Title
Orthographic influence on spoken word identification: Behavioral and fMRI evidence

Authors
Chiarello, Christine
Vaden, Kenneth I
Eckert, Mark A

Publication Date
2018-03-01

DOI
10.1016/j.neuropsychologia.2018.01.032

Peer reviewed

HHS Public Access

Author manuscript

Neuropsychologia. Author manuscript; available in PMC 2019 March 01.

Published in final edited form as:

Neuropsychologia. 2018 March ; 111: 103–111. doi:10.1016/j.neuropsychologia.2018.01.032.

Orthographic Influence on Spoken Word Identification: Behavioral and fMRI Evidence

Christine Chiarello[1], **Kenneth I. Vaden Jr.**[2], and **Mark A. Eckert**[2]

[1]University of California, Riverside

[2]Medical University of South Carolina

Abstract

The current study investigated behavioral and neuroimaging evidence for orthographic influences on auditory word identification. To assess such influences, the proportion of similar sounding words (i.e. phonological neighbors) that were also spelled similarly (i.e., orthographic neighbors) was computed for each auditorily presented word as the Orthographic-to-Phonological Overlap Ratio (OPOR). Speech intelligibility was manipulated by presenting monosyllabic words in multi-talker babble at two signal-to-noise ratios: +3 and +10 dB SNR. Identification rates were lower for high overlap words in the challenging +3 dB SNR condition. In addition, BOLD contrast increased with OPOR at the more difficult SNR, and decreased with OPOR under more favorable SNR conditions. Both voxel-based and region of interest analyses demonstrated robust effects of OPOR in several cingulo-opercular regions. However, contrary to prior theoretical accounts, no task-related activity was observed in posterior regions associated with phonological or orthographic processing. We suggest that, when processing is difficult, orthographic-to-phonological feature overlap increases the availability of competing responses, which then requires additional support from domain general performance systems in order to produce a single response.

Keywords

auditory word identification; cingulo-opercular system; literacy; lexical neighborhood effects; orthography

1. Introduction

Reading knowledge influences how speech is processed. Although spoken language is acquired prior to reading, literacy can affect some aspects of speech processing. For example, a word's orthography (spelling pattern) has an influence on how fast and accurately an auditory word is processed (Taft, 2011). Auditory lexical decisions are slower and less accurate for words with inconsistent sound-to-spelling mappings (Ziegler & Ferrand, 1998; Petrova, Gaskell, & Ferrand, 2011). In addition, auditory rhyme judgments are facilitated when word pairs have similar (pie-tie), as compared to dissimilar (rye-tie), spellings (Seidenberg & Tanenhaus, 1979). Likewise, greater phonological priming of

Correspondence: Christine Chiarello, Ph.D. Department of Psychology, University of California, Riverside, Riverside, CA 92521, USA, christine.chiarello@ucr.edu, Phone: (951) 827-5096, Fax: (951) 827-3985.

spoken words is obtained from words with more, than with less, spelling similarity (Chereau, Gaskell, & Dumay, 2007). Orthographic consistency effects have even been reported for task-irrelevant auditory words presented during a nonverbal noise detection task (Perre, Bertrand, & Ziegler, 2011; see also Pattamadilok et al., 2014 for a related finding). Such findings support a beneficial effect of orthography on spoken word processing when word pairs have similar spellings or when potential competitors for a single spoken word are spelled similarly. It appears that orthographic representations are activated, perhaps automatically (Taft et al., 2008; Perre et al., 2011), during speech recognition.

Orthographic consistency studies generally examine the influence of words with rimes (vowel and any terminal consonants) that are spelled similarly or differently (e.g., Ziegler & Ferrand, 1998; Ziegler, Ferrand, & Montant, 2004). Lexical neighborhood research explicitly examines how the processing of a target word is affected by the number of words with similar spellings or pronunciations, considering all of a word's segments[1]. Dense *phonological* neighborhoods have an inhibitory effect on auditory lexical decision (Vitevitch & Luce, 1999; Ziegler et al., 2003), object naming (Sadat et al., 2014), and shadowing (Dirks et al., 2001) tasks. However, the density of a word's *orthographic* neighborhood, that is the number of words sharing all but one letter with the target item, facilitates auditory word recognition in lexical decision and shadowing tasks (Ziegler, Muneaux, & Grainger, 2003). Such research examines the net size of a neighborhood rather than the extent to which orthographic and phonological neighborhoods overlap. Yet visual word recognition findings suggest that the number of such overlapping neighbors, rather than net neighborhood size, affects performance (Adelman & Brown, 2007). In the current investigation we take a similar approach to examine how orthography may influence aural word recognition accuracy by assessing the number of a spoken word's phonological neighbors that are also orthographic neighbors (orthophonic overlap - see Figure 1).

Two theoretical approaches are consistent with a role for orthography in spoken word processing (Petrova, Gaskell & Ferrand, 2011; Taft, 2011). One approach (interactive activation) claims that sublexical units for orthography and phonology mutually influence each other and are jointly activated from speech. When such units converge on the same sublexical and lexical representations, processing is facilitated (Taft, 2011). However, for words with inconsistent spellings, competing inconsistent representations will impair processing. Another approach postulates that orthographic knowledge restructures how phonological information for words is stored, with finer-grained phonological representations for words with consistent spellings (Ziegler, Petrova, & Ferrand, 2008). According to this view, the locus of orthographic influence on speech identification is within the phonological system (Taft, 2011). Thus far, evidence from behavioral studies has been unable to distinguish between these possibilities (e.g., Ziegler, Ferrand, & Montant, 2004; Petrova et al., 2011).

ERP data confirm an orthographic influence on speech recognition, with such effects detectable in waveforms by 300 – 400 ms post-onset (Perre & Ziegler, 2008; Perre et al.,

[1]Density calculations typically include words that differ from a target by a single phoneme, while consistency estimates only include words with identical rimes. The measures of interest in the current study are density estimates.

2009a,b). Across studies, these findings have been interpreted to support either the joint activation view (because orthographic and phonological priming effects differed in scalp topography, Perre et al., 2009a) or the phonological restructuring view (because the source of ERP spelling consistency effects was localized to left temporal-parietal sites associated with phonological processing, Perre et al., 2009b). Functional imaging during aural word recognition tasks could perhaps adjudicate between the theoretical approaches by examining the locus of brain activity associated with orthographic consistency/neighborhood size effects. The phonological restructuring view would predict predominant activity in regions associated with phonology, but not in areas associated with visual word processing. Only one prior fMRI study has explored this issue (Montant, Schön, Anton, & Ziegler, 2011). No ventral occipital-temporal cortex (vOTC) effects were observed when comparing activation for inconsistent to consistent word spelling conditions during an auditory lexical decision task. There was elevated left frontal operculum activity that was interpreted as support for phonological restructuring of the "speech network."

The Montant et al. (2011) frontal results, which included bilateral frontal operculum, anterior insula, and cingulate regions, can also be interpreted as involvement of a performance monitoring cingulo-opercular network that is typically engaged during challenging listening tasks (Eckert et al., 2016). Because this pattern of brain regions is recruited across verbal and nonverbal tasks (Dosenbach et al., 2008), it has also been interpreted as a domain general self-regulation system (Kelley, Wagner, & Heatherton, 2015) to guide behavior and optimize performance (Eckert et al., 2016). This performance monitoring and response selection perspective is generally consistent with the idea that orthographic neighborhood effects can occur at a decision or response selection stage of speech processing (Pattamadilok et al., 2007). Further support for this perspective may be seen in the lexical selection difficulty for orthographically inconsistent compared to consistent words that produces an effect at an FCz electrode location over medial-frontal cortex (Pattamadilok et al., 2009) where elevated theta power has been linked to an elevated threshold for response selection (Cavanagh et al., 2011).

