

Published in final edited form as:

*Neuropsychologia*. 2011 November ; 49(13): 3563–3572. doi:10.1016/j.neuropsychologia.2011.09.008.

## Inferior frontal sensitivity to common speech sounds is amplified by increasing word intelligibility

**Kenneth I. Vaden Jr., Ph.D., Stefanie E. Kuchinsky, Ph.D., Noam I. Keren, M.A., Kelly C. Harris, Ph.D., Jayne B. Ahlstrom, M.S., Judy R. Dubno, Ph.D., and Mark A. Eckert, Ph.D.**  
Hearing Research Program, Department of Otolaryngology, Head and Neck Surgery, Medical University of South Carolina

### Abstract

The left inferior frontal gyrus (LIFG) exhibits increased responsiveness when people listen to words composed of speech sounds that frequently co-occur in the English language (Vaden, Piquado, Hickok, 2011), termed high phonotactic frequency (Vitevitch & Luce, 1998). The current experiment aimed to further characterize the relation of phonotactic frequency to LIFG activity by manipulating word intelligibility in participants of varying age. Thirty six native English speakers, 19–79 years old (mean = 50.5, sd = 21.0) indicated with a button press whether they recognized 120 binaurally presented consonant-vowel-consonant words during a sparse sampling fMRI experiment (TR = 8 sec). Word intelligibility was manipulated by low-pass filtering (cutoff frequencies of 400 Hz, 1000 Hz, 1600 Hz, and 3150 Hz). Group analyses revealed a significant positive correlation between phonotactic frequency and LIFG activity, which was unaffected by age and hearing thresholds. A region of interest analysis revealed that the relation between phonotactic frequency and LIFG activity was significantly strengthened for the most intelligible words (low-pass cutoff at 3150 Hz). These results suggest that the responsiveness of the left inferior frontal cortex to phonotactic frequency reflects the downstream impact of word recognition rather than support of word recognition, at least when there are no speech production demands.

### 1. Introduction

During speech recognition, words may activate sublexical representations that correspond to syllables or phonological chunks that are distinct from lexical representations. The influence of sublexical processing has been demonstrated in speech recognition tasks including infant word learning (Jusczyk et al., 1993; 1994; Mattys & Jusczyk, 2001), segmentation of continuous speech (Saffran, Newport, Aslin, 1996; Saffran et al., 1997; 1999; McQueen, 2001), wordlikeness judgments (Frisch, Large, Pisoni, 2000; Bailey & Hahn, 2001), and same-different decisions (Luce & Large, 2001; Vitevitch, 2003). During speech recognition, sublexical processing effects may also be elicited by manipulating the phonotactic frequency of spoken words, measured by pair-wise phoneme counts for words (Vitevitch & Luce, 1998), while controlling for other psycholinguistic factors (Luce & Large, 2001; Vaden, Piquado, Hickok, 2011). Typical low phonotactic frequency words include *badge*, *couch*,

---

Address for manuscript correspondence: Mark A. Eckert or Kenneth I. Vaden Jr., Medical University of South Carolina, Department of Otolaryngology-Head and Neck Surgery, Hearing Research Program, 135 Rutledge Avenue, MSC 550, Charleston, SC 29425-5500, Telephone: 00-1-843-792-2774, eckert@musc.edu, vaden@musc.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

*look*, compared to high phonotactic frequency words: *bag*, *catch*, *load*. Sublexical tasks that are perceptual in nature (e.g., syllable categorization, phoneme detection) often elicit left inferior frontal gyrus (LIFG) activity (Heim et al., 2003; Siok et al., 2003; Callan, Kent, Guenther, 2000; Burton et al., 2005; Zaehle et al., 2007), which was unexpected since the predominance of LIFG findings has been more closely related to speech production (Hickok & Poeppel, 2004) and comprehension of connected speech (Peelle et al., 2010; Rodd, et al., 2005, 2010; Tyler et al., 2010).

In a recent fMRI experiment, Vaden et al. (2011) identified the influence of phonotactic frequency manipulations on LIFG activity when participants listened to real words during a non- word detection task. Participants were instructed to listen to words and button press only when a non-word was presented, which occurred for 8% of the trials. Activity was elevated with increasing phonotactic frequency, which was interpreted to reflect the engagement of motor representation for the words during speech recognition. These results could be consistent with the Association hypothesis, which proposes that sublexical processing effects in the LIFG primarily reflect passively activated motor representations that are associated with heard speech. One alternative explanation for the Vaden et al. results is that resources in inferior frontal cortex were recruited to facilitate lexical representation. The aim of the current experiment was to replicate the sublexical processing effects reported by Vaden et al. (2011) and determine the extent to which association and recruitment-based explanations could account for the influences of phonotactic frequency on LIFG activity.

The Association hypothesis proposes that motor codes are associated with heard speech, primarily due to auditory-phonological feedback that guides articulation, and that those connections are strengthened during speech production (Hickok & Poeppel, 2004; 2007). According to this view, words that are composed of highly frequent phoneme-pairs result in greater LIFG activity because of associations with motor representations, which were strengthened previously during speech production. Auditory feedback manipulations affect both speech production (Jones & Munhall, 2000; 2002) and auditory cortex responses that guide articulation (Houde et al., 2002; Tourville, Reilly, Guenther, 2008), and such a feedback system may reasonably account for auditory-motor associations (Hickok & Poeppel, 2004). Auditory word recognition may trigger related sublexical representations in the LIFG in proportion to their exposure during speech – which could explain why activity varied with the phonotactic frequency of words (Vaden et al., 2011). The Association hypothesis leads to predictions about how age and word intelligibility would interact with sublexical activity. We reasoned that more normal sounding, high intelligibility words would have greater associative value with motor codes than low intelligibility words, so increased speech intelligibility was predicted to enhance phonotactic frequency effects in LIFG. Since younger and older adults both produce speech, and auditory-motor associations are maintained through articulation and sensory feedback (cf. Kent et al., 1987; Nordeen & Nordeen, 1992; Doupe & Kuhl, 1999), sublexical activity was not expected to systematically change with age.

The Recruitment hypothesis proposes that speech production related resources in inferior frontal cortex facilitate lexical access or selection, especially when difficult listening conditions degrade speech intelligibility (Davis & Johnsrude, 2003; Wilson, 2009; Osnes, Hugdahl, Specht, 2011). The LIFG and pre-motor cortex are often activated during speech perception tasks, particularly when segmentation or sublexical distinctions are required (for review, Hickok & Poeppel, 2004). The engagement of LIFG may result from increasing the difficulty of lexical-phonological access (Prabhakaran et al., 2006; Righi et al., 2009), response control (Thompson-Schill et al., 1997; Snyder, Feigenson, Thompson-Schill, 2007), or declining speech intelligibility (Davis & Johnsrude, 2003; Giraud et al., 2004; Obleser et al., 2007; Eckert et al., 2008). Consistent with these neuroimaging studies,

Broca's aphasics appear to experience more difficulty in picture-matching tasks than age-matched controls, when presented with acoustically degraded speech (Moineau, Dronkers, Bates, 2005). These results are consistent with a large body of literature and the premise that LIFG is engaged in challenging listening conditions where demands are increased on response selection.

The Recruitment hypothesis also predicts sublexical effects during speech recognition as identified by Vaden et al. (2011), but partially as a result of continuous scanner noise and increased demands for high phonotactic frequency words during a pseudoword detection task. Sublexical processing might not otherwise be engaged when clear acoustic conditions favor direct lexical access; instead it may only facilitate speech recognition under challenging acoustic conditions (Mattys, White, Melhorn, 2005). The Recruitment hypothesis therefore predicts that decreasing speech intelligibility would result in increased sublexical processing activity in LIFG, which we can measure by manipulating phonotactic frequency. This account also predicts that sublexical processing and inferior frontal activity may increase in older adults, since the impact of age-related hearing loss and speech recognition declines (Dubno, Dirks, Morgan, 1984; Humes, Christopherson, 1991; Sommers, 1996; Dubno, et al., 2008) and central auditory system degradation (Harris et al., 2009) might be offset with increased involvement of prefrontal processing (Eckert, et al., 2008; 2010; Wong et al., 2009). Recruitment of sublexical representations might offer a potential mechanism to compensate for changes throughout the aging auditory system, if sublexical representations modulate word recognition.

The current study examined association and recruitment explanations for phonotactic frequency effects on LIFG activity. We predicted that LIFG activity would positively correlate with phonotactic frequency, and that the magnitude of this association would be affected by word intelligibility and participant age. If sublexical processing passively activates motor codes by association, then stronger phonotactic frequency effects would be expected when word intelligibility increases and these effects would not vary with age. However, if motor codes were recruited to modulate auditory activity, then the strongest phonotactic frequency effect would be expected in the least intelligible condition and this effect would increase in older compared to younger adults.

