 0                       | 3.38 | 68   |

Height threshold at $p < .001$ and pFWE at $p < .05$.

$^a$ Whole brain corrected.

$^b$ Small volume corrected.

Table 5. Cerebral Regions Showing Significant Signal Perfusion Changes during Omission Speech Errors ($n = 9$)

|                        | Peak MNI (x y z) | z Score | Cluster Size (n voxels) |
|------------------------|-----------------|---------|-------------------------|
| Omission errors > baseline |                 |         |                         |
| Right lingual gyrus$^a$ | 18              | -66     | -12                     | 3.96 | 342  |
| Left pSTG$^b$           | -54             | -34     | 16                      | 3.53 | 16   |

| Omission errors < baseline |                 |         |                         |
| None$^{ab}$                |                 |         |                         |

| Omission errors > matched correct trials |                 |         |                         |
| None$^{ab}$                   |                 |         |                         |

| Omission errors < matched correct trials |                 |         |                         |
| Left dorsal caudate nucleus$^a$        | -12             | 2       | 14                      | 4.57 | 580  |

Height threshold at $p < .001$ and pFWE at $p < .05$.

$^a$ Whole brain corrected.

$^b$ Small volume corrected.
The one-way ANOVA comparing omission errors with their matched correct trials did not reveal any significant perfusion signal increase but showed a perfusion signal decrease in the left dorsal caudate nucleus (Figure 4).

ROI Analyses

The $t$ contrast of semantic errors > baseline highlighted significant perfusion signal increases in the left IFG (both BAs 44 and 45), the left OFG, and the left pSTG (Table 4). The reverse contrast showed a significant perfusion signal decrease in the antMTG.

The one-way ANOVA contrasting semantic errors with their matched correct trials did not reveal any significant perfusion signal increase but showed perfusion signal decreases in the left pMTG and left AG (peaking more posteriorly and ventrally than in the contrast total errors < matched correct trials; Figure 5).

The $t$ contrast of omission errors > baseline revealed a significant perfusion signal increase in the left pSTG (Table 5). The reverse contrast did not show any significant signal decrease in ROIs.

The one-way ANOVA comparing omission errors with their matched correct trials did not reveal any significant perfusion signal increase or decrease in ROIs.

DISCUSSION

The aim of this study was to investigate the neural correlates of naturally occurring speech errors in healthy adults. We found that, independent of type, speech errors occurred more frequently on items whose name was acquired later in life, confirming prior reports in healthy participants and PWA (Brysbaert & Ellis, 2016; Abel et al., 2009). The analyses of perfusion signal changes showed that total speech errors involved a broad set of left-lateralized, frontal, parietal, and temporal regions that were much the same as those engaged during the production of correct responses. In addition, the analyses of specific error types compared to correct trials matched on various psycholinguistic variables revealed significant perfusion decreases in the left pMTG and AG for semantic errors and a significant perfusion decrease in the left dorsal caudate nucleus for omission (i.e., anomic) errors. Surprisingly, comparisons of errors and matched control trials did not reveal significant perfusion changes in either STG or ACC, despite their proposed roles in speech error detection/monitoring in prominent production models.

Overall, Speech Errors Engage a Similar Neural Network to Errorless Production

The regions engaged during total speech error production were very similar to those involved during the production of correct responses, as evidenced by the virtually identical distribution of perfusion signal changes compared to baseline in de Zubicaray and colleagues’ (2015) results on correct naming and the present results for total speech error production. The only significant difference revealed between total speech errors and matched correct trials was in the form of a signal decrease in the left AG, peaking at coordinates corresponding to its middle subdivision (at approximately $z = +30$) highlighted by Seghier (2013), a region linked with processing of thematic or combinatorial associations (see discussion of semantic errors below; de Zubicaray, Hansen, & McMahon, 2013; Seghier, Fagan, & Price, 2010; but see also Noonan, Jefferies, Visser, & Lambon Ralph, 2013).