Two other sources of data are relevant to the role of orthography in aural word recognition. Dehaene and colleagues (Dehaene et al., 2010; Dehaene, Cohen, Morais, & Kolinsky, 2015) have investigated how the acquisition of literacy affects the brain systems used to process spoken language. In individuals who became literate as adults, listening to spoken sentences or performing auditory lexical decisions was associated with increased activation bilaterally in the planum temporale (but with leftward asymmetry), as compared to matched illiterates (Dehaene et al., 2010). In addition, for auditory lexical decisions, literacy also was associated with increased activation in a region of the left inferior temporal cortex (the putative visual word form area), which the authors argue indicated recruitment of an orthographic code from speech. Similar findings were obtained when comparing child readers to age-matched non-readers during auditory sentence listening (Monzalvo & Dehaene-Lambertz, 2013). In that study, enhanced activation of the planum temporale was observed in 6-year-old readers compared to pre-readers. However, activation of the visual word form area was only increased in 9-year-old, compared to 6-year-old readers, implying that greater reading experience is required for literacy effects on speech in this region (Monzalvo & Dehaene-Lambertz, 2013). Dehaene and colleagues (2010, 2015) conclude from such studies that reading acquisition induces both restructuring of phonological representations in the planum temporale as well as activation of orthographic codes from speech.

Pattamadilok et al. (2010) applied transcranial magnetic stimulation (TMS) separately to the left supramarginal gyrus (SMG; a region associated with phonological processing) and to the left vOTC during auditory lexical decision. These stimulation sites were chosen based on independent experiments demonstrating disruption to phonological and orthographic processing, respectively. Stimulation to the left SMG, but not to the left vOTC, eliminated the response time benefit for orthographically consistent words. These authors propose that the SMG forms part of a “clean up” circuit that functions to aid perception in the face of noisy or distorted input. As reading is acquired, it provides an additional mapping onto phono-articulatory codes, strengthening their representations, and reducing the amount of clean up processing required to resolve the auditory input. Pattamodilok et al. (2010) conjecture that disruption of SMG clean up operations reduces or eliminates the processing advantage for orthographically consistent words. This is a novel interpretation, based on inhibitory transcranial stimulation of SMG, and is consistent with claims that orthographic influences may function to stabilize transient acoustic information (Ziegler, et al. 2008). The Pattamodilok claim might predict that under noisy conditions, orthographic influences on speech identification should be enhanced and associated with increased “clean up” SMG activity when stimuli are presented in relatively poor signal-to-noise conditions. In the current study this prediction was tested by aurally presenting words at two SNR levels relative to a continuous multi-talker babble.

To summarize, there is ample behavioral evidence for the role of orthography in spoken word processing, but such data have proven to be insufficient to adjudicate between competing theoretical interpretations. As noted earlier, behavioral and ERP data are equally consistent with the conjoint activation and phonological restructuring views. Extant neural data are intriguing, but not fully convincing. On the one hand, comparisons of literate vs less or non-literate persons suggest that auditory speech processing is associated with activation of ventral occipito-temporal cortex only after reading exposure (Dehaene et al., 2010, 2015). This finding is consistent with the view that orthographic representations are contacted from speech input. On the other hand, orthographic consistency effects were not observed in this region in fMRI auditory lexical decision (Montant et al., 2011), nor were they eliminated by TMS (Pattamadilok et al., 2010). There is some evidence favoring the phonological restructuring explanation, although there is a lack of consensus on which regions may subserve such processing as the left IFG (Montant et al., 2011), SMG (Pattamodilok et al., 2010), and planum temporale (Dehaene et al., 2010, 2015) have each been implicated, albeit using different methods and subject groups. Moreover, some of the fMRI results may be consistent with performance monitoring effects. Certainly it is possible that reading knowledge both restructures phonological representations and also permits activation of orthography from speech. However, we are far from understanding the consequences of literacy on the neural systems recruited for speech.

The current fMRI investigation attempted to address this issue using a slightly different approach. First, our metric of orthographic influences on aural word identification considered the overlap between the entire phonological and orthographic neighborhoods for target stimulus words, rather than dichotomizing a word's spelling consistency based only on the item's terminal segments[2]. This provides a continuous measure of orthographic influence across the entire spoken word and hence may be a more sensitive index than the consistency dichotomy. Because words were presented aurally, not visually, we examined phonological neighborhood density and effects related to orthographic overlap within the phonological neighborhood, rather than orthographic density. Second, to-be-identified words were presented with multi-talker babble using two signal-to-noise ratios (SNR) in order to vary speech intelligibility. This may represent a somewhat more natural listening context, as real world speech perception rarely occurs under optimal auditory conditions. Prior studies have indicated that more challenging listening conditions enhance neighborhood or consistency effects (Taler et al., 2010; see also Pattamodilok, DeMorais & Kolinsky, 2011).

Multi-talker babble was used in the current study, which is thought to activate phonological neighbors because this noise contains speech segments with varying intelligibility. These segments can provide evidence for some alternatives to the target word. For example, when the target word BITE is presented in multi-talker babble that contains an intelligible word TIE, then activation for the competitor TIGHT would be predicted to increase. A target word with fewer neighbors is less likely to be misrecognized due to spurious competitor activations, while a target from a dense neighborhood is more likely to have competitors that align with evidence from the signal and noise. The SNR manipulation in the current study also allowed us to examine a proposed "clean up" role for the SMG (Pattamodilok et al., 2010). Since greater phonological neighborhood density reduces aural word recognition, anything that strengthens the activation of phonological neighbors (e.g., orthographic overlap) should affect word identification to an even greater extent. This view predicts that high orthographic-to-phonological overlap produces increased SMG activation under more challenging listening conditions. Further, we might expect a domain-general effect if the neighborhood overlap results were observed in the more challenging SNR condition and involved activation of cingulo-opercular regions that support performance monitoring.

The current investigation tested predictions about the impact of orthographic- phonological overlap on word recognition and brain activity during an auditory word identification task. We predicted that word identification would be impaired for words with greater orthographic-to-phonological overlap when presented in the low SNR condition, in part based on evidence that high neighborhood density negatively impacts word recognition in more challenging SNR conditions (Taler et al., 2010). Words with lexical features that increase task difficulty (e.g., high orthographic-to-phonological overlap, low word frequency, or high phonological neighborhood density) should therefore elicit increased activity in performance monitoring cingulo-opercular regions. Alternatively, elevated activity across SNR conditions with increasing orthographic overlap might be expected due to a phonological restructuring mechanism (Montant et al., 2011). This view would predict orthographic overlap effects in SMG regions implicated in a "clean up" mechanism involving orthography to aid in the identification of aurally presented words (Pattamodilok

[2]Ziegler & Ferrand (1998) also computed a continuous measure of orthographic overlap, but then used this measure to create a dichotomous consistency variable used in the experimental design and analysis.

et al., 2010). We also predicted that words with high orthographic-to-phonological overlap would produce increased activity in occipito-temporal regions that are responsive to orthography (Dehaene et al., 2010, 2015). The performance monitoring predictions appear to be supported by our results.

2. Method

2.1. Participants

Thirty-seven (20 female) adult native English speakers participated (mean age = 40.9 yrs, range 19.9 – 78.8 years). All reported no difficulty learning to read. Their mean Edinburgh handedness questionnaire score was +68.5 [from a possible range of −100 (strongly left-handed) to +100 (strongly right-handed; Oldfield, 1971)]. The participants had an average of 15.9 years of education, an average socioeconomic status of 50.1 (possible range of 8 to 66; Hollingshead, 1983), and reported no history of neurological or psychiatric events. Their mean full scale IQ was 116.4 (Wechsler Adult Scale of Intelligence, Wechsler, 1999). Participants had mean pure tone thresholds < 21.5 dB HL from 0.25 kHz to 8 kHz (best ear, Madsen OB922 audiometer and TDH-39 headphones), no more than 14.5 dB differences between mean pure-tone thresholds for the right and left ears, and normal immittance measures. All of the participants demonstrated normal hearing below 2 kHz, with a sloping hearing loss for higher frequencies, which was largely driven by the older participants. Informed consent was obtained in compliance with the Institutional Review Board at the Medical University of South Carolina, and experiments were conducted in accordance to the Declaration of Helsinki.

