Published in final edited form as:

Neuroimage. 2010 January 1; 49(1): 1018–1023. doi:10.1016/j.neuroimage.2009.07.063.

Phonological Repetition-Suppression in Bilateral Superior Temporal Sulci

Kenneth I. Vaden Jr., Ph.D.,

Department of Cognitive Sciences, University of California at Irvine

L. Tugan Muftuler, Ph.D., and

Tu & Yuen Center for Functional Onco-Imaging, University of California at Irvine

Gregory Hickok, Ph.D.¹

Department of Cognitive Sciences, University of California at Irvine

Abstract

Evidence has accumulated that posterior superior temporal sulcus (STS) is critically involved in phonological processing during speech perception, although there are conflicting accounts regarding the degree of lateralization. The current fMRI experiment aimed to identify phonological processing during speech perception through repetition-suppression effects. Repetition-suppression occurs when brain activity decreases from repetitive presentation of stimulus characteristics, in regions of cortex that process those characteristics. We manipulated the degree of phonological repetition among words in short lists to obtain systematic decreases in brain response, indicative of phonological processing. The fMRI experiment presented seventeen participants with recorded wordlists, of low, medium, or high phonological repetition, defined by how many phonemes were shared among words. Bilaterally, middle STS demonstrated activity differences consistent with our prediction of repetition-suppression, as responses decreased systematically with each increase in phonological repetition. Phonological repetition-suppression in bilateral STS converges with neuroimaging evidence for phonological processing, and word deafness resulting from bilateral superior temporal lesions.

1. Introduction

The goal of the current study was to functionally identify the cortical bases of phonological processing in spoken word recognition. Phonological processes in speech perception result in discretely perceived utterances, based on the continuous and variable vocal sounds that we hear. A range of evidence from various sources has converged on the view that the superior temporal sulcus (STS) is a critical site in phonological processing during perception, although there are disputes about the degree of laterality (Hickok & Poeppel, 2007; Rauschecker & Scott, 2009). We will briefly review previous functional imaging studies relevant to this issue and then present a new experiment designed to map regions within the STS that are involved in phonological aspects of perception.

© 2009 Elsevier Inc. All rights reserved.

1Address for manuscript correspondence: Greg Hickok, University of California, SSPA 4109, Mail Code: 5100, Irvine, CA 92697, Telephone: 00-1-949-824-1409, Fax: 00-1-949-824-2307, greg.hickok@uci.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.

Early neuroimaging studies on the perception of auditorily presented speech consistently reported bilateral activity in superior temporal regions when contrasted with a baseline of silence or scanner noise (Wise et al., 1991; Zatorre et al., 1992; Mazoyer et al., 1993; Binder et al., 1994; Mellet et al., 1996). However, activations highlighted with these contrasts could involve a range of processing levels from acoustic processes to lexical-semantic or sentence-level computations. Based on those preliminary results, it was not clear which aspects of speech or audition were involved in the highlighted brain activity.

Subsequent studies attempted to identify those subregions that were specifically involved in phonological level processing during speech recognition. One common approach was to use cognitive subtractions, which sought to isolate speech-specific activations by contrasting speech sounds with a range of non-speech acoustic control stimuli such as noise bursts, tones, reversed speech, and spectrally rotated speech. The logic here is that speech-specific phonological systems could be isolated by subtracting out nonspeech acoustic processing. Despite the fact that many authors argued for a dominant or exclusive role for left superior temporal areas in speech recognition, most of these studies reported peaks in left and right superior temporal regions, see Price et al. (1996); Vouloumanos et al. (2001), Narain et al. (2003), Aleman et al. (2005), Meyer, Zysset, von Cramon, Alter (2005), and see Hickok & Poeppel (2007) for a summary. While it is true that the majority of cognitive subtraction-based studies found spatially broader activity in left than right superior temporal lobe, there is no a priori reason to believe that less focal activation translates to greater specialization (Hickok & Poeppel, 2007; Okada & Hickok, 2006).

One complication in interpreting activations highlighted by contrasting speech and nonspeech is that it relies on the assumption that only speech-selective regions are critical or specialized for speech processing (for example, Uppenkamp et al., 2006). This assumption may not be correct, because many mechanisms involved in speech are not exclusively recruited for speech (Price, Thierry, & Griffiths, 2005; Hickok & Poeppel, 2007). Speech sounds may be processed in multi-potent auditory regions (Zatorre & Gandour, 2008), so speech versus non-speech contrasts could weaken speech-related activity in regions with non-exclusive responses (Okada & Hickok, 2006). Even if distinct, dedicated speech resources exist in cortex, it is possible that these networks are intermixed with non-speech systems making it difficult to resolve speech-specific processing given typical functional imaging design and analysis schemes (spatial smoothing, group averaging, etc.). As a result, subregions that are critical to speech recognition may not reveal different responses to speech and nonspeech. Put another way, regions that show no difference in activation between speech and non-speech control stimuli may nonetheless be critically involved in speech processing and may even contain systems that are specialized for speech processing.