## 2. Method

### 2.1 Participants

Thirty-six<sup>1</sup> healthy adults participated in the study, 23 females and 13 males, 19–79 years old, mean age = 50.5 (sd = 21.0). All participants were native English speakers and right handed according to the Edinburgh handedness questionnaire, mean score = 91.4 (sd = 9.6), out of a possible –100 to 100 range, where 100 is strongly right handed (Oldfield, 1971). On average, the participants had 18.1 years of education (sd = 2.9); mean socioeconomic status = 52.4 (sd = 10) in a possible range of 8 to 66 (Hollingshead, 1975); and made fewer than three errors on the Mini Mental Status Examination, mean score was 29.6 out of 30 (sd = 0.6; Folstein, Robins, Helzer, 1983) demonstrating performance consistent with little or no cognitive impairment (cf. review by Tombaugh and McIntyre, 1992). Participants had no history of neurological or psychiatric events. 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.

---

<sup>1</sup>Originally 42 adults participated in the study, but technical issues prevented responses from being recorded for six of them. Since we were unsure of their task engagement, those subjects were excluded from analysis.

Pure-tone thresholds were measured with a Madsen OB922 audiometer and TDH-39 headphones, calibrated to ANSI standards (American National Standards Institute, 2004). Most of the participants ( $N = 20$ , mean age = 35.8,  $sd = 15.5$ ) had clinically normal hearing, with pure-tone thresholds of 25 dB HL or less from 250 Hz to 8000 Hz, while the others ( $N = 16$ , mean age = 68.8,  $sd = 9.2$ ) had mild to moderate hearing loss that affected hearing sensitivity mostly in the 4000 Hz to 8000 Hz frequency range. The pure-tone threshold measure used in fMRI and behavioral analyses was an average of the left and right ear thresholds for 250 Hz to 8000 Hz. Differences in thresholds between right and left ears did not exceed 15 dB at any frequency, and all participants had normal immittance measures.

## 2.2 Stimuli

The current experiment used 120 words (Supplemental Table 1) recorded by a male speaker that were selected from 400 monosyllabic consonant-vowel-consonant words by Dirks et al. (2001). Each item was selected from controlled ranges of word frequency: mean = 57.9,  $sd = 126.8$  (Kučera & Francis, 1967), neighborhood density: mean = 26.9,  $sd = 9.9$ , and phonotactic frequency: mean = 0.0018,  $sd = 0.0019$ , based on the Irvine Phonotactic Online Dictionary [IPHOD (Vaden, Halpin, Hickok, 2009)]. Specifically, phonotactic frequency was defined using the unweighted average biphoneme probability measure and neighborhood density was defined using the unweighted phonological neighborhood density measure. Each of the words was high pass filtered (200 Hz) and low pass filtered at one of four cutoff frequencies to parametrically vary word intelligibility (from least to most intelligible: 400 Hz, 1000 Hz, 1600 Hz, and 3150 Hz). Words did not repeat across intelligibility conditions, to avoid interactions between word intelligibility and priming or memory effects. This intelligibility manipulation was consistent with previous studies in which band-pass filter effects affected word recognition performance (Eckert et al., 2008; Harris et al., 2009).

Analyses of Variance (ANOVA) were performed to ensure that word intelligibility conditions were not confounded with other factors of interest. No significant differences between items in each word intelligibility condition were observed for Kučera-Francis word frequency [ $F(3,116) = 0.09$ ,  $p = 0.97$ ], IPHOD neighborhood density [ $F(3,116) = 0.15$ ,  $p = 0.93$ ], IPHOD phonotactic frequency [ $F(3,116) = 1.17$ ,  $p = 0.33$ ], or word duration [ $F(3,116) = 0.17$ ,  $p = 0.91$ ].

Two strategies were undertaken to limit the influence of individual hearing differences on the results of the experiment. All words were band-pass filtered above 200 Hz and below 3150 Hz to limit the impact of high frequency hearing loss while preserving the majority of the speech information. In addition, spectrally-shaped noise was digitally generated and continuously delivered during the experiment to provide generally equivalent masked thresholds for all participants. The noise was presented at 62.5 dB SPL and words were presented at 75 dB SPL. Because age-related differences in hearing might still contribute to word recognition differences, pure-tone thresholds from each participant were used to assess associations between age-related hearing loss, word recognition, and functional effects.

## 2.3 Experimental Procedure

Participants were given the task instructions outside of the scanner and reminded of the directions inside the scanner just prior to the experiment. They were instructed to listen carefully and indicate by button-press whether or not they recognized each word. Participants were informed that all the items were real words and that some words were harder to recognize than others. There were 8 sec between each volume acquisition onset (TR = 8 sec). During each listening trial, one word was presented two seconds after the previous volume acquisition. A white fixation cross appeared onscreen throughout word

recognition trials, turning red for two seconds to cue a button response after the word played. The timing of events is illustrated in Figure 1.

An event-related design and sparse sampling fMRI acquisition were used to aurally present one word per trial, in the absence of scanner noise. Each participant was presented with 120 words that were interspersed with 11 resting trials in the same conditions and order. Presentation order was pseudo-randomized with respect to the intelligibility, phonotactic frequency and neighborhood density of each word: correlation results demonstrated no significant associations between trial number and those characteristics (across all trials and within each intelligibility condition all R-squared values were ns). In total, 131 functional volumes were collected over 17 min 42 sec. Word recordings and background noise were combined using an audio mixer and presented to participants through piezoelectric insert earphones (Sensimetrics Corp.). Sound levels were calibrated prior to each scanning session using a precision sound level meter (Larson Davis 800B). Participants viewed the projector screen through a periscope mirror mounted above the head coil. Trial presentation was synchronized with the scanner using EPrime software (Psychology Software Tools Inc.) and an IFIS-SA control system (Invivo Corp.), and responses were collected with an ergonomically designed MRI-compatible button box (Magnetic Resonance Technologies Inc.).

## 2.4 Image acquisition

Structural and functional images were collected on a Philips 3T Intera scanner with an 8 channel SENSE headcoil (reduction factor = 2). Whole brain functional images were acquired using a T2\* weighted, single shot Echo Planar Imaging sequence (40 slices with a 64×64 matrix, TR = 8 sec, TE = 30 msec, TA = 1647 msec, slice thickness = 3.25 mm, gap = 0, sequential order). Voxel dimensions were 3×3×3.25 mm. The anatomical images consisted of 160 slices with a 256×256 matrix, TR = 8.13ms, TE = 3.7 msec, flip angle = 8 deg, slice thickness = 1.0 mm, and no slice gap. High resolution structural T1-weighted anatomical images were collected for the normalization procedure, which took advantage of the higher resolution of T1 images, as well as grey and white matter segmented images.

## 2.5 Neuroimaging data preprocessing and analysis

**Preprocessing**—Because participants were distributed over a large age range, a study-specific template was created to ensure optimal spatial alignment across younger to older adult images. Unified segmentation and diffeomorphic registration was performed in SPM5 [DARTEL (Ashburner & Friston, 2005; Ashburner, 2007)] to iteratively bias field correct and segment the T1-weighted images into their native space tissue components, and generate normalization parameters. A recursive diffeomorphic registration procedure preserved cortical topology according to a Laplacian model, which created invertible and smooth deformations to transform each individual's native space gray matter image to a common template space. The resultant structural template was therefore representative in size and shape of all the participants, and is shown in the results figures. Each individual's functional data were co-registered to their native space anatomical image using a mutual information optimization algorithm, then normalization parameters or flow fields derived for each individual's transformation to template space were applied to the corresponding functional data. The co-registration and normalization processes resulted in individual-level functional images that were spatially transformed into the study-specific template space.

Preprocessing was performed using SPM5 (Wellcome Department of Imaging Neuroscience): functional data were unwarped and realigned, normalized using parameters from the DARTEL procedure described above, and a Gaussian smoothing kernel of 8 mm was applied. Global mean signal fluctuations were detrended from the preprocessed



functional images using voxel-level Linear Model of the Global Signal (Macey et al., 2004). Next, we applied an algorithm described in Vaden, Muftuler, Hickok (2010) to identify extreme intensity fluctuations that occurred during each run. Cutoff values were all set to 2.5 standard deviations from the mean, identifying extreme noise in 4.65% of the functional images. These two outlier vectors were submitted to the first-level general linear model (GLM) as nuisance variables. We also included two motion-based nuisance variables in the first-level GLM that represented the magnitude of angular and translational displacement based on the Pythagorean Theorem (similar to Salek-Haddadi et al., 2006), which were calculated from the six realignment parameters generated by SPM realignment functions.

**Full model parametric analysis**—The preprocessed functional images were submitted to a parametric analysis in SPM5 at the individual level. We chose to model words in the four intelligibility conditions separately, since there were discrete differences in word intelligibility. Words in each intelligibility condition were modeled as separate events with onset and duration specified, and two parametric regressors specifying the phonotactic frequency and neighborhood density. Thus, the functional timecourses from each word intelligibility condition were analyzed using onsets and durations, convolved with the hemodynamic response function (HRF), in order to perform correlation analyses with density and phonotactic frequency across the words. This regression model also included 4 volume-wise nuisance regressors: two motion and two outlier vectors, described in the preprocessing section. Finally, the first level phonotactic frequency and density results from each participant were included in a second level random effects analysis to localize consistent phonotactic and density effects across participants.