Contrary to our expectations based on results from previous studies in healthy participants (e.g., Runnqvist et al., 2021; Hansen et al., 2019; Gauvin et al., 2016; see also Meeking & Scott, 2021), we did not observe any significant perfusion signal changes in brain regions proposed to be associated with error monitoring mechanisms during speech production, that is, either the STG or ACC. This suggests external manipulations employed to induce speech errors might be responsible for observations of activity in these regions, rather than naturally occurring speech error processes, as suggested by Meeking and Scott (2021). Of note, the STG has not been observed reliably even in error-eliciting paradigms (e.g., Gauvin et al., 2016) or in many studies investigating the
neural correlates of ToT states (Ubaldi et al., 2022; Huijbers et al., 2017; Shafto et al., 2007, 2010; Kikyo et al., 2001; Maril et al., 2001), and our results are consistent with the only other fMRI study of naturally occurring speech errors by Abel and colleagues (2009) who likewise did not observe STG activation for total errors. It is also worth noting that as strokes rarely result in lesions to ACC (e.g., Mandal et al., 2020), there is little direct evidence to support this region’s involvement in aphasic picture-naming errors as proposed by some computational models (e.g., Nozari, Dell, & Schwartz, 2011). A possible explanation for our failure to detect significant perfusion changes in STG or ACC could be that monitoring accounts have been primarily devised to explain the occurrence of semantic and phonological errors, that is, substitutions, and our total errors also contained omissions. However, we did not observe significant perfusion changes in these regions for the comparison of semantic errors and matched correct trials either.

The activation limited to the AG we observed also differs from Abel and colleagues’ (2009) findings of BOLD signal increases in the right SMA and MFG and the left insula for the identical contrast in their study. Several methodological factors are likely to explain these differences. First and foremost, our coverage did not include the most dorsal cortical slices so excluded part of the SMA, although it did include the MFG and insula. Second, although both studies analyzed comparable numbers of participants (n = 24 with 569 errors here, compared to n = 22 with 319 errors in Abel et al.), Abel and colleagues’ study included both left- and right-handed participants whereas ours were exclusively right-handed. This raises the possibility that Abel and colleagues’ results might reflect some contribution of atypical (i.e., right hemisphere or bilateral) language lateralization in their participants. Third, Abel and colleagues employed a continuous BOLD acquisition that provides a measure of task-related changes in the ratio of oxygenated to de-oxygenated blood, whereas we employed a measure of cerebral perfusion that is a signal more directly related to neural activity and less sensitive to speech-related susceptibility artifacts (see Heim & Specht, 2019). Finally, the interval between presentations of each picture differed between studies (6.5 sec here vs. 7–9 sec in Abel et al.), which might have allowed their participants more time for post-error-monitoring processes.

Speech Error Types Are Associated with Distinct Neural Correlates

Like the participants in Abel and colleagues’ study, our participants did not commit any purely phonological errors. Despite the fact that these errors occur relatively frequently in PWA (e.g., almost 10% of all errors in Schwartz et al., 2012), they are rarely observed in neurotypical participants even after TMS (Krägeloh-Mann et al., 2016; Hernandez-Pavon et al., 2014; Rösler et al., 2014; Liounmis et al., 2012). However, our participants did produce many semantic (i.e., mainly taxonomic) and, to a lesser extent, omission errors, allowing us to investigate their neural correlates.

Semantic Errors Are Associated with Perfusion Signal Decreases in the AG and pMTG

The pattern of perfusion signal changes for semantic errors compared to baseline was very similar to the pattern observed for total errors, which is consistent with the former error type representing more than half of total errors produced. Although no region showed perfusion increases for semantic errors compared to matched correct responses, we found significant clusters in the left AG and pMTG for the reverse contrast. These findings differ from many LSM studies in which lesions in the ATL were associated with semantic paraphasias (Chen et al., 2019; Mirman, Chen, et al., 2015; Mirman, Zhang, et al., 2015; Walker et al., 2011; Schwartz et al., 2009; see also Tochadse et al., 2018). Our findings are instead consistent with other LSM studies reporting the involvement of more posterior temporal and parietal regions. For example, Cloutman and colleagues (2009) reported that the production of semantic paraphasias during picture naming (with or without concomitant deficit in comprehension) was associated with damaged voxels in many left-lateralized regions, including the AG. More recently, Fridriksson and colleagues (2018) and Stark and colleagues (2019) both identified lesions in the left pMTG and AG as correlating with semantic errors during picture naming. Our results are also partially consistent with some DES studies. Both Duffau and colleagues (2005) and Corina and colleagues (2010) reported semantic paraphasias when stimulating the left pMTG. Stimulation applied at the junction between the pSTG and the SMG produced semantic errors in another study (Tate et al., 2014), and some researchers have even shown involvement of all temporal gyri (Perrone-Bertolotti et al., 2020; Miozzo et al., 2017). However, the left AG has not been reported by DES studies as a critical site for semantic errors in production (see Corina et al., 2010).