2.2. Stimuli

The word recognition task included 120 monosyllabic consonant-vowel-consonant words recorded by a male speaker (Dirks et al., 2001) that were presented through Sensimetrics piezoelectric insert ear phones. Words were not repeated across intelligibility conditions in order to avoid interactions between word intelligibility and priming or memory effects. The signal to noise ratio (SNR) of the words was manipulated by presenting a continuous multi-talker babble recording at a constant level of 82 dB SPL, and words at either 92 dB SPL (+10 dB SNR) or 85 dB SPL (+3 dB SNR). The multi-talker babble recording was originally prepared as part of the SPIN test (Kalikow et al., 1977) and consists of twelve talkers, which results in energetic masking equivalent to steady state noise at the same SNR (Miller, 1947; Carhart et al., 1969; Wilson et al., 2012). The level of target words and multi-talker babble were calibrated separately in the scanner control room prior to each scanning session.

Neighborhood and word frequency information[3] for each auditory stimulus word was acquired from the English Lexicon Project (ELP) website (http://elexicon.wustl.edu) in order to obtain phonological and orthographic neighborhoods from the same corpus. For each word, a list was generated of phonological and orthographic neighbors, and words that overlapped across these neighborhoods were identified (see Figure 1). The log frequencies

[3]Log word frequency obtained from an auditory corpus (http://www.ugent.be/pp/experimentele-psychologie/en/research/documents/subtlexus) was strongly correlated with the log ELP word frequency, $r = .88$ across the word list.

for each of these orthographic-to-phonological overlap words were determined and then summed. For stimulus words with alternate spellings (e.g., fair/fare), the orthographic neighbors for both spellings were obtained and included in the overlap calculation. The log frequencies for all phonological neighbors were also obtained and then summed. The summed overlap frequencies were then divided by this value to yield the orthography-to-phonology overlap ratio (OPOR). OPOR values for stimulus words ranged from 0 to .796 (mean = .362). Correlations were observed between OPOR and phonological neighborhood density (PND) values [across SNR conditions: $r = 0.32$, $p < 0.001$; +3 dB SNR words: $r = 0.26$, $p = 0.04$, +10 dB SNR words: $r = 0.37$, $p = 0.003$], although the results of our control analyses indicate the PND did not account for the OPOR effects described below.

2.3. Experimental procedure

The word recognition experiment and scanning protocol (sparse acquisition with TR = 8.6 sec) used in the current study were also used in Vaden et al. (2013 in Vaden et al. (2015). Word recognition was assessed using two epochs of 60 trials during which the words were presented in multi-talker babble (+3 or +10 dB SNR conditions) using a sparse sampling image acquisition (Figure 2). This sparse sampling design resulted in an inter-trial interval of 8.6 seconds so that the word stimuli could be presented in the absence of scanner noise, enabling better control over the SNR. Each SNR condition occurred for 4–6 consecutive trials to limit predictability for the onset of the next SNR condition block. Participants were instructed to respond vocally, repeating each word aloud or to say “nope” if the word was not recognized. No performance feedback was provided during the experiment. A visual prompt (“get ready”) was displayed to cue participants to the start of word recognition epochs. In addition, a crosshair changed color from white to red to cue participants about when to respond. Participants viewed the projector screen through a periscope mirror. The experimental design consisted of two word recognition epochs and three rest intervals at the beginning, middle, and end of the experiment.

2.4. Word recognition scores and analyses

Two raters listened to participant responses during the experiment, scoring responses as correct only if the participant repeated the word exactly as presented. Discrepant ratings (4.39%) were resolved based on audio recordings from an MRI-compatible microphone (Magnetic Resonance Technologies Inc.). The analysis omitted 2.7% of all trials based on unintelligible or missing responses, and “nope” responses (1.1%) were scored as incorrect. Logistic regression analyses were performed using General Linear Mixed Models (GLMM) to determine whether the likelihood of correct word identification on each trial was predicted by OPOR, SNR, and their interaction (R statistics software, lme4 package, version 3.3.1). Control analyses were performed to ensure that word frequency and phonological neighborhood density did not account for effects of OPOR on the likelihood of correct word identification.

2.5. Image acquisition, preprocessing, and analyses

A Siemens Tim Trio 3T scanner and 32-channel head coil were used to collect structural and functional images at the Medical University of South Carolina Center for Biomedical Imaging. T1-weighted whole brain structural images were acquired in 160 slices with a

256×256 matrix, TR = 8.13 msec, TE = 3.7 msec, flip angle = 8°, slice thickness = 1.0 mm, and no slice gap. A T2* weighted, single shot echo-planar imaging sequence was used to acquire 180 whole brain functional images (36 slices with 64×64 matrix, TR = 8.6 sec, TE = 35 msec, acquisition time: TA = 1647 msec, slice thickness = 3.0 mm, gap = 0, sequential order, GRAPPA-parallel imaging with acceleration factor = 2). Each functional image had 3.0 mm isomorphic voxels.

A study-specific template (Avants & Gee, 2004; Vaden et al., 2012) was created from the structural T1-weighted images using the Advanced Normalization Tools ANTS version 2.1; Avants, Tustison, & Song, 2011). The ANTS-derived normalization parameters for the T1-weighted images were used to spatially transform each subject's co-registered functional images into a study-specific template space. Voxel coordinates with significant peak effects were converted into MNI space by normalizing the study-specific template to the MNI template with ANTS and applying the resultant deformation parameters to the peak voxel coordinates in study specific space.

SPM8 (http://www.fil.ion.ucl.ac.uk/spm) was used to realign, unwarp, and co-register functional images to corresponding structural scans from each individual prior to normalization and prior to smoothing with a Gaussian kernel (FWHM = 8 mm). Global mean BOLD contrast fluctuations were assumed to be of no interest and were de-trended from the pre-processed functional images with a linear model of the global signal (Macey et al., 2004). Furthermore, an algorithm detailed in Vaden, Muftuler, & Hickok (2010) was used to identify functional images with voxel or volume intensities that exceeded 2.5 standard deviations from the mean time series intensity. An average of 4.50% of the functional images contained outlier noise, and these volumes were controlled for in the participant or first-level General Linear Model (GLM) analyses using binary nuisance regressors. Four nuisance vectors that quantified head position and head motion were entered into the GLM as covariates of no interest (Kuchinsky et al., 2012; Wilke, 2012). Specifically, the Pythagorean Theorem was used to calculate the difference in head position from the middle of the time series (rotation and translation) and trial-level head movement (rotation and translation) based on the six head position vectors output from the SPM motion correction (https://www.nitrc.org/projects/pythagoras). The rationale for this approach to enhance statistical correction for motion artifacts is that SPM motion correction records the total change in head position with six redundant vectors, which do not quantify trial-level head movements.

Voxel-level GLMs were performed to test predictions about the degree to which OPOR was related to brain activity in cingulo-opercular, SMG, and vOTC regions. The GLM included events with timing information that was convolved with the hemodynamic response function to predict BOLD changes during task trials for which words were presented in babble in each of the SNR conditions, trials that presented babble without words to identify, and transition trials at the beginning of each epoch to control for orienting effects. Model parameters for the word recognition trials included word identification (correct/incorrect), and lexical characteristics [OPOR, phonological neighborhood density (PND), word frequency (WF)] that were used to investigate effects within, across, or differing by SNR condition. To conservatively estimate effects of OPOR or PND, two GLMs were performed where OPOR or PND were entered as the last parameter in the model (Mumford et al., 2015). For example, OPOR was the last variable entered in the GLM when examining the degree to which OPOR unique variance related to variance in brain activity that was not accounted for by the other parameters. Thus, we were able to characterize the specific effects of OPOR in each SNR, while controlling for variance that could be related to task performance, PND, and WF.