Group level t-statistic maps were thresholded using a t threshold that corresponded to  $p = 0.001$  (uncorrected) and cluster-level  $p = 0.05$  (Family Wise Error corrected). Age was also included as a group level covariate to determine the extent that significant effects were influenced by age. Cluster level correction used voxel extent thresholds based on Random Field Theory ( $\sim$  image smoothness FWHM  $\approx 9.4 \text{ mm}^3$  and each resel  $\approx 30$  voxels). The MNI coordinates for each peak t-statistic within clusters were obtained by transforming each group level statistic map with normalization parameters generated in SPM5 by warping the co-registered study-specific gray matter template to match an MNI gray matter template and applying the normalization parameters to the statistic images.

**Region of Interest (ROI) analyses**—A LIFG ROI mask was functionally defined based on the significant phonotactic frequency effect that was detected with the parametric analysis, as described above. The first analysis assessed the correlation between the age of participants and their ROI-averaged phonotactic frequency effect. The second ROI analysis treated the phonotactic frequency effect from each intelligibility condition as a dependent variable, and tested for differences related to word intelligibility condition.

**Control analyses**—The full model parametric analysis was modified to demonstrate that phonotactic frequency effects were not directly attributable to other factors. Two submodel analyses were performed which included either phonotactic frequency or neighborhood density in the GLM. Among monosyllabic English words, PF and ND are positively correlated because short words that share a phoneme pair often differ by one phoneme (Vitevitch, et al., 1999; Storkel, 2004). Since these variables exhibited a modest association among the words used in the experiment ( $R^2 = 0.041$ ,  $p = 0.027$ ), the submodel approach could potentially highlight unique correlations otherwise obscured by collinearity (Wilson, Isenberg, Hickok, 2009). A different concern was that since phonotactic frequency sometimes affects task performance, it could drive frontal activation more directly related to cognitive load. Two response-based control analyses were implemented to separate activity reflective of the difficulty that the participant experienced from stimulus driven effects. The

reaction time control analysis included an additional parameter in the GLM for each word, which specified the response latency from a participant if it occurred within the two second interval. Two participants were excluded due to missing reaction time data ( $N = 34$ ). The recognition control analysis examined whether phonotactic frequency effects were still detected when the GLM only modeled words that the participant recognized. Words in the lowest intelligibility condition (400 Hz) were not modeled due to insufficient power (mean < 33% recognition) and four participants were excluded since they indicated recognition in fewer than 33% of all trials ( $N = 32$ ).

### 3. Results

#### 3.1 Intelligibility, phonotactic frequency effects in reported recognition

Word recognition (percent correct) was measured as the proportion of button responses that indicated that the participant recognized the presented word. Word recognition was significantly related to word intelligibility (mean percent recognition responses from lowest to highest filter condition: 32%, 67%, 81%, 97%). A repeated-measures ANOVA demonstrated significant differences in word recognition based on intelligibility,  $F(3,105) = 136.29$ ,  $p < 0.001$ , and pairwise Tukey tests demonstrated significant differences between all conditions,  $p < 0.001$ . There were no significant correlations between word recognition and age ( $R^2 = 0.07$ ,  $p = 0.11$ ), nor recognition and mean pure tone thresholds ( $R^2 = 0.04$ ,  $p = 0.25$ ). Within each filter condition, word recognition was not significantly correlated with age (all  $R^2 < 0.09$ ,  $p > 0.11$ ), nor was recognition significantly correlated with mean pure tone threshold (all  $R^2 < 0.06$ ,  $p > 0.15$ ). These results confirmed an effect of filter condition on word recognition, while age and hearing did not appear to impact a participant's reported word recognition.

The logistic regression analysis (Baayen, Davidson, Bates, 2008) detected a significant main effect of word intelligibility condition ( $Z = 27.5$ ,  $p < 0.001$ ), as well as significant effects of phonotactic frequency in the 3150 Hz filter condition ( $Z = -2.87$ ,  $p = 0.004$ ) and neighborhood density in the 400 Hz filter condition ( $Z = -4.17$ ,  $p < 0.001$ ) and 1000 Hz filter conditions ( $Z = 2.74$ ,  $p = 0.006$ ). Model testing demonstrated that intelligibility conditions, phonotactic frequency in 3150 Hz filter trials, and neighborhood density in 400 Hz and 1000 Hz trials were each factors that contributed significantly to the model's fit, all  $\chi^2 > 7.42$  and  $p$ 's < 0.006. A significant change in the goodness of fit was not observed with the inclusion in the model of phonotactic frequency or neighborhood density within the other intelligibility conditions, word frequency, participant ages, and participant mean pure tone thresholds (all  $\chi^2 < 3.68$  and  $p$ 's > 0.06). In summary, increasing phonotactic frequency resulted in lower reported word recognition in the highest intelligibility condition (3150 Hz), but was not significantly related to performance in the other three conditions.

#### 3.2 Pars triangularis: phonotactic frequency replication

Words that consisted of more frequent phoneme pairs elicited increased LIFG responses. The LIFG contained one cluster with activity that correlated positively with phonotactic frequency [first peak:  $t(34) = 4.82$ ,  $p < 0.001$ , MNI coordinates = -46, 41, 8; second peak MNI coordinates = -37, 33, 10] and survived multiple comparison correction at the cluster level (corrected  $p < 0.001$ ; 76 voxels). Shown in Figure 3A, this cluster extended from the medial pars triangularis (PTr) into the lateral inferior frontal sulcus and medial anterior insula. Whole-brain analyses did not demonstrate any significant associations between cortical responses and the age covariate, neighborhood density, or negatively with phonotactic frequency, after correcting for multiple comparisons (cluster-corrected  $p = 0.05$  threshold). Significant task-correlated effects (i.e., listening - rest) did not include clusters

that overlapped with phonotactic frequency effects (Supplementary Figure), which suggests that phonotactic frequency modulated sub-threshold activity, as in Vaden et al. (2011).

### 3.3 No relationship between age or hearing, and phonotactic frequency effects

Correlations were performed using a functionally defined ROI based on the main phonotactic frequency effect in LIFG that was identified using the full model. Phonotactic frequency effect contrast values from within that ROI were averaged within each individual then correlated with participant ages, which demonstrated a non-significant association ( $R^2 = 0.049$ ,  $p = 0.19$ ; Figure 4). The LIFG-phonotactic frequency parameter estimate was also unrelated to mean pure tone threshold ( $R^2 = -0.048$ ,  $p = 0.20$ ). In summary, neither age nor hearing differences in our sample could explain differences in sublexical activity among individuals.

### 3.4 Greatest phonotactic frequency effects for the most intelligible speech

The LIFG ROI was also used to determine the extent to which sublexical activity was dependent on word intelligibility. Phonotactic frequency effects that were modeled separately for each word intelligibility condition, were averaged within the ROI for each individual, then tested for differences between intelligibility conditions. A repeated-measures ANOVA demonstrated a main effect of word intelligibility on sublexical activity [ $F(3,105) = 6.30$ ,  $p < 0.001$ ]. Tukey tests revealed  $p$ 's  $< 0.05$  for all pairwise comparisons with the most intelligible condition (3150 Hz), but no other significant pairwise effects were observed ( $p$ 's  $> 0.48$ ) for the other word intelligibility conditions. Phonotactic frequency did significantly engage LIFG in the other intelligibility conditions, but not to the same extent as the 3150 Hz condition (Figure 4). Phonotactic frequency effects modeled in each intelligibility condition were significantly greater than zero,  $t(35) > 2.15$ ,  $p < 0.04$ , except for the 1600 Hz words,  $t(35) = 1.35$ ,  $p = 0.18$ .

To ensure that intelligibility related changes in activity did not underlie the change in correlation with phonotactic frequency, the LIFG ROI was used to calculate mean contrast values for event-related effects (i.e., listening minus rest), modeled separately by word intelligibility. An ANOVA demonstrated no significant effect of word intelligibility condition on overall event-related activity in LIFG [ $F(3,105) = 0.37$ ,  $p = 0.77$ ], nor were there significant pairwise differences [all Tukey-corrected  $F(1,105) < 0.9$  and  $p > 0.77$ ]. These results indicated that while the association between LIFG activity and phonotactic frequency significantly increased for the most intelligible speech, that interaction could not simply be explained by increasing or decreasing regional activity due to the word intelligibility manipulations. We performed two additional correlation tests to establish that age and hearing differences were not significantly related to activity in any word intelligibility condition, both found  $R^2 < 0.03$  and  $p > 0.29$ .

### 3.5 Control Analyses

The submodel control analyses included either phonotactic frequency or neighborhood density as a parametric regressor, but not both. The phonotactic frequency-only submodel demonstrated only one significant cluster that was in the same location as the positive phonotactic frequency effect in the main analysis [peak  $t = 5.32$ , MNI coordinates =  $-44$ ,  $39$ ,  $7$ ; second peak =  $-38$ ,  $35$ ,  $9$ ] and included 125 voxels (corrected cluster  $p < 0.001$ ). Significant clusters were not observed with the neighborhood density-only submodel. The submodel analyses demonstrated a specific effect of phonotactic frequency, and not neighborhood density, on LIFG activity. A paired t-test demonstrated a significantly greater effect of phonotactic frequency than neighborhood density for the separately modeled contrast values [ $t(35) = 5.28$ ,  $p < 0.001$ ].