Several hypotheses could be postulated to explain the deactivation of the left AG and pMTG during the production of semantic errors. One possibility is that it could be related to a failure in semantic control, that is, executive processes that manipulate semantic knowledge in a task- and time-appropriate way (see Jeffries, 2013), as both the pMTG and AG have been implicated in this process (see Lambon Ralph et al., 2017; Davey et al., 2015, 2016; Noonan et al., 2013; Whitney, Kirk, O’Sullivan, Lambon Ralph, & Jeffries, 2011; see also Humphreys, Lambon Ralph, & Simons, 2021). However, we consider this explanation unlikely for a simple task such as picture naming that does not place strong demands on executive control processes. The relation between AG and semantic control has also been more consistently reported in the most dorsal/anterior part of the AG (Kuhnke et al., 2022;
of absence. Nonetheless, the finding of significantly increased pSTG activation is interesting because participants in whom semantic knowledge is assumed to be preserved and accessible.

Omission Errors Are Related to Perfusion Signal Decrease in the Caudate Nucleus

Unlike semantic errors that were found to activate virtually identical left-lateralized, language-related regions as total errors when compared to baseline, the comparison of omission errors to baseline revealed only two clusters with significant perfusion signal increases, in the right lingual gyrus and left pMTG. The reduced network is likely attributable to the comparison being conducted in a subset of participants who did not hear themselves produce a verbal response, unlike correct trials and semantic errors. The peak coordinates (−54, −34, 16) correspond closely to those reported by Meekings and Scott (2021) in their meta-analysis of studies employing altered auditory feedback during production (−54, −30, 10). This might therefore reflect the operation of an auditory error map as proposed by speech motor control accounts (e.g., Kearney & Guenther, 2019).

The analysis contrasting omission errors with matched correct trials revealed a single cluster in the left caudate nucleus that survived whole-brain correction for multiple comparisons. This result contrasts with the more extensive set of cortical regions reported by studies in patients with lesions (i.e., anomie errors in LSM) as well as for TMS and ToT states in neurotypical subjects. Again, this may reflect low statistical power. A possible explanation for the involvement of the left caudate nucleus in omission errors lies in the proposed role for this region in domain-general movement planning and execution (Jankowski, Scheef, Hüppe, & Boecker, 2009; Gerardin et al., 2004; Mendez, Adams, & Lewandowski, 1989; see also Radanovic & Mansur, 2017; Price, 2010). Several studies have suggested that the left caudate nucleus may be involved in the cognitive control of language (e.g., Crinion et al., 2006; see also Friederici, 2006), especially for inhibiting inappropriate responses, as indicated by its activation in functional imaging studies during word interference (Ali, Green, Kherif, Devlin, & Price, 2010) or when inhibiting the unintended language in bilinguals (Hervais-Adelman, Moser-Mercer, Michel, & Golestani, 2015; Abutalebi et al., 2008), and by the occurrence of perseverations upon electrical stimulations or lesions in this region (Nys, van Zandvoort, van der Worp, Kappelle, & de Haan, 2006; Robles, Gatignol, Capelle, Mitchell, & Duffau, 2005; Kreisler et al., 2000; see also Radanovic & Mansur, 2017). Although omission errors can result from many underlying cognitive mechanisms (Chen et al., 2019; Tochadse et al., 2018; see also Dell, Lawler, Harris, & Gordon, 2004, for computational accounts of omission errors), their subcortical locus of activation here suggests that they may have originated from a failure to initiate a correct articulatory–motor response or at inhibiting an unintended (i.e., incorrect) response.

Limitations

The error data reported here were acquired during performance of a continuous picture-naming paradigm designed to elicit a cumulative semantic interference effect in naming latencies (de Zubicaray et al., 2015; Howard et al., 2006). Consequently, although the errors we analyzed occurred on both filler and experimental trials, it is possible that this task configuration influenced the activation patterns we observed for errors. However, it is worth noting that any large series of pictures of everyday objects to be named will invariably involve presenting more than one exemplar from multiple categories (e.g., Abel et al., 2009), which will result in similar interference effects in latencies. For example, Gordon and Cheimariou (2013) retrospectively analyzed naming latencies from a randomized picture-naming task containing a wide variety of items and categories. They found similar interactions between semantic categories and item order resulting in interference effects. Furthermore, although the cumulative semantic interference effect manifests reliably in naming latencies in healthy participants, a similar effect has not been reliably reported for accuracy (Belke, 2013; Navarrete, Mahon, & Caramazza, 2010; Howard et al., 2006).