Group-level single sample t-tests were performed on the contrast maps from the subject-level GLMs and each also included a control covariate for age. For these voxel-based group-level analyses, a combined voxel statistic and cluster size approach determined significant effects (Friston et al., 1994). Permutation testing (10,000) was performed to control for multiple comparisons and identify significant effects ($Z = 3.09$, $p < 0.001$ peak threshold) using a method that limits false positive error rates (Eklund et al., 2014, 2016; https://github.com/wanderine/BROCCOLI).

Regions of interest analyses were performed to determine the extent to which significant effects of OPOR were present in SMG, vOTC, and inferior frontal gyrus locations that have been identified in previous studies as having a potential relation with orthographic consistency. The SMG region of interest (MNI: −51, −31, 26) was selected because of evidence that TMS interfered with an orthographic consistency effect for this region, which had been targeted because it was responsive during a rhyme judgment task (Pattamadilok et al., 2010). The vOTC (MNI: −41, −56, −20) region of interest was also based on Pattamadilok et al. (2010) results where words elicited increased activity relative to a fixation condition, generally consistent with orthographic effects on vOTC activity (e.g., Dehaene et al., 2010; Grill-Spector et al., 2001). Finally, we also included left inferior frontal gyrus (IFG) regions of interest (MNI: −42, 20, 0; −50, 4, 8) where Montant et al. (2011) suggest phonological restructuring can occur (although this region also activates due to task difficulty in speech recognition tasks, Vaden et al., 2013). These coordinates were transformed into the space of the ANTS study-specific template. A 10 mm diameter sphere was then created around each of the coordinates so that average beta value for the OPOR contrasts could be obtained for each participant. Bonferroni correction was performed for the 12 one-sample t-tests (4 regions by 3 contrasts) to determine the extent to which there were significant OPOR effects in these regions of interest. The contrasts examined OPOR effects within each SNR as well as the difference between the two SNRs.

3. Results

3.1. OPOR Predicts Word Identification

Logistic regression demonstrated that word identification was significantly higher for the +10 dB SNR condition (89.9% correct) compared to the +3 dB SNR condition (64.6% correct), $Z = 17.90$, $p < 0.001$. Across conditions, word identification was lower with increasing age, $Z = -2.91$, $p < 0.005$. Model-testing showed that age × OPOR and age × SNR × OPOR interactions did not significantly contribute to model fit, so those were not included in the final GLMM. A significant SNR × OPOR interaction was observed, $Z = 7.16$, $p < 0.001$, which was due to lower accuracy for higher OPOR words in the +3 dB SNR condition ($Z = -10.97$, $p < 0.001$). Word identification was not significantly related to

OPOR in the +10 dB SNR condition, although there was a trend for higher overlap to facilitate performance ($Z = 1.66$, $p = 0.10$). Furthermore, the control analyses determined that PND, WF, and mean tone thresholds (0.2 to 8 kHz, best ear) did not account for the OPOR effects. In summary, these results are consistent with the prediction that higher OPOR in a difficult listening condition can limit word identification.

3.2. OPOR Predicts BOLD Contrast: Whole Brain Analyses

There were no regions sensitive to the OPOR variable that did not interact with SNR. Significant SNR interactions with OPOR were observed for cingulo-opercular activity (permutation-corrected cluster $p_{FWE} < .05$: Table 1, Figure 3). The SNR interaction effect was due both to elevated activity in the +3dB SNR condition, and reduced activity in the +10 dB SNR condition, for higher orthophonic ratio words. These effects were independent of word recognition errors (see description above of GLM), and occurred where there were also significant effects of SNR on cingulo-opercular activity (e.g., SNR: Peak-Z ≥ 5.49: left IFG −47, 20, 6; right IFG 33, 27, 5; paracingulate −3, 34, 39; shown in Supplemental Figure 1). The OPOR effect was also independent of PND and word frequency when these variables were included in the model (Table 1; additional findings for control variables reported in Table S1), and did not depend on whether or not the age covariate was included in the model. In addition, the left vOTC and left SMG did not exhibit significant associations with OPOR within or between SNR conditions for this whole brain analysis using family-wise error correction. These findings are consistent with the performance monitoring account, but do not provide evidence for differential involvement of either phonological or visual-orthographic systems.

3.3. OPOR Predicts BOLD Contrast: *A Priori* Regions of Interest

There were no significant effects in the left SMG and left lateral IFG regions of interest for the OPOR contrasts (Table 2). Consistent with the voxel-based results, there were significant associations between OPOR and the IFG/anterior insula (AI) region of interest between SNR conditions. The left IFG/AI region overlapped spatially with cingulo-opercular cortex, so it was not surprising that it demonstrated a consistent pattern of significant effects.

4. Discussion

The current study obtained both behavioral and neuroimaging evidence for orthographic influence on auditory word identification. Examining orthographic influences using a continuous variable such as the orthographic-to-phonological overlap ratio (OPOR), while controlling for partially correlated lexical variables in linear mixed models, enabled us to account for variation due to words as well as speakers, as has been recently recommended (Sadat et al., 2014). Using a continuous neighborhood-based measure of orthographic-to-phonological overlap, words with higher overlap had lower identification rates in the challenging +3 dB SNR condition. Similarly, BOLD contrast varied with OPOR bilaterally in inferior frontal, insula, and anterior cingulate regions, increasing under high levels of noise and decreasing under more favorable listening conditions. The findings could not be attributed to variations in potentially correlated effects of word frequency or phonological neighborhood density. These results lend additional weight to the conclusion that orthographic knowledge, obtained via reading experience, alters how auditory language is processed. However, as we discuss below, the current results did not support anatomical predictions derived from either phonological restructuring or orthographic activation models of speech processing. Rather, the orthographic influence on speech that we observed had a more indirect effect by modulating the recruitment of more domain general performance monitoring regions.

4.1 Influences of Orthographic-to-Phonological Overlap

The effects of OPOR on word identification can be understood within an interactive activation account of neighborhood effects. Encountering a word will activate neighborhoods of memory representations that share phonological and/or orthographic features (see Figure 1). Representations that receive activation from both systems (overlap items) will be enhanced relative to those that cannot benefit from the overlap. Under easy listening conditions (high SNR), the target word will likely receive the highest level of activation relative to competitors that share features, and may receive some support from sublexical features shared with neighbors (Ziegler et al., 2003). However, more challenging listening conditions (lower SNR) will provide weaker support for the target representation relative to competitors. Furthermore, some of the competitors may receive spurious activation from the background babble, making them strong candidates for response selection. Because words with higher orthophonic ratios will have more of these stronger competitors, word identification will be more error-prone. Our behavioral findings are consistent with this interpretation, at least for the +10 dB SNR condition.

However, one might have expected higher OPOR values to facilitate performance under easier listening conditions given prior studies of orthographic influence on spoken word processing (e.g., Chereau et al., 2003; Ziegler et al., 2003). At best, we only observed a nonsignificant trend in that direction. The fact that, even in the easier listening condition, words in the present study were experienced with some background noise may have contributed to the absence of measurable facilitation. This interpretation receives some support from a study (Pattamadilok et al., 2011), whose findings might at first blush seem contradictory to those presented here. Pattamadilok et al. compared a no noise condition to a condition with Gaussian white noise yielding an SNR of +12.5 dB, and observed orthographic facilitation under noise. The current investigation compared two multi-talker babble noise conditions with SNRs of +10 (low noise) and +3 (high noise). Hence the noise condition of Pattamadilok et al. and the low noise condition of the current study are more comparable, and this is supported by the accuracies obtained in these conditions in each study (Pattamadilok et al., approx. 90.3% correct; current study 89.9% correct). Thus, although Pattamadilok obtained orthographic facilitation under noise, our low noise findings are not inconsistent with their result (marginal facilitation and reduced BOLD). However, we are cautious about making direct comparisons across studies with both different orthographic measures and different noise manipulations. Clearly additional studies are needed to address how SNR manipulations can affect the role of lexical factors in aural word recognition. Our behavioral and fMRI findings of OPOR X SNR interactions imply that lexical influences on speech identification may depend on the intelligibility of various listening environments.