Another disadvantage of the cognitive subtraction methodology arises when interpreting differences between linguistic and acoustic processes. Contrasting speech with nonspeech often conflates many linguistic distinctions, making it unclear how to characterize highlighted activity (Okada & Hickok, 2006; Poeppel, Idsardi, van Wassenhove, 2008). Speech-nonspeech differences could be exclusively acoustic, phonetic, phonological, lexical, or some combination of linguistic and perceptual dimensions. Even when designs try to isolate a particular aspect of speech, there may be spreading activation that causes unintended language processes to enter into the cognitive subtraction. The interpretation of activity differences becomes even more convoluted when different tasks are applied to the speech and nonspeech controls (see a review of task effects in Hickok & Poeppel, 2007).

Some neuroimaging experiments have used an alternative approach; they manipulated phonological variables to modulate brain activity within phonological processing during speech recognition, rather than comparing speech and nonspeech activity. These experiments

have typically added support to the bilateral speech perception account (Benson et al., 2001; Wilson and Iacoboni, 2006; Obleser et al., 2008; Formisano et al, 2008), but others have produced mixed results. For example, Okada and Hickok (2006) observed increased activity in bilateral posterior STS when subjects listened to words with high phonological neighborhood density versus words with low density. In contrast, Prabhakaran et al. (2006) only found consistent density effects in left SMG. The current study aimed to use a different phonological manipulation to resolve the question of left or bilateral phonological activity during speech perception.

In the current fMRI experiment, we used repetition-suppression effects to highlight phonological processing. Repetition-suppression occurs when brain regions sensitive to a stimulus characteristic receive repetitive input, resulting in decreased response to that characteristic (Grill-Spector, Henson, Martin, 2006). The precise mechanisms underlying neural repetition effects are not known, but previous fMRI experiments have applied repetition-suppression to isolate or disambiguate speech processes for characteristics such as linguistic content and prosody (Dehaene-Lambertz et al., 2006), bilingual language resources (Klein & Zatorre et al., 2006), and lexical semantics (Rissman, Eliassen, Blumstein, 2003). Repetition-suppression effects have been observed when subjects passively listened to syllables (Hasson, Skipper, Nusbaum, Small, 2007) or sentences (Hasson, Nusbaum, Small, 2006), named pictures (Graves et al., 2008), or performed lexical decisions on spoken words that were studied prior to scanning (Gagnepain et al., 2008). We predicted that phonological repetition in word lists would result in repetition-suppression effects, allowing us to highlight subregions of cortex that process that information.

Performance on a range of speech tasks provides evidence that speech mechanisms are sensitive to phonological repetition, often worsening performance when greater numbers of phonemes are shared among words. Such effects on the perceptual side include short term priming, when repeating a spoken word is slowed by having heard a similar sounding prime word milliseconds earlier (Dufour & Peereman, 2003). Long term phonological priming has also been observed, when lexical decisions slowed on words several minutes after hearing a prime word (Sumner & Samuel, 2007). Phonological repetition has also been found to degrade phoneme recognition in selective adaptation experiments (Eimas & Corbit, 1973; Samuel, 1997). In memory, short term recall is reduced when words sound similar to one another (Baddeley, 1966; Sperling & Speelman, 1967; Sperling, 1968). Production of speech also shows effects of phonological repetition among tongue-twisters and spoonerisms, when similar phoneme sequences become interposed in articulated utterances. Many aspects of speech show effects of phonological repetition, which provides evidence that cortical resources sensitive to phonological content may systematically change their dynamics when confronted with repetitive input.

The current fMRI study targeted phonological processing activity during speech recognition, using repetition-suppression effects. Based on earlier neuroimaging and psycholinguistic results showing repetition effects in speech perception, we predicted that phonological processing regions would demonstrate decreasing activity related to increasing phonological repetition in wordlists.² We predicted that these phonological repetition-suppression effects would occur bilaterally, in posterior portions of superior temporal lobe, on the basis of earlier neuroimaging evidence for activity related to lexical-phonological processing in these regions.

²We acknowledge that phonological repetition effects could reflect decreased load-dependent processing activity in STS. For example, greater repetition among words may increase the predictability of phonemes, thereby facilitating phoneme parsing during word segmentation, or similar processes. Thus, a load-dependent explanation could also predict reduced activity for increasing phonological repetition, which is completely consistent with our claim that phonological processing activity is susceptible to the repetition effects.