Most of our methodological choices (e.g., creating sets of matched correct trials for participants with eight errors or more) replicated those of Abel and colleagues’ (2009) earlier BOLD fMRI study, facilitating comparisons with their findings. However, we acknowledge that determining the optimal number of trials needed to elicit...
reliable error-related activation depends on a combination of factors such as experimental paradigm, sample size, effect size, field strength, contrast mechanism (perfusion vs. BOLD signal), signal-to-noise ratio, and so forth. To our knowledge, only one 3-T BOLD fMRI study has investigated this issue, and it showed that error-related activation elicited with a go/no-go task can be reliable with as few as eight trials in 20 participants, although the precise numbers varied across their ROIs located solely in the frontal cortex (Steele et al., 2016). Although our whole-brain FWE corrected results ensure that Type 1 errors are low, we cannot exclude the possibility of Type 2 errors. Hence, the optimal combination of participants/trials to detect error-related activation during picture naming will need to be addressed in future fMRI studies.

Finally, although we matched our sets of correct and error stimuli on a range of variables such as numbers of syllables and phonemes, AoA, lexical frequency, and visual complexity per previous studies, we were not able to match them on other variables known to influence picture naming performance such as name agreement or concept familiarity (e.g., Perret & Bonin, 2019). This is because normative data for these variables were not available for the Howard and colleagues (2006) stimuli that we used. Future studies of error production during picture naming should therefore consider employing stimuli from more recent normative databases that include these variables (e.g., Krautz & Keuleers, 2022).

Conclusion

Overall, our findings indicate that the brain regions activated during speech errors in healthy participants are broadly similar to those activated during errorless production. However, we also found that some regions were relatively less activated, as indicated by reduced perfusion signal, for specific error types compared to correct responses matched for various psycholinguistic variables. These included the left AG and pMTG for semantic errors and the left caudate nucleus for omission errors. Our results also provide little evidence for a role of monitoring mechanisms during naturally occurring speech errors in healthy participants, indicating they may only be strongly engaged when errors are produced as a function of external manipulations. Furthermore, they indicate the production stages/mechanisms responsible for generating speech errors in healthy participants are less variable than those in PWA.

Acknowledgments

This research was supported by Australian Research Council Discovery Project grants (DP1092619 and DP200100127).

Reprint requests should be sent to Angelique Volfart, School of Psychology and Counselling, Faculty of Health, Queensland University of Technology, A Wing, O Block, Kelvin Grove Campus, Brisbane, QLD 4059, Australia, or via e-mail: angelique.volfart@qut.edu.au.

Data and Code Availability Statement

Behavioral data that support the findings of this study are available upon request from the corresponding author. Individual fMRI data are not publicly available because of ethical restrictions, but group-level t maps are available upon request.

Author Contributions

Angelique Volfart: Formal analysis; Visualization; Writing—Original draft; Writing—Review & editing. Katie L. McMahon: Conceptualization; Data curation; Funding acquisition; Investigation; Methodology; Project administration; Resources; Writing—Review & editing. David Howard: Conceptualization; Methodology; Resources; Writing—Review & editing. Greig I. de Zubicaray: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Writing—Original draft; Writing—Review & editing.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the Journal of Cognitive Neuroscience (JoCN) during this period were M(an)/M = .407, W(oman)/M = .32, M/ W = .115, and W/ W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, JoCN, 34:1, pp. 1–3). Consequently, JoCN encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article’s gender citation balance.

Note

1. These five studies involve additions to an identical sample of patients and are thus not fully independent.

REFERENCES

Abel, S., Dressel, K., Kümmerer, D., Saur, D., Mader, I., Weiller, C., et al. (2009). Correct and erroneous picture naming responses in healthy subjects. Neurosci. Lett., 463, 167–171. https://doi.org/10.1016/j.neulet.2009.07.077, PubMed: 19647038

Abutalebi, J., Annoni, J.-M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., et al. (2008). Language control and lexical competition in bilinguals: An event-related fMRI study. Cerebral Cortex, 18, 1496–1505. https://doi.org/10.1093/cercor/bhm182, PubMed: 17947346
Gauvin, H. S., De Baene, W., Brass, M., & Hartsuiker, R. J. (2016). Conflict monitoring in speech processing: An fMRI study of error detection in speech production and perception. *Neuroimage*, 126, 96–105. https://doi.org/10.1016/j.neuroimage.2015.11.037, PubMed: 26608243

Gerardin, E., Pochn, J.-B., Poline, J.-B., Tremblay, L., Van de Moortele, P.-F., Levy, R., et al. (2004). Distinct striatal regions support movement selection, preparation and execution. *Neuroreport*, 15, 2327–2331. https://doi.org/10.1097/00001756-200410250-00005, PubMed: 15640749