4.2 Functional Anatomy of OPOR

Our fMRI findings provide converging evidence for orthophonic neighborhood effects on spoken word identification. BOLD contrast increased with OPOR at the more challenging SNR, and decreased with OPOR under more favorable SNR conditions. These results could not be attributed to potentially correlated effects of phonological density or word frequency. They were also independent of performance effects (whether an individual response is correct or not), ruling out an interpretation that they reflected error-related response processes. Although we obtained robust OPOR effects, the regions responsive to this variable provide no evidence to support either an orthographic or a phonological locus. Neither the whole brain nor the ROI analyses detected OPOR-related brain activity in occipito-temporal regions associated with orthographic processing, consonant with the results of Montant et al. (2011). In addition, activity in SMG or other areas associated with phonological processing did not exhibit strong associations with OPOR. We did observe a left IFG BOLD response that was sensitive to orthographic-to-phonological overlap, similar to the region Montant et al. (2011) found to respond to orthographic consistency in an auditory lexical task. However, rather than indicating orthographic modification of the speech network (Montant et al., 2011), the activity we observed within the left IFG was clearly embedded within a larger opercular region of activity that also included the anterior insula and dorsal cingulate. As discussed further below, the current findings implicate a more domain-general performance monitoring system that is sensitive to word difficulty from many sources, including that induced by orthographic-to-phonological overlap. Such a system would be globally recruited whenever task difficulty increases, regardless of whether a response on a given trial was correct or not.

It may be useful to consider further why left temporal-parietal and vOTC regions, that are consistently associated with phonological and orthographic processing across a variety of tasks, did not evidence involvement when neighborhood effects were varied in the current study. If lexical and sublexical features are represented in a highly distributed manner across numerous units within a large region, it may be difficult to observe mass-univariate changes in the BOLD signal related to subtle variations in hypothesized neighborhood effects (see Binder et al., 2005; Protopapas et al., 2016 for related arguments). If similar distributed units are weakly activated for a variety of word inputs, neighborhood-type contrasts may not reveal differential neural activity. This does not invalidate neighborhood models as conceptual frameworks for word representation and processing, but it does suggest a potential mismatch in granularity when attempting to instantiate such models in the brain.

We also note that the current results document yet one more unsuccessful attempt to adjudicate between the orthographic activation and phonological restructuring accounts of orthographic influences on auditory word processing (Taft, 2011). Prior behavioral, ERP, and neuroimaging findings have not consistently supported one account over the other, and the present findings supplement this consensus. Although the cumulative data provide substantial evidence that reading knowledge affects auditory word processing, neither theoretical proposal appears superior. This situation suggests that current experimental designs and theoretical frameworks may be inadequate to address the question (orthographic vs phonological locus). To better understand how reading knowledge influences auditory language processing, perhaps neuroimaging studies should explore lexical variables that do not elicit performance differences.

Our findings do support the claim that enhancement of orthographic influences on speech identification under noisy conditions may involve “clean up” operations (Pattamadilok et al., 2010) to recover the identity of the presented word. However, the SMG and other putative phonological regions did not exhibit strong evidence for such an effect. Rather, as we discuss below, the current results implicate the recruitment of more domain-general regions to monitor performance. This may perhaps result in the signaling to other brain regions to initiate greater top-down control to resolve conflict between competing response candidates. It may be important to note here that evidence for a “clean up” mechanism was obtained with TMS rather than fMRI (Pattamadilok et al., 2010). TMS could have widespread effects on the function of regions connected to SMG, including inferior frontal regions where we observed significant associations between OPOR and BOLD contrast. Future functional connectivity studies designed to characterize differences in network structure between OPOR conditions might reveal that SMG network structure can be modified by orthographic manipulations.

4.3 Cingulo-Opercular System

Robust effects of OPOR were obtained in several prefrontal regions comprising the cingulo-opercular network (Dosenbach et al., 2006). This system has been implicated in a variety of performance monitoring functions including conflict monitoring (Botvinick et al., 2001), domain general task control (Sestieri et al., 2014), and processing of ambiguous stimuli (Neta, Schlaggar, & Petersen, 2014). Our findings complement this research by demonstrating that this system is variably recruited depending on the extent of overlap between a word’s orthographic and phonological neighborhoods, and the acoustic quality of the listening environment. In the more favorable +10 dB SNR listening condition, when the target word would be much more accessible than competitors, cingulo-opercular activation decreased for high overlap words. While OPOR did not significantly predict word identification accuracy in the +10 dB SNR, effects related to correct/incorrect responses (performance monitoring) as well as response uncertainty and conflict have been observed in cingulo-opercular regions for a broad range of perceptual tasks, including speech recognition (Eckert et al., 2016). Thus, the OPOR effect for words in the +10 dB SNR condition could reflect relatively greater demands or more response uncertainty for low OPOR words. Conversely, in the noisier +3 dB SNR condition, when greater overlap would strengthen accessibility of competitors, cingulo-opercular activity increased for words with higher OPOR. To the extent that OPOR and SNR modulate the accessibility of potential response competitors, stimulus ambiguity and the demand for conflict monitoring should increase. In sum, our fMRI findings are readily interpretable within current views of the function of the cingulo-opercular system. They further indicate that subtle variations in bottom-up linguistic features (extent of orthographic-to-phonological neighborhood overlap) can modulate the recruitment of top-down monitoring systems, at least when response selection is required.

4.4 Limitations and Future Directions

The sparse sampling fMRI procedure used here may well have had an impact on our findings. Participants were cued to respond only within a specific interval and response times could not be used as an additional index of processing, as is typical in behavioral investigations. Another consideration is that, on some trials participants would have to withhold their response while waiting for the cue. Such temporary response inhibitions may well have influenced the brain systems recruited to perform the task. The sparse sampling method may also have limited sensitivity to effects in brain regions where the hemodynamic response is more variable across subjects, or is maximal on either side of the 5.5s sampling of the response. It will be important to investigate the effects of OPOR across a variety of tasks, listening conditions, and imaging paradigms in order to assess the generalizability of the current results.

Further, we did not examine the nature of the errors participants made. Our interpretation predicts that, under challenging listening conditions, high OPOR words should tend to be misheard as competitors, with words sharing features produced as erroneous responses. Competitors with high orthographic-to-phonological overlap should be produced more frequently than other phonological neighbors. More detailed examination of error responses can be addressed in subsequent studies. In addition, because neighborhood-based effects on speech processing may be attributed to alterations across highly distributed representations, techniques such as multi-voxel pattern analysis may be more sensitive than the mass univariate methods we used. Alternatively, perhaps many more trials are necessary to observe effects in temporal-parietal or vOTC regions than to observe the robust frontal responses in the current study. Finally, word recognition was lower in the older adults who were likely to have experienced declines in the brain regions that were the focus of this study. Although age did not interact with OPOR to affect word identification, it is possible that the inclusion of potentially noisier data from older adults may have obscured a positive OPOR influence in the +10 dB SNR condition or results from the region of interest analyses. Because cingulo-opercular regions undergo consistent structural declines with age (Fiell et al., 2009), the OPOR effects would have been likely underestimated by including older adults. However, we have observed that sublexical frequency effects demonstrate little age-related variance in left inferior frontal cortex activity (Vaden et al., 2011a; 2011b), suggesting that some lexical manipulations have relatively consistent effects across the lifespan. Moreover, vocabulary knowledge is relatively stable across the adult lifespan, despite evidence of age-related structural declines in language pathways (Teubner-Rhodes et al., 2016).