2. Method

Participants

Seventeen volunteer subjects, aged 19 to 29 (Mean = 23.7, SD = 3.75) years participated in the experiment. There were eight males and nine females. All participants were right-handed, native English speakers, free of neurological disease and had normal hearing by self report. All subjects gave informed consent under the protocol approved by the Institutional Review Board of the University of California at Irvine.

Procedure

Each subject participated in a single experiment session lasting approximately 1 hour at the Phillips 3T scanner at University of California at Irvine, and was paid \$30 for participation. Informed consent and health screening were obtained just prior to the session. Volume was adjusted to comfortable levels prior to initiating the experiment. The subjects were instructed to listen to each wordlist, and press a button only if that list contained one or more pseudowords. Before the experiment began, they were informed that few trials required a response, thus the majority of trials only required them to listen carefully and press no buttons. Before each functional run, we asked subjects if they could hear the stimuli clearly, then adjusted the volume or equipment if necessary.

Design

During the fMRI experiment, subjects were presented auditorily with wordlists comprised of low, medium, or high levels of phonological repetition. Phonological repetition was manipulated by how many phonemes were shared in each wordlist. This is explained in more detail in the stimuli section, and examples of different repetition level wordlists can be seen in Table 1.

Experiment sessions consisted of eight runs of blocked trials, although one subject only completed six runs because of time limitations. Each run contained nine wordlists in each repetition level: low, medium, or high, for a total of 27 experiment trials per run. Two catch trials containing pseudowords also occurred randomly during each run, to motivate participants to listen carefully to the words. Catch trials each presented two words and two pseudowords with low or medium phonological repetition³. Conditions were pseudo-randomly ordered with a constraint that no conditions repeated back to back. Table 1 summarizes trials and volumes collected per condition.

Each blocked trial had a jittered duration equal to 8.4, 10.5, or 12.6 seconds (4, 5, 6 TR), and was synchronized to start with the acquisition of a functional image. Trials included a relatively fixed-length auditory wordlist presentation, followed by a variable rest period. During wordlist presentation, words were presented with a silent ISI of 150 milliseconds, so the entire stimulus presentation lasted 3.17 seconds on average, equalized across conditions. High repetition trial duration Mean = 3169 msec, SD = 12; medium trial duration Mean = 3163 msec, SD = 3; and low trial duration Mean = 3163 msec, SD = 3. Figure 1 illustrates the jittered block design.

³Omitting a high-repetition catch trial (one repeating pseudoword) might allow subjects to form a no-response bias in the task since repeating item lists never contained pseudowords. However, our design fostered the contingent expectation that catch trials containing a repeating pseudoword might occur sometime in the experiment. Medium and low repetition catch trials were distributed unevenly within each run, to reduce predictability and associations between repetition level and pseudoword detection. There were only 16 catch trials out of 232, providing scant evidence of a missing repeating pseudoword catch trial condition. Since catch trials were unpredictable on the basis of repetition, the best strategy would be to direct attention to the presented words instead of repetition form. On a related note, an fMRI study by Horner & Henson (2008) found that priming in performance and repetition-suppression (RS) in frontal lobe change when associations are learned, while perceptual RS effects appear unaffected by associations. Thus, if subject associated single repeating item lists with pseudoword absence, it might not affect phonological RS during perception.

Stimuli

Phonological repetition was manipulated through the number of shared or unshared phonemes among four CVC words that were selected to create wordlists with fixed low, medium, or high levels of phonological repetition. Low phonological repetition wordlists contained four dissimilar sounding words that shared neither consonant-vowel (CV) beginnings, nor VC endings.⁴ Medium repetition wordlists contained four similar sounding words that shared CV beginnings, so only the final consonant changed among words. Medium repetition wordlists did not rhyme, since CV onsets cross the syllabic onset-nucleus boundary. This also avoided conflating phonological repetition with rime repetition, so effects were less likely to be related to syllable-structure processing. Finally, high phonological repetition wordlists each consisted of a repeating word, simply presented four times in a row.

We balanced lexical factors unrelated to our hypothesis by using the same pool of 112 words to produce all three conditions. The word pool consisted of 28 sets of four phonologically similar words, sharing CV beginnings. Low and medium phonological repetition wordlists were constructed by reordering those 28 sets to satisfy the CV beginning or VC ending constraints, as described above. All words appeared equally often in those two conditions. For each high phonological repetition wordlist, where a word repeated, the word was chosen so each CV beginning was heard an equal number of times throughout the experiment. Another constraint on word selection for high phonological repetition trials was to select the word with the lowest total previous presentations during earlier trials, which could prevent familiarity developing from unbalanced exposure.