Gordon, J. K., & Cheinmariou, S. (2013). Semantic interference in a randomized naming task: Effects of age, order, and category. *Cognitive Neuropsychology*, 30, 476–494. https://doi.org/10.1080/02643294.2013.877437, PubMed: 24499271

Grande, M., Meffert, E., Schoenberger, E., Jung, S., Frauenrath, T., Huber, W., et al. (2012). From a concept to a word in a syntactically complete sentence: An fMRI study on spontaneous language production in an overt picture description task. *NeuroImage*, 61, 702–714. https://doi.org/10.1016/j.neuroimage.2012.03.087, PubMed: 22504766

Halai, A. D., Woolams, A. M., & Lambon Ralph, M. A. (2018). Triangulation of language-cognitive impairments, naming errors and their neural bases post-stroke. *NeuroImage: Clinical*, 17, 465–473. https://doi.org/10.1016/j.nicl.2017.10.037, PubMed: 29159095

Hansen, S. J., McMahon, K. L., & de Zubicaray, G. I. (2019). Neural mechanisms for monitoring and halting of spoken word production. *Journal of Cognitive Neuroscience*, 31, 1946–1957. https://doi.org/10.1162/jocn_a_01462, PubMed: 31418536

Heim, S., & Specht, K. (2019). Studying language with functional magnetic resonance imaging (fMRI). In G. I. de Zubicaray & O. Schiller (Eds.), *The Oxford handbook of neurolinguistics* (pp. 72–95). New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190672027.013.4

Hernandez-Pavon, J. C., Mäkelä, N., Lehtinen, H., Lioumis, P., Huijbers, W., Papp, K. V., LaPoint, M., Wigman, S. E., Dagley, A., Gordon, J. K., & Cheimariou, S. (2013). Semantic interference in a randomized naming task: Effects of age, order, and category. *Cognitive Neuropsychology*, 30, 476–494. https://doi.org/10.1080/02643294.2013.877437, PubMed: 24499271

Krautz, A. E., & Keuleers, E. (2022). LinguaPix database: A megastudy of picture-naming norms. *Behavior Research Methods*, 54, 941–954. https://doi.org/10.3758/s13428-021-01651-0, PubMed: 34378177

Kreibis, A., Godefroy, O., Delaire, C., Debachy, B., Leclercq, M., Pruve, J.-P., et al. (2000). The anatomy of aphasia revisited. *Neurology*, 54, 1117–1123. https://doi.org/10.1212/WNL.54.5.1117

Krieg, S. M., Sollmann, N., Tanigawa, N., Foeschler, A., Meyer, B., & Ringel, F. (2016). Cortical distribution of speech and language errors investigated by visual object naming and navigated transcranial magnetic stimulation. *Brain Structure and Function*, 221, 2259–2286. https://doi.org/10.1007/s00429-015-1042-7, PubMed: 25894631

Kuhnke, P., Chapman, C. A., Cheung, V. K. M., Turker, S., Graessner, A., Martin, S., et al. (2022). The role of the angular gyrus in semantic cognition: A synthesis of five functional neuroimaging studies. *Brain Structure and Function*. https://doi.org/10.1007/s00429-022-02493-y, PubMed: 35476027

Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18, 42–55. https://doi.org/10.1038/nrn.2016.150, PubMed: 27681854

Lambon Ralph, M. A., Moriarty, L., & Sage, K. (2002). Anomaia is simply a reflection of semantic and phonological impairments: Evidence from a case-series study. *Aphasiology*, 16, 56–82. https://doi.org/10.1080/02687040143000448

Lambon Ralph, M. A., Sage, K., Jones, R. W., & Mayberry, E. J. (2010). Coherent concepts are computed in the anterior temporal lobes. *Proceedings of the National Academy of Sciences, U.S.A.*, 107, 2717–2722. https://doi.org/10.1073/pnas.0907307107, PubMed: 20133790

Levitt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, 3, 223–232. https://doi.org/10.1016/S1364-6613(99)01319-4

Lewis, G. A., Poeppel, D., & Murphy, G. L. (2019). Contrasting semantic versus inhibitory processing in the angular gyrus: An fMRI study. *Cerebral Cortex*, 29, 2470–2481. https://doi.org/10.1093/cercor/bhy118, PubMed: 29878066

Lioumis, P., Zhdanov, A., Mäkelä, N., Lehtinen, H., Wilenius, J., Neuvonen, T., et al. (2012). A novel approach for documenting naming errors induced by navigated transcranial magnetic stimulation. *Journal of Neuroscience Methods*, 204, 349–354. https://doi.org/10.1016/j.jneumeth.2011.11.003, PubMed: 22108143