The performance-based control analyses included additional reaction time parameters, or restricted the analysis according to reported word recognition. There were reaction time effects throughout the left inferior frontal cortex, but these effects were independent of the phonotactic frequency effect. Again, a LIFG cluster was positively correlated with phonotactic frequency, [peak  $t = 4.83$ , MNI coordinates =  $-39, 29, 9$ ; second peak =  $-45, 40, 8$ ], corrected cluster  $p = 0.006$  (45 voxels). Only one overlapping voxel was identified in both the reaction time and phonotactic frequency results. Figure 3C illustrates the spatially distinct distribution of phonotactic frequency and reaction time effects. In the second analysis, phonotactic frequency effects were observed in a LIFG cluster [peak  $t = 4.35$ , MNI coordinates =  $-46, 42, 10$ ; second peak =  $-31, 35, -3$ ] at a lower statistical threshold of  $p < 0.005$  ( $t = 2.74$ ,  $df = 31$ ) and a corrected cluster extent threshold of  $p < 0.05$  (70 voxels). Both of the response-based control analyses demonstrated that cognitive load did not account for the spatially distinct phonotactic frequency effects.

#### 4. Discussion

The results of this study demonstrated elevated LIFG activity in response to words with increasing phonotactic frequency. Importantly, the association between phonotactic frequency and LIFG activity was greatest when speech was highly intelligible. Together with the absence of an aging effect and unique effects of reaction time on inferior frontal activity, these results were more consistent with the predictions of an association-based mechanism for which sublexical processing involves the passive activation of motor-related sublexical representations that are associated with heard speech.

The current experiment replicated Vaden et al. (2011), despite differences in the design of the two studies. In contrast to the younger adults studied in the Vaden et al. study, the participants in this study included a broad range of younger and older adults, which provided a means to determine that sublexical processing is present throughout adulthood. We presented a single word each trial (TR) to prevent segmentation demands from potentially driving IFG activity (e.g., McQueen, 2001; McNealy, Mazziotta, Dapretto, 2006), in contrast to word lists used in Vaden et al. (2011). The current experiment also took additional measures to eliminate scanner noise interference with stimulus presentation and reduce the impact of individual differences in hearing threshold, which allowed us to explicitly manipulate word intelligibility. Scanner noise and individual hearing differences otherwise might increase the need for sublexical processing or account for individual differences, based on a recruitment mechanism. Additional differences between the studies included task demands, which could have biased responses toward greater sensitivity to lexical or sublexical manipulations (Vitevitch, 2003). The word recognition task used in this study reduced potential sublexical processing bias from exaggerating LIFG activity, compared to previous experiments which employed some non-word stimuli (Papoutsis et al., 2009; Vaden et al., 2011). Since these methodological changes did not change the main result, they strengthened the conclusion of both studies that variation in phonotactic frequency modulates LIFG activity during spoken word recognition. Importantly, the common button press design for expressing decisions in each study, rather than an oral response, was likely a key design feature common to both studies that enhanced the likelihood that the phonotactic frequency finding was replicated. We return to this issue below.

The current study was designed to test two potential interpretations of the Vaden et al. (2011) findings, which included 1) recruitment of LIFG during high phonotactic frequency conditions and 2) passive activation of motor plans that are frequently engaged when producing speech by high phonotactic frequency sounds. The results of this study suggest

that, at least with respect to word listening studies in which no overt retrieval is required, the results may be better explained by an association-based explanation.

### Limited evidence for phonotactic frequency recruitment

The Recruitment hypothesis predicted that older adults would exhibit the most robust LIFG relation with phonotactic frequency, for the following reasons. First, older adults experience increased task demands across perceptual and cognitive tasks compared to younger adults (Dubno, Dirks, Morgan, 1984; Humes, Christopherson, 1991; Dubno, et al., 2008), which at least partially explains why frontal cortex tends to exhibit elevated activity in older adults (Eckert, et al., 2008; Wong et al., 2009). Second, an increased phonotactic frequency relation with LIFG activity in older adults would reflect recruitment of LIFG to perform the word listening task. However, we did not observe age-related phonotactic frequency effects on LIFG activity.

Phonotactic frequency influences were most pronounced in the most intelligible listening condition (3150 Hz), as opposed to the least intelligible listening condition (400 Hz) which was predicted based on a recruitment account. Furthermore, phonological neighborhood density of the words was related to reported recognition in the two most difficult word intelligibility conditions. Neighborhood density effects on word recognition performance were more consistent with a recruitment explanation, since words from dense phonological neighborhoods were more difficult to distinguish when intelligibility was degraded. This result provides an additional dissociation between phonotactic frequency and neighborhood density and further suggests that the phonotactic frequency modulation of LIFG response was not due to LIFG recruitment to perform the task during the least intelligible presentations.

As discussed in the introduction, previous studies have demonstrated that lexical demands could increase LIFG activity. This implied that phonotactic frequency might indirectly modulate activity in relation to uncertainty in word recognition and increased processing times. However, when we excluded unrecognized words from the whole brain analysis, we observed that reducing the range of difficulty did not eliminate the phonotactic frequency effects. Furthermore, the phonotactic frequency effect was robust and unique when the reaction time parameter was added to the model, suggesting that variation in reaction time did not account for the phonotactic frequency effect. We are cautious in the interpretation of the reaction time finding because there was a short delay before participants responded, which restricted the range of the reaction time data. However, these results are incompatible with the view that difficulty mediated the relation between phonotactic frequency and LIFG activity.

### Inferior frontal activity is modulated by task demands

Again, long reaction times were related to increased LIFG activation across subjects, but this reaction time effect did not account for the phonotactic frequency association with LIFG activation. Nonetheless, it is possible that task demands influenced the phonotactic frequency and LIFG findings. The current study and Vaden et al. (2011) involved tasks that did not require overt speech production. In contrast, Papoutsis et al., (2009) observed increased activity in the pars opercularis for low phonotactic frequency compared to high phonotactic frequency nonwords during an overt speech repetition task. We have observed similar results to the Papoutsis findings when participants perform an overt speech repetition task involving words (Vaden, preliminary observation). One explanation for these results is that the same LIFG response selection system is differentially modulated depending on task demands. In particular, neural resources that are important for motor planning may be up-regulated for low phonotactic frequency sounds when planning for articulation (Aichert &

Ziegler, 2004), but show a general elevation in response to high phonotactic frequency sounds during passive listening (Vaden et al., 2011). One explanation accounting for the multidimensional response properties of the LIFG is that the LIFG may be recruited when more effort is required to articulate low phonotactic frequency sounds, and passively activated by association to high phonotactic frequency sounds in the absence of overt articulatory demands.

The engagement of a frontal attention-related system could also explain the phonotactic frequency results. The phonotactic frequency association with LIFG activity could reflect a global shift in frontal attention-related systems in response to words of increasing phonotactic frequency. For example, a salience or ventral attention network is hypothesized to monitor important environmental stimuli (Dosenbach et al., 2006; 2007; Menon & Uddin, 2010). High phonotactic frequency stimuli may elicit increased engagement of a salience or ventral attention network. In support of this explanation, the medial LIFG and anterior insula region associated with phonotactic frequency is also considered part of a bilaterally organized salience network (Seeley et al., 2007). This salience pattern of activity is typically associated with unexpected events (Corbetta & Shulman, 2002), performance errors (Menon et al., 2001; Klein et al., 2007) and increased task difficulty (Eckert et al., 2009), and to long reaction times (Binder et al., 2004).

A salience network explanation is appealing since phonotactic frequency modulated activity in the left anterior insula and appeared to have affected recognition difficulty. However, the findings of this study were not consistent with that view. Because LIFG response was positively correlated with phonotactic frequency, this explanation would require the counter-intuitive possibility that high phonotactic frequency words (e.g., den) were more salient than low phonotactic frequency words (e.g., wave), which were composed of comparatively more unusual phoneme pairs. This explanation would predict a strong coupling of phonotactic frequency effects to recognition responses, but instead we observed that the phonotactic frequency effect was not eliminated when the model was restricted to recognition response trials. Response latencies correlated with clusters in the pars orbitalis and lateral anterior insula in a spatial distribution that was more consistent with saliency or difficulty. Finally, task difficulty-driven responses in anterior insula are often bilateral or right-lateralized, which is not spatially consistent with the left lateralized phonotactic frequency effects that were observed in the present study and in Vaden et al. (2011).

### **Evidence for association effects**

The Association hypothesis proposes that LIFG is passively activated by sublexical speech segments during word recognition, reflecting exposure to frequently paired motor plans and auditory word representations. The phonotactic frequency was most strongly related to LIFG activity in the most intelligible listening condition, which elicits the greatest response in auditory cortex compared to the other filter conditions (Eckert et al., 2008). The phonotactic frequency result could therefore be dependent on the quality of temporal lobe representations that are passed on to LIFG. Furthermore, the strength of phonotactic frequency effects did not vary with age, which is consistent with an Association hypothesis tenet that speech production maintains auditory-motor associations (Vaden et al., 2011), across the lifespan.