4.5 Conclusions

Literacy impacts how speech is processed. The language system becomes increasingly tuned to orthographic, as well as phonological, features as greater reading experience is acquired (Dehaene et al., 2010, 2015). The current study examined the role of a continuous measure of orthographic to phonological neighborhood overlap to obtain relevant behavioral and neuroimaging evidence under both easy and challenging listening conditions. Words with higher overlap were more difficult to identify, and evidenced increased BOLD responses, under noisy conditions, relative to words with less overlap. Contrary to prior theoretical accounts, increased responses were not observed in regions associated with either orthographic or phonological processing. Rather, cingulo-opercular activation varied with SNR and orthographic-to-phonological overlap. We suggest that such feature overlap may increase the availability of competing responses, which then requires additional support from domain general performance systems in order to produce a single response.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported (in part) by the National Institutes of Health (NIH)/National Institute on Deafness and Other Communication Disorders (P50 DC 00422), MUSC Center for Biomedical Imaging, South Carolina Clinical and Translational Research (SCTR) Institute, NIH/National Center for Research Resources (NCRR) Grant number UL1 RR029882, and a University of California Riverside Academic Senate Travel and Research Award. This investigation was conducted in a facility constructed with support from Research Facilities Improvement Program (C06 RR14516) from the NIH/NCRR.

References

Adelman JS, Brown GDA. Phonologic neighbors, not orthographic neighbors, determine word naming latencies. Psychonomic Bulletin & Review. 2007; 14:455–459. [PubMed: 17874587]

Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. Conflict monitoring and cognitive control. Psychological Review. 2001; 108:624–652. [PubMed: 11488380]

Binder JR, Medlar DA, Desai R, Conant LL, Liebenthal E. Some neurophysiological constraints on models of word naming. NeuroImage. 2005; 27:677–693. [PubMed: 15921937]

Avants BB, Gee JC. Geodesic estimation for large deformation anatomical shape averaging and interpolation. NeuroImage. 2004; 23(Suppl 1):S139–S150. [PubMed: 15501083]

Avants, BB., Tustison, NJ., Song, G. Advanced normalization tools. Vrsion 2.1. 2011. Available at http://www.picsl.upenn.edu/ANTS

Carhart R, Tillman TW, Greetis ES. Perceptual masking in multiple sound backgrounds. Journal of the Acoustical Society of America. 1969; 45:694–703. [PubMed: 5776931]

Cavanagh JF, Wiecki TV, Cohen MX, Figueroa CM, Samanta J, Sherman SJ, Frank MJ. Subthalamic nucleus stimulation reversed mediofrontal influence over decision threshold. Nature Neuroscience. 2011; 14:1462–1467. [PubMed: 21946325]

Chereau C, Gaskell G, Dumay N. Reading spoken words: Orthographic effects in auditory priming. Cognition. 2007; 102:341–360. [PubMed: 16480971]

Dehaene S, et al. How learning to read changes the cortical networks for vision and language. Science. 2010; 330:1359–1364. [PubMed: 21071632]

Dehaene S, Cohen L, Morais J, Kolinsky R. Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. Nature Reviews Neuroscience. 2015; 16:234–244. [PubMed: 25783611]

Dirks DD, Takayanagi S, Moshfegh A. Effects of lexical factors on word recognition among normal-hearing and hearing-impaired listeners. Journal of American Academy of Audiology. 2001; 12:233–244.

Dosenbach NUF, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Schlaggar BL. A core system for the implementation of task sets. Neuron. 2006; 50:799–812. [PubMed: 16731517]

Dosenbach NU, Fair DA, Cohen AL, Schlaggar BL, Petersen SE. A dual-networks architecture of top-down control. Trends in Cognitive Sciences. 2008; 12(3):99–105. [PubMed: 18262825]

Eckert MA, Teubner-Rhodes S, Vaden KI. Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. Ear and Hearing. 2016; 37:101S–110S. [PubMed: 27355759]

Eklund A, Dufort P, Villani M, Laconte S. BROCCOLI: Software for fast fMRI analysis on many-core CPUs and GPUs. Frontiers in Neuroinformatics. 2014; 8:24. [PubMed: 24672471]

Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proceedings of National Academy of Sciences. 2016; 113:7900–7905.

Fjell AM, Westlye LT, Amlien I, Espeseth T, Reinvang I, Raz N, Dale AM. High consistency of regional cortical thinning in aging across multiple samples. Cerebral Cortex. 2009; 19:2001–2012. [PubMed: 19150922]

Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC. Assessing the significance of focal activations using their spatial extent. Human Brain Mapping. 1994; 1:210–220. [PubMed: 24578041]

Grill-Spector K, Kourtzi Z, Kanwisher N. The lateral occipital complex and its role in object recognition. Vision Research. 2001; 41:1409–1422. [PubMed: 11322983]

Kalikow DN, Stevens KN, Elliott LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. Journal of the Acoustical Society of America. 1977; 61:1337–1351. [PubMed: 881487]

Kelley WM, Wagner DD, Heatherton TF. In search of a human self-regulation system. Annual Review of Neuroscience. 2015; 38:389.

Kuchinsky SE, Vaden KI, Keren NI, Harris KC, Ahlstrom JB, Dubno JR, Eckert MA. Age and word intelligibility affect cross-modal signaling in a word listening task. Cerebral Cortex. 2012; 22:1360–1371. [PubMed: 21862447]

Macey PM, Macey KE, Kumar R, Harper RM. A method for removal of global effects from fMRI time series. NeuroImage. 2004; 22:360–366. [PubMed: 15110027]

Miller G. The masking of speech. Psychological Bulletin. 1947; 44:105–129. [PubMed: 20288932]

Montant M, Schön D, Anton JL, Ziegler JC. Orthographic contamination of Broca's area. Frontiers in Psychology. 2011; 2:378.doi: 10.3389/fpsyg.2011.00378 [PubMed: 22207859]

Monzalvo K, Dehaene-Lambertz G. How reading changes children's spoken language. Brain and Language. 2013; 127:356–365. [PubMed: 24216407]

Mumford JA, Poline JB, Poldrack RA. Orthogonalization of regressors in fMRI models. PLoS One. 2015; 10(4):e0126255. [PubMed: 25919488]

Neta M, Schlaggar BL, Petersen SE. Separable responses to error, ambiguity, and reaction time in cingulo-opercular task control regions. NeuroImage. 2014; 99:59–68. [PubMed: 24887509]

Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia. 1971; 9:97–113. [PubMed: 5146491]

Pattamadilok C, Knierim IN, Kawabata Duncan KJ, Devlin JT. How does learning to read affect speech perception? Journal of Neuroscience. 2010; 30:8435–8444. [PubMed: 20573891]

Pattamadilok C, Morais J, Colin C, Kolinsky R. Unattentive speech processing is influenced by orthographic knowledge: Evidence from mismatch negativity. Brain and Language. 2014; 137:103–111. [PubMed: 25190330]

Pattamadilok C, De Morais JJ, Kolinsky R. Naming in noise: the contribution of orthographic knowledge to speech repetition. Frontiers in Psychology. 2011; 2:361.doi: 10.3389/fpsyg. 2011.00361 [PubMed: 22164152]

Pattamadilok C, Perre L, Dufau S, Ziegler JC. On-line orthographic influences on spoken language in a semantic task. Journal of Cognitive Neuroscience. 2009; 21:169–179. [PubMed: 18476763]

Pattamadilok C, Kolinski R, Ventura P, Radeau M, Morais J. Orthographic representations in spoken word priming: No early automatic activation. Language and Speech. 2007; 50:505–531. [PubMed: 18330215]