Liu, T. T., & Wong, E. C. (2005). A signal processing model for arterial spin labeling functional MRI. *NeuroImage*, 24, 207–215. https://doi.org/10.1016/j.neuroimage.2004.09.047, PubMed: 15586612

Mandal, A. S., Fama, M. E., Skipper-Kallal, L. M., DeMarco, A. T., Lacey, E. H., & Turkeltaub, F. E. (2020). Brain structures and cognitive abilities important for the self-monitoring of speech errors. *Neurobiology of Language*, 1, 319–338. https://doi.org/10.1162/nol_a_00015, PubMed: 34676371

Maril, A., Wagner, A. D., & Schacter, D. L. (2001). On the tip of the tongue: An event-related fMRI study of semantic retrieval

Volfart et al. 125
failure and cognitive conflict. Neuron, 31, 653–660. https://doi.org/10.1016/S0896-6273(01)00396-9
Meehlings, S., & Scott, S. K. (2021). Error in the superior temporal gyrus: A systematic review and activation likelihood estimation meta-analysis of speech production studies. Journal of Cognitive Neuroscience, 33, 422–444. https://doi.org/10.1162/jocn_a_01661, PubMed: 33076327
Meier, E. L., Sheppard, S. M., Goldberg, E. B., Kelly, C. R., Walker, A., Ubelacker, D. M., et al. (2022). Dysfunctional tissue correlates of unrelated naming errors in acute left hemisphere stroke. Language, Cognition and Neuroscience, 37, 330–347. https://doi.org/10.1080/23737982.2021.1980593, PubMed: 35605076
Mendez, M. F., Adams, N. L., & Lewandowski, K. S. (1989). Neurobehavioral changes associated with caudate lesions. Neurology, 39, 349–359. https://doi.org/10.1212/WNL.39.3.349, PubMed: 2927642
Mizio, M., Williams, A. C., McKhann, G. M., & Hamberger, M. J. (2017). Topographical gradients of semantics and phonology revealed by temporal lobe stimulation. Human Brain Mapping, 38, 688–703. https://doi.org/10.1002/hbm.23409, PubMed: 27635492
Mirman, D., Chen, Q., Zhang, Y., Wang, Z., Faseyitan, O. K., Coslett, H. B., et al. (2015). Neural organization of spoken language revealed by lesion-symptom mapping. Nature Communications, 6, 6762. https://doi.org/10.1038/ncomms7762, PubMed: 25879574
Mirman, D., Zhang, Y., Wang, Z., Coslett, H. B., & Schwartz, M. F. (2015). The ins and outs of meaning: Behavioral and neuroanatomical dissociation of semantically-driven word retrieval and multimodal semantic recognition in aphasia. Neuropsychologia, 76, 208–219. https://doi.org/10.1016/j.neuropsychologia.2015.02.014, PubMed: 25681739
Mirnave, E., Mahson, B. Z., & Caramazza, A. (2010). The cumulative semantic cost does not reflect lexical selection by competition. Acta Psychologica, 134, 279–289. https://doi.org/10.1016/j.actpsy.2010.02.009, PubMed: 20547062
Neta, M., Miezin, F. M., Nelson, S. M., Dubis, J. W., Dosenbach, N. U. F., Schlaggar, B. L., et al. (2015). Spatial and temporal characteristics of error-related activity in the human brain. Journal of Neuroscience, 35, 253–266. https://doi.org/10.1523/JNEUROSCI.1313-14.2015, PubMed: 25581109
Noonan, K. A., Jefferies, E., Visser, M., & Lamber Ralph, M. A. (2013). Going beyond inferior prefrontal involvement in semantic control: Evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. Journal of Cognitive Neuroscience, 25, 1824–1850. https://doi.org/10.1162/jocn_a_00442, PubMed: 23859646
Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. Cognitive Psychology, 63, 1–33. https://doi.org/10.1016/j.cogpsych.2011.05.001, PubMed: 21652015
Nys, G. M. S., van Zandvoort, M. J. E., van der Worp, H. B., Kappelle, L. J., de Haan, E. H. F. (2006). Neuropsychological and neuroanatomical correlates of perseverative responses in subacute stroke. Brain, 129, 2148–2157. https://doi.org/10.1093/brain/awl199, PubMed: 16870885
Okada, K., Matchin, W., & Hickok, G. (2018). Neural evidence for predictive coding in auditory cortex during speech production. Psychonomic Bulletin & Review, 25, 425–430. https://doi.org/10.3758/s13423-017-1284-x, PubMed: 28397076
Perret, C., & Bonin, P. (2019). Which variables should be controlled for to investigate picture naming in adults? A Bayesian meta-analysis. Behavior Research Methods, 51, 2533–2545. https://doi.org/10.3758/s13428-018-1100-1, PubMed: 30066263