An association-based explanation for our results is also consistent with evidence that speech recognition processes are hierarchically organized within the temporal lobe (e.g., Binder et al., 2000; Davis & Johnsrude, 2003; Okada et al., 2010), even though LIFG has been implicated previously in perceptual aspects of speech processing (e.g., Fridriksson et al., 2008). In the current study and in Vaden et al. (2011), LIFG regions were not significantly activated during word listening, despite significant correlations with phonotactic frequency

(Supplementary Figure). While high phonotactic frequency words increased the responsiveness of LIFG regions, it is unclear the extent to which low phonotactic frequency words suppressed LIFG regions. These results suggest that LIFG activity is contingent upon phonotactic frequency, although appears to be influenced by word intelligibility conditions that elicit increased temporal lobe responsiveness and therefore further consistent with an association explanation.

There is also convincing evidence that inferior frontal cortex lesions do not typically affect speech recognition as severely as production (Hickok, 2009; Hickok et al., 2008; Lotto, Hickok, Holt, 2009; Bates et al., 2003; Scott, McGettigan, Eisner, 2009). Lesion and neuroimaging evidence suggest that both LIFG and anterior insula are important to motor planning and fluent articulation (Dronkers, 1996; Wise, Greene, Büchel, Scott, 1999; Bates et al., 2003; Moser et al., 2009; but see also Bonilha et al., 2006; Hillis et al., 2004 for evidence that LIFG is critical, not insula). The results of aphasia studies also demonstrate that sublexical task deficits occur in conjunction with speech production deficits more often than word comprehension (Miceli et al., 1980; Caplan, Gow, Makris, 1995). Together, these findings suggest that auditory input into frontal regions important for speech planning underlies the phonotactic frequency effects observed in this study.

### Models of speech recognition

While the results of the current experiment are best explained by an association-based mechanism in comparison to a recruitment-based mechanism, they may also be consistent with other models of word recognition. Specifically, increased phonotactic frequency effects for the most intelligible speech condition could relate to the Neighborhood Activation (Luce et al. 2000), the TRACE (McClelland & Elman, 1986), Short List (Norris & McQueen, 2008), or Cohort (Marslen-Wilson, 1987) models of speech recognition in which sublexical information is hypothesized to contribute to word recognition in high intelligibility conditions. We examined the possibility that cohort size contributed to the phonotactic frequency effects observed in this study. Cohort size was estimated for each word by counting the number of phonological transcriptions of words within IPHOD ([www.iphod.com](http://www.iphod.com)) that shared all phonemes prior to the uniqueness point, which is the first phoneme that distinguishes a word from all others. Among the words used in the present study, cohort size was correlated with neighborhood density,  $R^2 = 0.15$ ,  $p < 0.001$ , and phonotactic frequency,  $R^2 = 0.10$ ,  $p < 0.001$ . However, no significant cohort size effects were observed in the LIFG regardless of whether the phonotactic frequency regressor was included in the regression model. In addition, phonotactic frequency effects were not diminished by controlling for cohort size; the LIFG cluster contained 117 voxels in this submodel analysis. This supported our conclusion that phonotactic frequency effects were not confounded with competitive cohort activations or phonological neighborhood activations.

### Limitations of the study

Some design features of the current experiment limit the strength of our conclusions. A subvocalization control condition would more directly preclude the possibility that internal speech drove sublexical activity during the word recognition task. In addition, the word intelligibility manipulations consisted of subtractive low pass filters, whose effects may differ from environmental noise, which are typically additive in nature. It is also possible that phonotactic frequency effects occur specifically when intelligible speech is presented in a noisy background. Background noise was presented across filter conditions in this study, whereas scanner noise was present across the experiment in the Vaden et al., (2011) study. Future studies could be designed to examine the impact of additive noise (e.g., babble, traffic) and the absence of background noise on sublexical processing in LIFG. Readers

should also consider that our participants included older adults with near-normal hearing and good health compared to population norms. Our recruitment procedures screened for conditions that would ensure safe fMRI experimentation, avoid potentially confounding medication effects on the BOLD response, and reduce differences in individual hearing abilities. Nonetheless, the present experiment established basic observations concerning sublexical processing in LIFG, which may potentially inform predictions in future studies with related questions about different sample populations (e.g., hearing impaired, dyslexic, or aphasic adults).

Finally, the specific location of the LIFG results is also important for interpreting the functional significance of the results. An association-based mechanism for the phonotactic frequency results might be expected to occur in the pre-central sulcus and/or anterior insula (Price, 2010). While we observed group effects in the medial pTr/anterior insula, a small pre-central sulcus group effect was present only when an uncorrected statistical threshold of  $p < 0.01$  was used. The most robust effects across the group extend from the pTr into the inferior frontal sulcus. Visual inspection of each subject's results demonstrated that phonotactic frequency effects could be observed in the inferior frontal sulcus, pars triangularis, pars opercularis, anterior insula, and less frequently in the pars orbitalis. We predict that this varied pattern of individual subject effects is due to variation in the primary termination sites of fibers from temporal lobe cortex. To test this prediction, future high resolution imaging studies involving a single subject design could focus on the sensitivity of LIFG regions receiving temporal lobe fiber projections to phonotactic frequency.

## Conclusion

The current study replicated and expanded on the Vaden et al. (2011) phonotactic frequency findings by demonstrating that sublexical processing in LIFG significantly increased for more intelligible speech. We found no relation between age and phonotactic frequency effects in LIFG, which indicates that sublexical processing was relatively stable across the adult lifespan in this cross-sectional study. Together, the results of this study are consistent with the view that auditory-motor associations passively activate sublexical representations in the LIFG during speech recognition.

## 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/National Institute on Deafness and Other Communication Disorders (P50 DC 00422), South Carolina Clinical and Translational Research (SCTR) Institute, with an academic home at the Medical University of South Carolina, National Institutes of Health/National Center for Research Resources (UL1 RR029882), and the MUSC Center for Advanced Imaging Research. This investigation was conducted in a facility constructed with support from Research Facilities Improvement Program (C06 RR14516) from the National Center for Research Resources, National Institutes of Health. We also thank the study participants for assistance with data collection.

## Works Cited

- Aichert I, Ziegler W. Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*. 2004; 88:148–159. [PubMed: 14698739]
- American National Standards Institute. *American National Standard Specifications for Audiometers*. New York: American National Standards Institute; 2004.
- Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage*. 2007; 38:95–113. [PubMed: 17761438]



- Ashburner J, Friston KJ. Unified segmentation. *Neuroimage*. 2005; 26:839–851. [PubMed: 15955494]
- Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*. 2008; 59:390–412.
- Bailey TM, Hahn U. Determinants of wordlikeness: phonotactics or lexical neighborhoods? *Journal of Memory and Language*. 2001; 44:568–591.
- Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, Dronkers NF. Voxel-based lesion-symptom mapping. *Nature Neuroscience*. 2003; 6(5):449–450.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET. Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*. 2000; 10:512–528. [PubMed: 10847601]
- Binder JR, Liebenthal E, Possing ET, Medler DA, Ward BD. Neural correlates of sensory and decision processes in auditory object identification. *Nature Neuroscience*. 2004; 7(3):295–301.
- Bonilha L, Moser D, Rorden C, Baylis GC, Fridriksson J. Speech apraxia without oral apraxia: can normal brain function explain the physiopathology? *NeuroReport*. 2006; 17(10):1027–1031. [PubMed: 16791097]
- Burton MW, Paul T, LoCasto C, Krebs-Noble D, Gullapalli RP. A systematic investigation of the functional neuroanatomy of auditory and visual phonological processing. *NeuroImage*. 2005; 26:647–661. [PubMed: 15955475]
- Callan DE, Kent RD, Guenther FH. An auditory-feedback-based neural network model of speech production that is robust to developmental changes in the size and shape of the articulatory system. *Journal of Speech, Language, and Hearing Research*. 2000; 43:721–736.
- Caplan D, Gow D, Makris N. Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. *Neurology*. 1995; 45:293–298. [PubMed: 7854528]
- Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*. 2002; 3:201–215.
- Davis MH, Johnsrude IS. Hierarchical processing in spoken language comprehension. *Journal of Neuroscience*. 2003; 23:3423–3431. [PubMed: 12716950]
- Dirks DD, Takayanagi S, Moshfegh A, Noffsinger PD, Fausti SA. Examination of the neighborhood activation theory in normal and hearing-impaired listeners. *Ear & Hearing*. 2001; 22:1–13. [PubMed: 11271971]
- Dosenbach NUF, Visscher KM, Palmer ED, Miezen FM, Wenger KK, Kang HC, Burgund ED, Grimes AL, Schlagger BL, Peterson SE. A core system for the implementation of task sets. *Neuron*. 2006; 50:799–812. [PubMed: 16731517]
- Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlagger BL, Peterson SE. Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences of the USA*. 2007; 104(26):11073–11078. [PubMed: 17576922]
- Doupe AJ, Kuhl PK. Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*. 1999; 22:567–631.
- Dronkers NF. A new brain region for coordinating speech articulation. *Nature*. 1996; 384:159–161. [PubMed: 8906789]
- Dubno JR, Dirks DD, Morgan DE. Effects of age and mild hearing loss on speech recognition in noise. *Journal of the Acoustical Society of America*. 1984; 76(1):87–96. [PubMed: 6747116]
- Dubno JR, Lee F, Matthews LJ, Ahlstrom JB, Horwitz AR, Mills JH. Longitudinal changes in speech recognition in older persons. *Journal of the Acoustical Society of America*. 2008; 123(1):462–475. [PubMed: 18177174]
- Eckert MA, Walczak A, Ahlstrom J, Denslow S, Horwitz A, Dubno JR. Age-related effects on word recognition: reliance on cognitive control systems with structural declines in speech-responsive cortex. *Journal of the Association for Research in Otolaryngology*. 2008; 9:252–259. [PubMed: 18274825]
- Eckert MA, Menon V, Walczak A, Ahlstrom J, Denslow S, Horwitz A, Dubno JR. At the heart of the ventral attention system: the right anterior insula. *Human Brain Mapping*. 2009; 30:2530–2541. [PubMed: 19072895]