Perre L, Bertrand D, Ziegler JC. Literacy affects spoken language in a non-linguistic task: an ERP study. Frontiers in Psychology. 2011; 2:274.doi: 10.3389/fpsyg.2011.00274 [PubMed: 22025917]

Perre L, Ziegler JC. On-line activation of orthography in spoken word recognition. Brain Research. 2008; 1188:132–138. [PubMed: 18062940]

Perre L, Midgley K, Ziegler JC. When beef primes reef more than leaf: Orthographic information affects phonological priming in spoken word recognition. Psychophysiology. 2009a; 46:739–746. [PubMed: 19386047]

Perre L, Pattamadilok C, Montant M, Ziegler JC. Orthographic effects in spoken language: On-line activation or phonological restructuring? Brain Research. 2009b; 1275:73–80. [PubMed: 19376099]

Petrova A, Gaskell MG, Ferrand L. Orthographic consistency and word-frequency effects in auditory word recognition: new evidence from lexical decision and rime detection. Frontiers in Psychology. 2011; 2:263.doi: 10.3389/fpsyg.2011.00263 [PubMed: 22025916]

Protopapas A, Orfanidou E, Taylor JSH, Karavasilis E, Kapnoula EC, Panagiotaropoulou G, Kelekis D. Evaluating cognitive models of visual word recognition using fMRI: Effects of lexical and sublexical variables. NeuroImage. 2016; 128:328–341. [PubMed: 26806289]

Sadat J, Martin CD, Costa A, Alario FX. Reconciling phonological neighborhood effects in speech production through single trial analysis. Cognitive Psychology. 2014; 68:33–58. [PubMed: 24291531]

Seidenberg MS, Tanenhaus MK. Orthographic effects on rhyme monitoring. Journal of Experimental Psychology: Human Learning and Memory. 1979; 5:546–554.

Sestieri C, Corbetta M, Spadone S, Romani GL, Shulman GL. Domain-general signals in the cingulo-opercular network for visuospatial attention and episodic memory. Journal of Cognitive Neuroscience. 2014; 26:551–568. [PubMed: 24144246]

Taft M. Orthographic influences when processing spoken pseudowords: theoretical implications. Frontiers in Psychology. 2011; 2:140.doi: 10.3389/fpsyg.2011.00140 [PubMed: 21886628]

Taft M, Castles A, Davis C, Lazendic G, Nguyen-Hoan M. Automatic activation of orthography in spoken word recognition: Pseudohomograph priming. Journal of Memory and Language. 2008; 58:366–379.

Taler V, Aaron GP, Steinmetz LG, Pisoni DB. Lexical neighborhood density effects on spoken word recognition and production in healthy aging. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences. 2010; 65B:551–560.

Teubner-Rhodes S, Vaden KI, Cute SL, Yeatman JD, Dougherty RF, Eckert MA. Aging-resilient associations between the arcuate fasciculus and vocabulary knowledge: Microstructure or morphology? Journal of Neuroscience. 2016; 36:7210–7222. [PubMed: 27383595]

Vaden KI, Gebregziabher M, Kuchinsky SE, Eckert MA. Multiple imputation of missing fMRI data in whole brain analysis. NeuroImage. 2012; 60:1843–1855. DOI: 10.1016/j.neuroimage.2012.01.123 [PubMed: 22500925]

Vaden KI Jr, Muftuler LT, Hickok G. Phonological repetition-suppression in bilateral superior temporal sulci. NeuroImage. 2010; 49:1018–1023. [PubMed: 19651222]

Vaden KI Jr, Piquado T, Hickok G. Sublexical properties of spoken words modulate activity in Broca's area but not superior temporal cortex: implications for models of speech recognition. Journal of Cognitive Neuroscience. 2011; 23:2665–2674. [PubMed: 21261450]

Vaden KI Jr, Kuchinski SE, Cute SL, Ahlstrom JB, Dubno JR, Eckert MA. The cingulo-opercular network provides word-recognition benefit. Journal of Neuroscience. 2013; 33:18979–18986. [PubMed: 24285902]

Wechsler, D. Manual for the Wechsler abbreviated intelligence test (WASI). San Antonio, TX: The Psychological Corporation; 1999.

Wilke M. An alternative approach towards assessing and accounting for individual motion in fMRI timeseries. NeuroImage. 2012; 59:2062–2072. [PubMed: 22036679]

Wilson RH, Trivette CP, Williams DA, Watts KL. The effects of energetic and informational masking on the Words-in-Noise Test (WIN). Journal of the American Academy of Audiology. 2012; 23:522–533. [PubMed: 22992259]

Vitevitch MS, Luce PA. Probabilistic phonotactics and neighborhood activation in spoken word recognition. Journal of Memory and Language. 1999; 40:374–408.

Ziegler JC, Ferrand L. Orthography shapes the perception of speech: The consistency effect in auditory word recognition. Psychonomic Bulletin & Review. 1998; 5:683–689.

Ziegler J, Ferrand L, Montant M. Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. Memory & Cognition. 2004; 32:732–741. [PubMed: 15552350]

Ziegler JC, Muneaux M, Grainger J. Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. Journal of Memory and Language. 2003; 48:779–793.

Ziegler JC, Petrova A, Ferrand L. Feedback consistency effects in visual and auditory word recognition: Where do we stand after more than a decade? Journal of Experimental Psychology: Learning, Memory, and Cognition. 2008; 34:643–661.

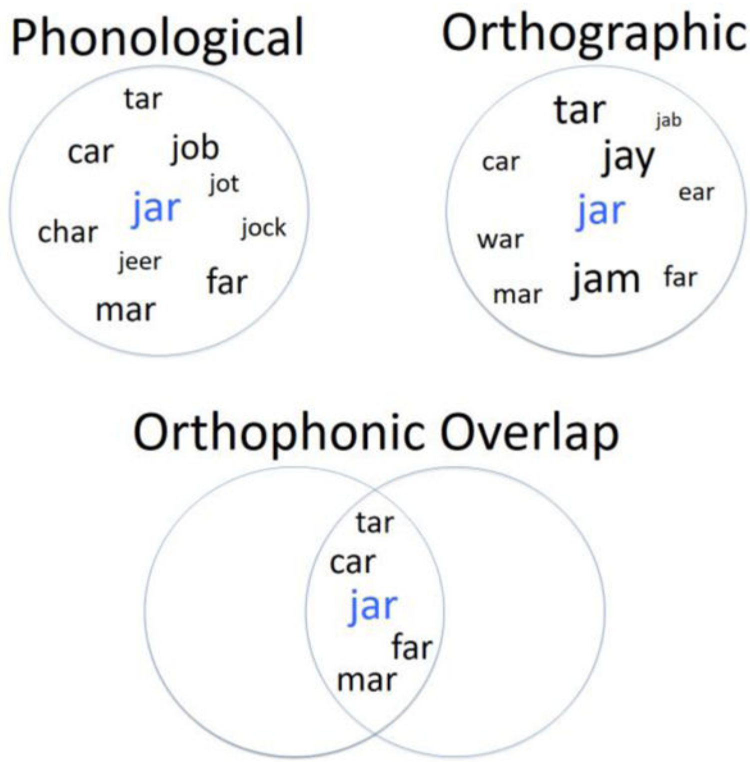

Figure 1.

Illustration of phonological and orthographic neighborhoods for 'jar' and the orthophonic overlap between them. Differences in word frequency approximated by font size. Not all neighbors are shown.

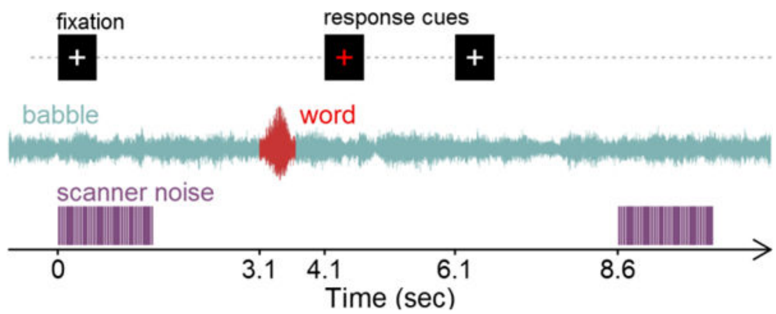

Figure 2.