Perrone-Bertolotti, M., Alexandre, S., Jobb, A.-S., De Palma, L., Baciu, M., Mairesse, M.-P., et al. (2020). Probabilistic mapping of language networks from high frequency activity induced by direct electrical stimulation. Human Brain Mapping, 41, 4113–4126. https://doi.org/10.1002/hbm.25112, PubMed: 32697353
Price, C. J. (2010). The anatomy of language: A review of 100 MRI studies published in 2009. Annals of the New York Academy of Sciences, 1191, 62–88. https://doi.org/10.1111/j.1749-6632.2010.05444.x, PubMed: 20392276
Price, A. R., Bonner, M. F., Peele, J. E., & Grossman, M. (2015). Converging evidence for the neuroanatomical basis of combinatorial semantics in the angular gyrus. Journal of Neuroscience, 35, 3276–3284. https://doi.org/10.1523/JNEUROSCI.3446-14.2015, PubMed: 25698762
Radianovic, M., & Mansur, L. L. (2017). Aphasia in vascular lesions of the basal ganglia: A comprehensive review. Brain and Language, 173, 20–32. https://doi.org/10.1016/j.bandl.2017.05.005, PubMed: 28570947
Robles, S. G., Gatigopol, P., Capelle, L., Mitchell, M.-C., & Dufau, H. (2005). The role of dominant striatum in language: A study using intraoperative electrical stimulations. Journal of Neurology, Neurosurgery and Psychiatry, 76, 940–946. https://doi.org/10.1136/jnnp.2004.045948, PubMed: 15965199
Rösler, J., Niraula, B., Strack, V., Zdunczyk, A., Schilt, S., Savolainen, P., et al. (2014). Language mapping in healthy volunteers and brain tumor patients with a novel navigated TMS system: Evidence of tumor-induced plasticity. Clinical Neurophysiology, 125, 526–536. https://doi.org/10.1016/j.clinph.2013.08.015, PubMed: 24051073
Runqvist, E., Chanoine, V., Strijkers, K., Pattamadilok, C., Bonnard, M., Nazarian, B., et al. (2021). Cerebellar and cortical correlates of internal and external speech error monitoring. Cerebral Cortex Communications, 2, 100538, https://doi.org/10.1093/texcom/tgab038, PubMed: 34296182
Sakreida, K., Lange, I., Willmes, K., Hein, S., Binkofski, F., Clusmann, H., et al. (2018). High-resolution language mapping of Broca’s region with transcranial magnetic stimulation. Brain Structure and Function, 223, 1297–1312. https://doi.org/10.1007/s00429-017-1550-8, PubMed: 29116426
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. https://doi.org/10.1093/brain/awx300, PubMed: 23176622
Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Brecher, A., Faseyitan, O. K., Dell, G. S., et al. (2011). Neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. Proceedings of the National Academy of Sciences, U.S.A., 108, 8520–8524. https://doi.org/10.1073/pnas.1014935108, PubMed: 21540329
Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Faseyitan, O., Brecher, A., Dell, G. S., et al. (2009). Anterior temporal involvement in semantic word retrieval: Voxel-based lesion-symptom mapping evidence from aphasia. Brain, 132, 3411–3427. https://doi.org/10.1093/brain/awp284, PubMed: 19942676
Seghier, M. L. (2013). The angular gyrus: Multiple functions and multiple subdivisions. The Neuroscientist, 19, 43–61. https://doi.org/10.1177/1073858412440596, PubMed: 22547530
Seghier, M. L., Fagan, E., & Price, C. J. (2010). Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. Journal of Neuroscience, 30, 16800–16817. https://doi.org/10.1523/JNEUROSCI.3377-10.2010, PubMed: 21159592
Shafo, M. A., Burke, D. M., Stamatakis, E. A., Tam, P. P., & Tyler, L. K. (2007). On the tip-of-the-tongue: Neural correlates of
increased word-finding failures in normal aging. *Journal of Cognitive Neuroscience, 19*, 2060–2070. https://doi.org/10.1162/jocn.2007.19.12.2060, PubMed: 17892392

Shafto, M. A., Stamatakis, E. A., Tam, P. P., & Tyler, L. K. (2010). Word retrieval failures in old age: The relationship between structure and function. *Journal of Cognitive Neuroscience, 22*, 1530–1540. https://doi.org/10.1162/jocn.2009.21321, PubMed: 19642890