- Eckert MA, Keren NI, Roberts DR, Calhoun VD, Harris KC. Age-related changes in processing speed: unique contributions of cerebellar and prefrontal cortex. *Frontiers in Human Neuroscience*. 2010; 4(10):1–14. [PubMed: 20204154]
- Folstein MF, Robins LN, Helzer JE. The Mini-Mental State Examination. *Archives of General Psychiatry*. 1983; 40(7):812. [PubMed: 6860082]
- Fridriksson J, Moser D, Ryalls J, Bonilha L, Rorden C, Baylis G. Modulation of frontal lobe speech areas associated with the production and perception of speech movements. *Journal of Speech, Language, and Hearing Research*. 2008; 52(3):812–819.
- Frisch SA, Large NR, Pisoni DB. Perception of wordlikeness: effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language*. 2000; 42:481–496. [PubMed: 21738287]
- Giraud AL, Kell C, Thierfelder C, Sterzer P, Russ MO, Preibisch C, Kleinschmidt A. Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. *Cerebral Cortex*. 2004; 14:247–255. [PubMed: 14754865]
- Goldrick M, Larson M. Phonotactic probability influences speech production. *Cognition*. 2008; 107:1155–1164. [PubMed: 18096149]
- Harris KC, Dubno JR, Keren NI, Ahlstrom JB, Eckert MA. Speech recognition in younger and older adults: a dependency on low-level auditory cortex. *Journal of Neuroscience*. 2009; 29(19):6078–6087. [PubMed: 19439585]
- Heim S, Opitz B, Müller K, Friederici AD. Phonological processing during language production: fMRI evidence for a shared production-comprehension network. *Cognitive Brain Research*. 2003; 16(2): 285–296. [PubMed: 12668238]
- Hickok G. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*. 2009; 21(7):1229–1243. [PubMed: 19199415]
- Hickok G, Poeppel D. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*. 2004; 92:67–99. [PubMed: 15037127]
- Hickok G, Poeppel D. The cortical organization of speech processing. *Nature Reviews Neuroscience*. 2007; 8(5):393–402.
- Hickok G, Okada K, Barr W, Pa J, Rogalsky C, Donnelly K, Barde L, Grant A. Bilateral capacity for speech sound processing in auditory comprehension: evidence from Wada procedures. *Brain and Language*. 2008; 107(3):179–184. [PubMed: 18976806]
- Hillis AE, Work M, Barker PB, Jacobs MA, Breese EL, Maurer K. Re-examining the brain regions crucial for orchestrating speech articulation. *Brain*. 2004; 127(7):1479–1487. [PubMed: 15090478]
- Hollingshead, AB. Four factor index of social status. New Haven, CT: Yale University; 1975.
- Houde JF, Nagarajan SS, Sekihara K, Merzenich MM. Modulation of the auditory cortex during speech: an MEG study. *Journal of Cognitive Neuroscience*. 2002; 14(8):1125–1138. [PubMed: 12495520]
- Humes LE, Christopherson L. Speech identification difficulties of hearing-impaired elderly persons: the contributions of auditory processing deficits. *Journal of Speech and Hearing*. 1991; 34:686–693.
- Jones JA, Munhall KG. Perceptual calibration of F0 production: evidence from feedback perturbation. *Journal of the Acoustical Society of America*. 2000; 108(3):1246–1251. [PubMed: 11008824]
- Jones JA, Munhall KG. Learning to produce speech with an altered vocal tract: the role of articulatory feedback. *Journal of the Acoustical Society of America*. 2002; 113(1):532–543. [PubMed: 12558289]
- Jusczyk PW, Friederici AD, Wessels JMI, Svenkerund VY, Jusczyk AM. Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*. 1993; 32:402–420.
- Jusczyk PW, Luce PA, Charles-Luce J. Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*. 1994; 33:630–645.
- Kent RD, Osberger MJ, Netsell R, Hustedde CG. Phonetic development in identical twins differing in auditory function. *Journal of Speech and Hearing Disorders*. 1987; 52:64–75. [PubMed: 3807347]
- Klein TA, Endrass T, Kathmann N, Neumann J, von Cramon DY, Ullsperger M. Neural correlates of error awareness. *NeuroImage*. 2007; 34:1774–1781. [PubMed: 17185003]

- Kučera, H.; Francis, WN. Computational analysis of present-day American English. Providence: Brown University Press; 1967.
- Levelt WJM, Wheeldon L. Do speakers have access to a mental syllabary? *Cognition*. 1994; 50:239–269. [PubMed: 8039363]
- Lotto AJ, Hickok GS, Holt LL. Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences*. 2009; 13(3):110–114. [PubMed: 19223222]
- Luce PA, Pisoni DB. Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*. 1998; 19:1–36. [PubMed: 9504270]
- Luce PA, Goldinger SD, Auer ET, Vitevitch MS. Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics*. 2000; 62(3):615–625. [PubMed: 10909252]
- Luce PA, Large NR. Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes*. 2001; 16:565–581.
- Macey PM, Macey KE, Kumar R, Harper RM. A method for removal of global effects from fMRI time series. *Neuroimage*. 2004; 22(1):360–366. [PubMed: 15110027]
- Marslen-Wilson WD. Functional parallelism in spoken word-recognition. *Cognition*. 1987; 25:71–102. [PubMed: 3581730]
- Mattys SL, Jusczyk PW. Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception and Performance*. 2001; 27(3):644–655. [PubMed: 11424651]
- Mattys SL, White L, Melhorn JF. Integration of multiple speech segmentation cues: a hierarchical framework. *Journal of Experimental Psychology: General*. 2005; 134(4):477–500. [PubMed: 16316287]
- McClelland JL, Elman JL. The TRACE model of speech perception. *Cognitive Psychology*. 1986; 18:1–86. [PubMed: 3753912]
- McNealy K, Mazziotta JC, Dapretto M. Cracking the language code: neural mechanisms underlying speech parsing. *Journal of Neuroscience*. 2006; 26(29):7629–7639. [PubMed: 16855090]
- McQueen JM. Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*. 2001; 39:21–46.
- Menon V, Adelman NE, White CD, Glover GH, Reiss AL. Error-related brain activation during a Go/NoGo response inhibition task. *Human Brain Mapping*. 2001; 12:131–143. [PubMed: 11170305]
- Menon V, Uddin LQ. Saliency, switching, attention and control: a network model of insula function. *Brain Structure & Function*. 2010; 214:1863–2653.
- Miceli G, Gainotti G, Caltagirone C, Masullo C. Some aspects of phonological impairment in aphasia. *Brain and Language*. 1980; 11:159–169. [PubMed: 7427715]
- Moineau S, Dronkers NF, Bates E. Exploring the processing continuum of single-word comprehension in aphasia. *Journal of Speech, Language, and Hearing Research*. 2005; 48:884–896.
- Moser D, Fridriksson F, Bonilha L, Healy EW, Baylis G, Baker JM, Rorden C. Neural recruitment for the production of native and novel speech sounds. *NeuroImage*. 2009; 46:549–557. [PubMed: 19385020]
- Nordeen KW, Nordeen EJ. Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behavioral and Neural Biology*. 1992; 57(1):58–66. [PubMed: 1567334]
- Norris D, McQueen JM. Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*. 2008; 115(2):357–395. [PubMed: 18426294]
- Obleser J, Wise RJS, Dresner MA, Scott SK. Functional integration across brain regions improves speech perception under adverse listening conditions. *Journal of Neuroscience*. 2007; 27(9):2283–289. [PubMed: 17329425]
- Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*. 1971; 9:97–113. [PubMed: 5146491]
- Okada K, Rong F, Venezia J, Matchin W, Hsieh IH, Saberi K, Serences JT, Hickok G. Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. *Cerebral Cortex*. 2010.1093/cercor/bhp318
- Osnes B, Hugdahl K, Specht K. Effective connectivity demonstrates involvement of premotor cortex during speech perception. *Neuroimage*. 2011; 54:2437–2445. [PubMed: 20932914]