The experiment was designed to accommodate a sparse acquisition fMRI acquisition (TR = 8.6 s) for a collected volume (0 to 1.65 s), with presentations fixed in their timing relative to the scans. A white crosshair fixation cue appeared at the beginning of each trial ($t = 0$ s), followed by the aural presentation of a single CVC word over the insert earphones ($t = 3.1$ s). The fixation cue changed color to red at the onset of the response interval ($t = 4.1$ s), then became white again when no time remained for a response ($t = 6.1$ s). The multi-talker babble was presented throughout each trial.

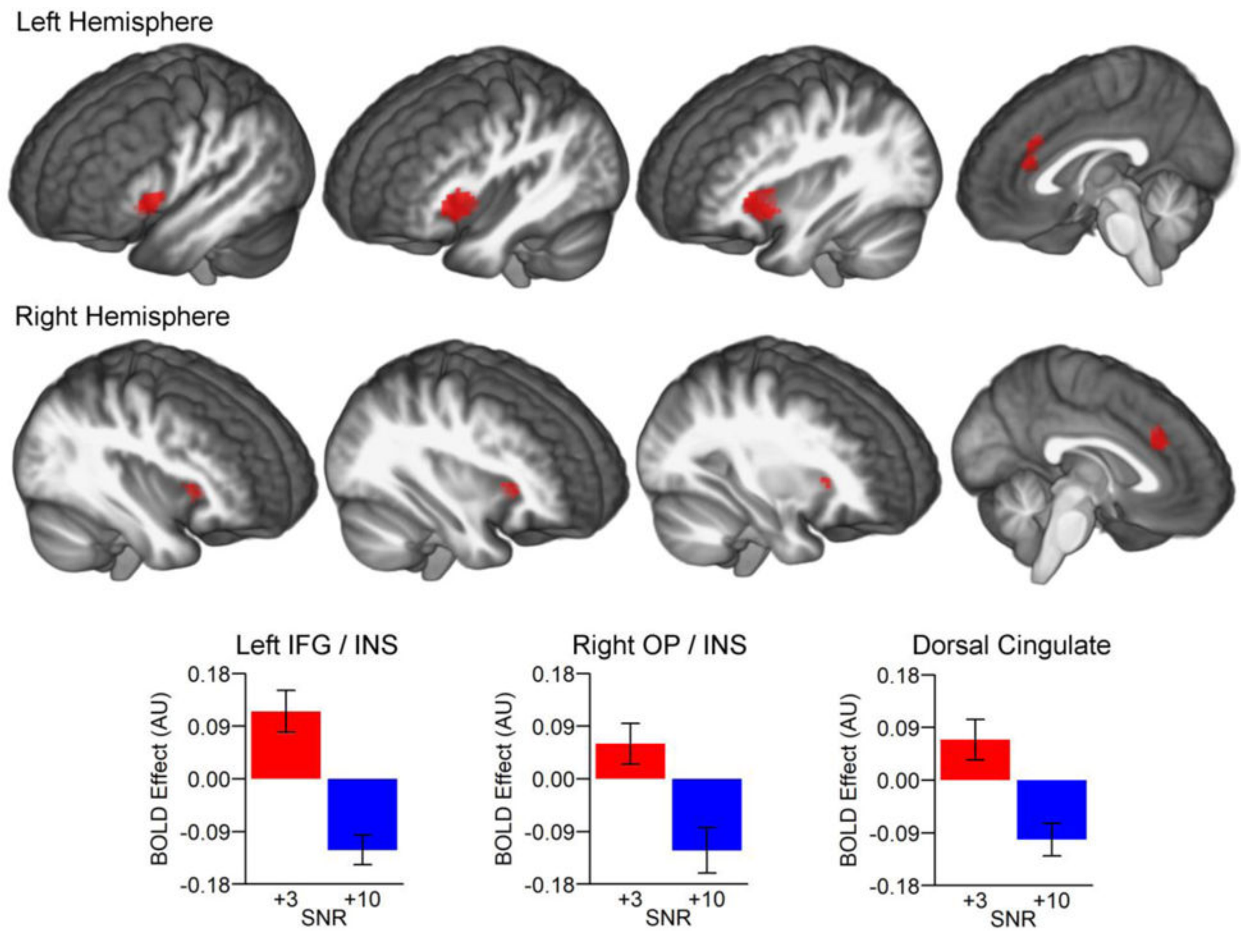

Figure 3.

Significant negative SNR × OPOR interactions were observed for BOLD contrast within the cingulate and left inferior frontal gyrus regions shown in red (top). The statistical map was submitted to a voxel-level threshold: Z = 3.09, $p_{UNC} = 0.001$ in combination with a permutation-corrected cluster-extent threshold: $p_{FWE} = 0.05$. Based on the same statistical threshold, there was an additional cluster in the right frontal operculum (OP) and insula [Peak-Z = 3.48; MNI: 33, 27, 5; 29 voxel extent $p_{FWE} = 0.003$(SPM)] that was not significant based on the permutation test. Each cingulo-opercular region exhibited increased BOLD contrast for the words with higher OPOR in the +3 dB SNR and decreased BOLD for words with higher OPOR in the +10 dB SNR condition. This interaction was consistent for the inferior frontal gyrus and insula (IFG/INS) bilaterally, in addition to the dorsal cingulate, as shown in the subplots (bottom).

Table 1

OPOR interaction with SNR, statistically independent from phonological neighborhood density and word frequency effects, predicts significant brain activity.

Description of Contrast, Cluster Extent	Peak Z	# Voxels	Peak MNI
Negative OPOR × SNR Interaction			
Dorsal Cingulate/Paracingulate	3.94	73	−3, 34, 39
L Inferior Frontal Gyrus, Anterior Insula	5.82	228	−47, 20, 6

Note: Significant results were based on a combined voxel statistic: $Z = 3.09$, $p_{UNC} = 0.001$, and permutation-corrected cluster extent: $p_{FWE} < 0.05$. MNI: Montreal Neurological Institute coordinates. Additional clusters were observed in the right frontal operculum and insula [Peak-Z = 3.48; MNI: 33, 27, 5; 29 voxels, SPM cluster extent $p_{\mathrm{FWE}} = 0.003$] and in the left occipital pole [Peak-Z = 3.80; MNI: −9, −96, 25; 29 voxels, SPM cluster extent $p_{\mathrm{FWE}} = 0.003$], which were not significant based on permutation testing; L: left, R: right, otherwise bilateral.

Table 2

Regional OPOR effects on BOLD within and between SNR conditions.

Contrast	Left vOTC −41, −56, −20	Left SMG −51, −32,26	Left Lateral IFG −50,4,8	Left IFG/AI −42,20,0
+3 dB vs. +10 dB SNR	−0.94	0.60	1.90	6.09 ***
+3 dB SNR	−1.29	−0.67	0.45	2.76
+10 dB SNR	0.15	−1.28	−2.00	−4.84 ***

Notes: The table presents results from one-sample *t*-tests ($df = 36$) on the average beta value from each region of interest for the association between OPOR and BOLD contrast (activity), within and between SNR conditions. Asterisks denote Bonferroni-corrected test significance;

*** $p < 0.001$; Abbreviations for ROI names include vOTC: ventral Occipital-Temporal Cortex; SMG: Supramarginal Gyrus; IFG: Inferior Frontal Gyrus; AI: Anterior Insula. MNI coordinates for each spherical ROI are shown below the name of the region.