Stark, B. C., Basilakos, A., Hickok, G., Rorden, C., Bonilha, L., & Fridriksson, J. (2019). Neural organization of speech production: A lesion-based study of error patterns in connected speech. *Cortex, 117*, 228–246. https://doi.org/10.1016/j.cortex.2019.02.029, PubMed: 31005024

Steele, V. R., Anderson, N. E., Claus, E. D., Bernat, E. M., Rao, V., Assaf, M., et al. (2016). Neuroimaging measures of error-processing: Extracting reliable signals from event-related potentials and functional magnetic resonance imaging. *Neuroimage, 132*, 247–260. https://doi.org/10.1016/j.neuroimage.2016.02.046, PubMed: 26908319

Székely, A., & Bates, E. (2000). Objective visual complexity as a variable in studies of picture naming. *Center for Research in Language Newsletter, 12*, 3–33.

Tate, M. C., Herbet, G., Moritz-Gasser, S., Tate, J. E., & Duffau, H. (2014). Probabilistic map of critical functional regions of the human cerebral cortex: Broca’s area revisited. *Brain, 137*, 2773–2782. https://doi.org/10.1093/brain/awu168, PubMed: 24970097

Thirion, B., Pinel, P., Mériaux, S., Roche, A., Dehaene, S., & Poline, J.-B. (2007). Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses. *Neuroimage, 35*, 105–120. https://doi.org/10.1016/j.neuroimage.2006.11.054, PubMed: 17239619

Tochadse, M., Halai, A. D., Lambon Ralph, M. A., & Abel, S. (2018). Unification of behavioural, computational and neural accounts of word production errors in post-stroke aphasia. *Neuroimage: Clinical, 18*, 952–962. https://doi.org/10.1016/j.nicl.2018.03.031, PubMed: 29876280

Ubaldi, S., Rabini, G., & Fairhall, S. L. (2022). Recruitment of control and representational components of the semantic system during successful and unsuccessful access to complex factual knowledge. *Journal of Neuroscience, 42*, 4879–4890. https://doi.org/10.1523/JNEUROSCI.2485-21.2022, PubMed: 35552235

Ullsperger, M., Harsay, H. A., Wessel, J. R., & Ridderinkhof, K. R. (2010). Conscious perception of errors and its relation to the anterior insula. *Brain Structure and Function, 214*, 629–643. https://doi.org/10.1007/s00429-010-0261-1, PubMed: 20512371

Vaughan, J. T., Adriany, G., Garwood, M., Yacoub, E., Duong, T. Q., DelaBarre, L., et al. (2002). Detunable transverse electromagnetic (TEM) volume coil for high-field NMR. *Magnetic Resonance in Medicine, 47*, 990–1000. https://doi.org/10.1002/mrm.10141, PubMed: 11979579

Walker, G. M., Schwartz, M. F., Kimberg, D. Y., Faseyitan, O., Brecher, A., Dell, G. S., et al. (2011). Support for anterior temporal involvement in semantic error production in aphasia: New evidence from VLSM. *Brain and Language, 117*, 110–122. https://doi.org/10.1016/j.bandl.2010.09.008, PubMed: 20961612

Wang, Z., Aguirre, G. K., Rao, H., Wang, J., Fernández-Seara, M. A., Childress, A. R., et al. (2008). Empirical optimization of ASL data analysis using an ASL data processing toolbox: ASLtbx. *Magnetic Resonance Imaging, 26*, 261–269. https://doi.org/10.1016/j.mri.2007.07.003, PubMed: 17826940

Wang, J., Alsop, D. C., Li, L., Listerud, J., Gonzalez-At, J. B., Schnall, M. D., et al. (2002). Comparison of quantitative perfusion imaging using arterial spin labeling at 1.5 and 4.0 Tesla. *Magnetic Resonance in Medicine, 48*, 242–254. https://doi.org/10.1002/mrm.10211, PubMed: 12210932

Whitney, C., Kirk, M., O’Sullivan, J., Lambon Ralph, M. A., & Jefferyes, E. (2011). The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gGyrus. *Cerebral Cortex, 21*, 1066–1075. https://doi.org/10.1093/cercor/bhq180, PubMed: 20851853

Xue, G., Aron, A. R., & Poldrack, R. A. (2008). Common neural substrates for inhibition of spoken and manual responses. *Cerebral Cortex, 18*, 1923–1932. https://doi.org/10.1093/cercor/bhm220, PubMed: 18245044