- Papoutsis M, de Zwart JA, Jansma JM, Pickering MJ, Bednar JA, Horwitz B. From phonemes to articulatory codes: an fMRI study of the role of Broca's area in speech production. *Cerebral Cortex*. 2009; 19(9):2156–2165. [PubMed: 19181696]
- Peelle JE, Troiani V, Wingfield A, Grossman M. Neural processing during older adults' comprehension of spoken sentences: age differences in resource allocation and connectivity. *Cerebral Cortex*. 2010; 20:773–782. [PubMed: 19666829]
- Prabhakaran R, Blumstein SE, Myers EB, Hutchison E, Britton B. An event-related fMRI investigation of phonological-lexical competition. *Neuropsychologia*. 2006; 44(12):2209–2221. [PubMed: 16842827]
- Price CJ. The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*. 2010; 1191:62–88. [PubMed: 20392276]
- Righi G, Blumstein SE, Mertus J, Worden MS. Neural systems underlying lexical competition: an eye tracking and fMRI study. *Journal of Cognitive Neuroscience*. 2009; 22(2):213–224. [PubMed: 19301991]
- Rodd JM, Davis MH, Johnsrude IS. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*. 2005; 15:1261–1269. [PubMed: 15635062]
- Rodd JM, Longe OA, Randall B, Tyler LK. The functional organization of the fronto-temporal language system: Evidence from syntactic and semantic ambiguity. *Neuropsychologia*. 2010; 48:1324–1335. [PubMed: 20038434]
- Saffran JR, Newport EL, Aslin RN. Word segmentation: the role of distributional cues. *Journal of Memory and Language*. 1996; 35:606–621.
- Saffran JR, Newport EL, Aslin RN, Tunick RA, Barrueco S. Incidental language learning: listening (and learning) out of the corner of your ear. *Psychological Science*. 1997; 8:101–195.
- Saffran JR, Newport EL, Johnson EK, Aslin RN. Statistical learning of tone sequences by human infants and adults. *Cognition*. 1999; 70:27–52. [PubMed: 10193055]
- Salek-Haddadi A, Diehl B, Hamandi K, Merschhemke M, Liston A, Friston K, Duncan JS, Fish DR, Lemieux L. Hemodynamic correlates of epileptiform discharges: an EEG-fMRI study of 63 patients with focal epilepsy. *Brain Research*. 2006; 1088:148–166. [PubMed: 16678803]
- Scott SK, McGettigan C, Eisner F. A little more conversation, a little less action - candidate roles for the motor cortex in speech perception. *Nature Reviews Neuroscience*. 2009; 10:295–302.
- Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD. Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*. 2007; 27(9):2349–2356. [PubMed: 17329432]
- Siok WT, Jin S, Fletcher P, Tan LH. Distinct brain regions associated with syllable and phoneme. *Human Brain Mapping*. 2003; 18:201–207. [PubMed: 12599278]
- Snyder HR, Feigenson K, Thompson-Schill SL. Prefrontal cortical response to conflict during semantic and phonological tasks. *Journal of Cognitive Neuroscience*. 2007; 19(5):761–775. [PubMed: 17488203]
- Sommers MS. The structural organization of the mental lexicon and its contribution to age-related changes in spoken word recognition. *Psychology and Aging*. 1996; 11:333–341. [PubMed: 8795062]
- Storkel HL. Methods for minimizing the confounding effects of word length in the analysis of phonotactic probability and neighborhood density. *Journal of Speech, Language, and Hearing Research*. 2004; 47:1454–1468.
- Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proceedings of the National Academy of Sciences, USA*. 1997; 94:14792–14797.
- Tombaugh TN, McIntyre NJ. The Mini-Mental State Examination: a comprehensive review. *Journal of the American Geriatrics Society*. 1992; 40:922–935. [PubMed: 1512391]
- Tourville JA, Reilly KJ, Guenther FH. Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*. 2008; 39(3):1429–1443. [PubMed: 18035557]
- Tyler LK, Shafto MA, Randall B, Wright P, Marslen-Wilson WD, Stamatakis EA. Preserving syntactic processing across the adult life span: the modulation of the frontotemporal language

- system in the context of age-related atrophy. *Cerebral Cortex*. 2010; 20:352–364. [PubMed: 19505991]
- Vaden, KI.; Halpin, HR.; Hickok, GS. Irvine Phonotactic Online Dictionary, Version 1.4. 2009. [Data file]. Available from <http://www.iphod.com>
- Vaden KI, Piquado T, Hickok GS. 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 Available online. 10.1162/jocn.2011.21620
- Vaden KI, Muftuler LT, Hickok G. Phonological repetition-suppression in bilateral superior temporal sulci. *Neuroimage*. 2010; 49(1):1018–23. [PubMed: 19651222]
- Vitevitch MS. The influence of sublexical and lexical representations on the processing of spoken words in English. *Clinical Linguistics and Phonetics*. 2003; 17(6):487–499. [PubMed: 14564833]
- Vitevitch MS, Luce PA. When words compete: levels of processing in spoken words. *Psychological Science*. 1998; 9(4):325–329.
- Vitevitch MS, Luce PA, Pisoni DB, Auer ET. Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*. 1999; 68:306–311. [PubMed: 10433774]
- Wilson SM. Speech perception when the motor system is compromised. *Trends in Cognitive Sciences*. 2009; 13(8):329–330. [PubMed: 19646917]
- Wilson SM, Isenberg AL, Hickok G. Neural correlates of word production stages delineated by parametric modulation of psycholinguistic variables. *Human Brain Mapping*. 2009 Published Online: April 2009.
- Wise RJS, Greene J, Büchel C, Scott SK. Brain regions involved in articulation. *Lancet*. 1999; 353:1057–1061. [PubMed: 10199354]
- Woodcock, RW.; McGrew, KS.; Mather, N. Woodcock-Johnson Test of Achievement (WJ-III-R). Itasca, IL: Riverside; 2001.
- Wong PCM, Jin XJ, Gunasekera GM, Abel R, Lee ER, Dhar S. Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia*. 2009; 47(3):693–703. [PubMed: 19124032]
- Zaehle T, Geiser E, Alter K, Jancke L, Meyer M. Segmental processing in the human auditory dorsal system. *Brain Research*. 2007; 1220:179–190. [PubMed: 18096139]



### Research Highlights

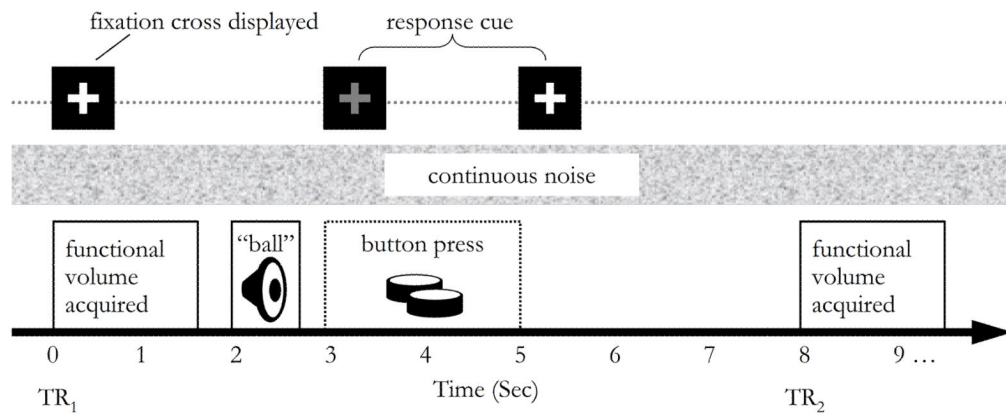
Left inferior frontal cortex is particularly responsive to common speech sounds.

We probed this finding by manipulating speech intelligibility and varying age.

The most intelligible words produced the strongest frontal response to speech sounds.

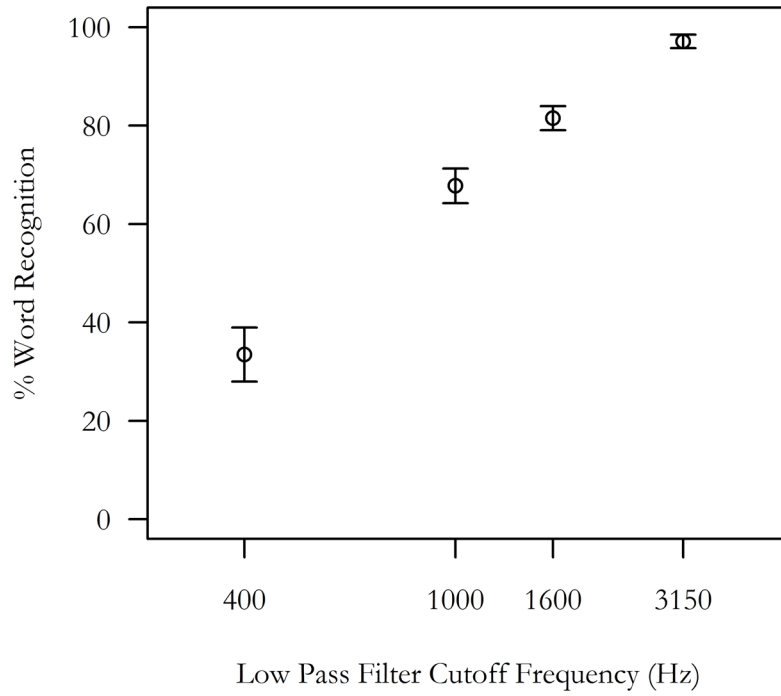
Age (19–79) did not predict the speech sound effect on inferior frontal cortex.

These task specific frontal results suggest an associative not a recruitment impact.

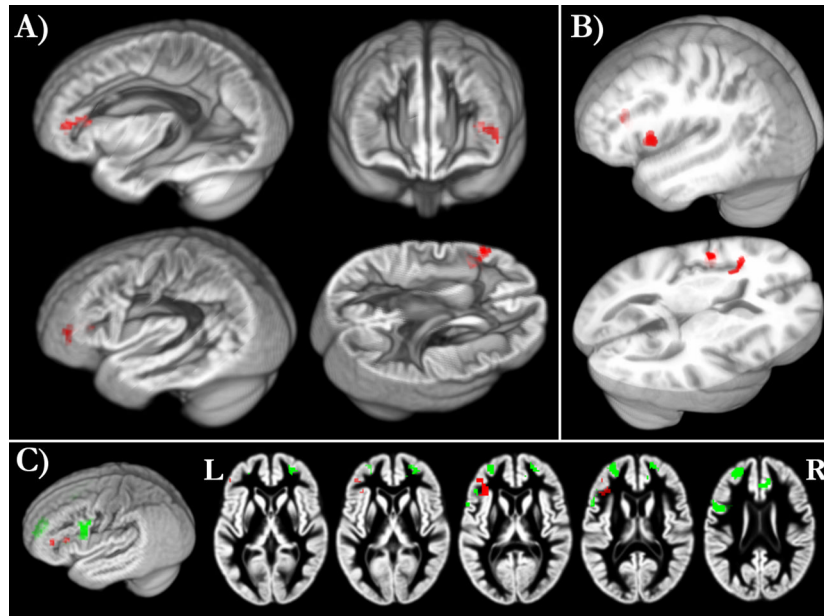


**Figure 1.**  
Presentation timing

A sparse-sampling design was used to collect one functional volume every 8 sec. At the beginning of each trial, a white fixation cross appeared onscreen ( $t = 0$  sec). The participant was presented with a single word (e.g., “ball”), 0.4 sec after the scanner noise ended ( $t = 2$  sec). The fixation cross turned red to cue the participant to button press during a two-second response interval ( $t = 3$ – $5$  sec). The timing ensured that participants were able to hear and attended to the task in the absence of scanner noise. Digitally generated broadband noise was played continuously in the background throughout the experiment, to minimize the impact on word recognition of individual differences in hearing.



**Figure 2.**  
Intelligibility and speech recognition  
Word recognition responses indicated that the low pass filter manipulation resulted in parametric increases in word intelligibility. Mean percent word recognition responses are plotted for each low pass filter cutoff frequency. Error bars denote  $\pm 1$  SEM.

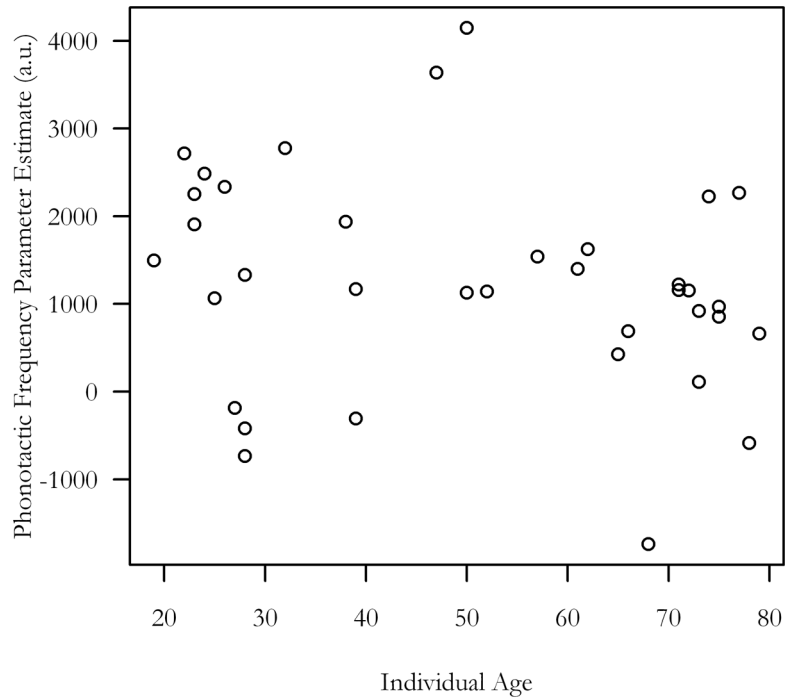


**Figure 3.**

Positive correlation with phonotactic frequency in the pars triangularis

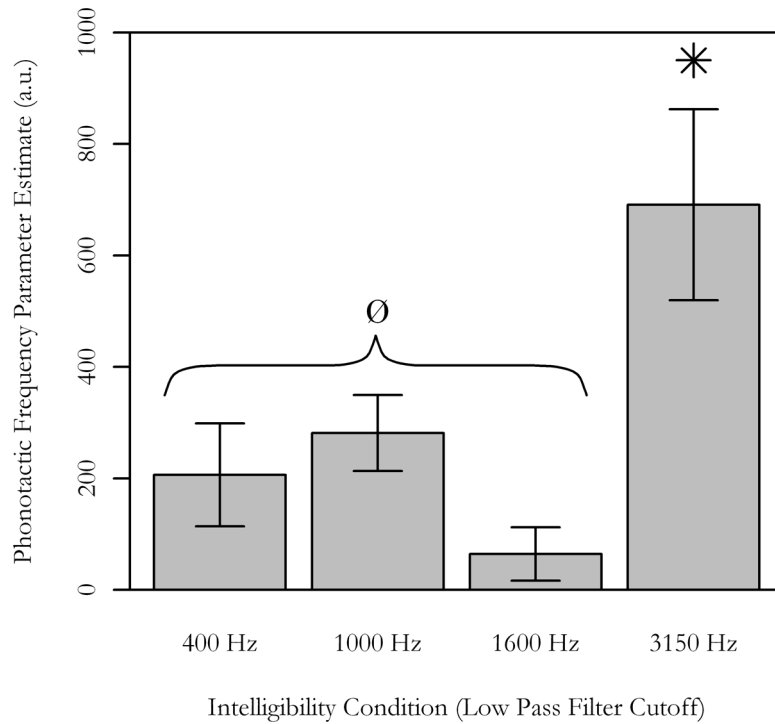
Words that consisted of more frequent phoneme pairs elicited the highest activity in the PTR.

A) Voxels shown in red passed a threshold of  $t(34) = 3.35$ ,  $p = 0.001$  uncorrected, with cluster-size corrected  $p < 0.001$ , displayed on the rendered grey matter template in study-specific space. B) Joint phonotactic frequency effects in the current results and Vaden et al. (2011). In order to compare the significant positive correlations in both studies within a common stereotactic space, we segmented the template used in Vaden et al. (2011) then normalized the gray matter template and results from the current study into that space using ANTS (<http://www.picsl.upenn.edu/ANTS>). We assessed the joint statistic maps at a threshold of  $p = 0.001$  (uncorrected), based on the conjunction of each study's positive phonotactic frequency statistic maps, thresholded at  $p = 0.01$ . Different thresholds were needed for each study since they differed in degrees of freedom:  $t(33) = 3.36$  for the current study and  $t(16) = 3.69$  for Vaden et al (2011). Significant voxels (red) formed two clusters in the left IFG extending from the lateral pars triangularis into medial pars opercularis, and lateral inferior frontal sulcus to medial anterior insula. This illustrates voxels in both studies in a common stereotactic space that increased activity in response to high phonotactic frequency words, with the caveat that different preprocessing steps were used in each study, including normalization methods and smoothing kernel. We elaborate on differences that related to the significance of this replication in the discussion section. C) The reaction time control analysis demonstrated a significant positive correlation between reaction time and activity in neighboring frontal sub-regions including pars orbitalis, lateral anterior insula, and motor cortex (green), with only one voxel overlap with the phonotactic frequency effect (red). Voxels were thresholded at  $t(33) = 3.36$ ,  $p = 0.001$  uncorrected, cluster size  $p = 0.05$  corrected (30 voxels minimum).



**Figure 4.**  
Age and sublexical activity  
Age was not significantly related to the LIFG phonotactic frequency parameter estimate, from first-level analyses ( $R^2 = 0.05$ ,  $p = 0.2$ ).





**Figure 5.**

Word intelligibility and sublexical activity

The phonotactic frequency effect in the LIFG was dependent on word intelligibility. Pairwise tests revealed that phonotactic frequency had the greatest impact on LIFG activity for the 3150 Hz condition, compared to each of the other conditions. The ROI mask was functionally defined using the full model phonotactic effect (76 voxels in LIFG). The phonotactic frequency beta values (shown as arbitrary units, a.u.) were modeled separately by word intelligibility condition then averaged from each individual's resultant first-level contrast within the ROI. Error bars denote  $\pm 1$  SEM. Single sample t-tests demonstrated that phonotactic frequency effects were significantly greater than zero in the 400 Hz, 1000 Hz, and 3150 Hz intelligibility conditions,  $t(35) > 2.15$ ,  $p$ 's  $< 0.04$ , but not in the 1600 Hz condition,  $t(35) = 1.35$ ,  $p = 0.18$ .