The selection and recording process ensured natural-sounding recordings with low variability. Stimuli were chosen on the basis of shared phonemes using the Irvine Phonotactic Online Dictionary (Vaden, Hickok & Halpin, 2005), excluding any words with plural or past tense endings. The voice of the speaker was an adult, male native English speaker. Each word was recorded more than ten times using a Shure amplifier and PC equipped with Audacity software, which saved recordings to a single channel wav file at 44.1 kHz sampling rate. We selected recorded word tokens falling closest to the mean duration, then RMS normalized their amplitudes. Edited word recordings had a mean duration of 678 ms (SD = 48 ms). The longest word recording, *lobe*, lasted 810 ms, and the shortest word, *bat*, lasted 495 ms. These recordings were homogenous enough to sound like recorded lists, instead of individually recorded items, when they were presented sequentially in the experiment trials.

fMRI scan procedure and preprocessing

We used the 3T Phillips Inera Achieva MRI scanner at the Research Imaging Center, University of California at Irvine to collect structural and functional images. Anatomical and functional images were oriented to the anterior commissure-posterior commissure line. Anatomical, 1.5mm³ isomorphic images were collected using a T1 weighted sequence following the eight experiment runs. Functional volumes consisted of 2.3 × 2.3 × 3mm voxels, 34 slices for whole brain coverage, acquired in interleaved order with no gap. Other specifications for the EPI sequence are: TR = 2.1s, TE = 26ms; Flip Angle = 90; FOV = 200; 150 volumes acquired in 315 sec per run. The scanner was equipped with an eight channel SENSE head-coil, which increased sensitivity to temporal lobe, though SENSE factor = 0. Cogent 2000 (Romaya, 2003) synchronized sound delivery through Resonance Technologies MR-compatible headphones. Figure 1 illustrates the jittered block design, with stimulus presentation synchronized with scanner pulse delivery. MR compatible response boxes were used for subjects to press a button when they detected pseudowords. Although responses were

⁴Low phonological repetition wordlists only shared one vowel among two of the four words (“cat, lap”), one time per run. Put another way, in each run 106/108 phonemes did not repeat in a particular CVC position in these lists.

not recorded due to a technical error, we ensured that subjects were alert and responsive by talking to them on the intercom between runs.

Preprocessing functional data and statistical analyses were performed using SPM5 (Wellcome Department of Imaging Neuroscience). Prior to analysis, we performed slice-timing correction, motion correction in six dimensions, co-registration aligned the anatomical to the middle functional image in the series. Next, segmentation and normalization fitted anatomical and functional images to the MNI template. Spatial smoothing was performed on functional scans with a 5mm FWHM Gaussian kernel. LMGS, voxel-level Linear Model of the Global Signal (Macey et al., 2004) detrended global mean signal fluctuations from the preprocessed functional images that were presumably related to respiration or heart rate.

The final step of preprocessing was an outlier detection algorithm which implemented two methods to independently detect outliers at the volume level, on a run-by-run basis. The first method searched for volumes whose global intensity values exceeded 2.75 standard deviations from the mean. The second method detected volumes that contained an excessive number of extreme voxel values. A voxel's intensity was extreme whenever its intensity exceeded 3.1 standard deviations from its mean timecourse. Volumes that contained a number more than 2.35 standard deviations from the average number of extreme voxels in a volume were excluded. Optimal standard deviation cutoffs were determined empirically by selecting values that maximized signal in voxels in the vicinity of primary auditory cortex for all listening conditions minus rest in the first run of each subject's data. Each method's results were submitted to the GLM as a nuisance variable. Our noise detection algorithm identified 5.5 images per run, or 3.68% of the total functional volumes collected. Only 12% of the extreme volumes identified were detected by both, indicating that the two approaches were largely sensitive to different noise sources.

fMRI analyses

We performed a parametric analysis to localize phonological repetition-suppression and repetition-enhancement effects using SPM5 (Wellcome Department of Imaging Neuroscience). In the General Linear Model, all trials were modeled as blocked events from a single listening condition, with a single parameter that coded [1, 2, 3] for increasing phonological repetition levels. Nuisance regressors contained catch trials, run-wise constants, six motion-correction parameters, and the two extreme volume vectors from the preprocessing step described earlier. Catch trials were excluded from analysis. The resultant t-statistic map reflects the variance accounted for by the phonological repetition parameter regressor, as it correlated with the BOLD signal.

Individual t-static maps were submitted to a random-effects analysis, to find regions that exhibited phonological repetition-suppression across subjects. Repetition-suppression predicted that responses would be highest for the *Low* phonological repetition wordlists since they did not repeat phonemes, and lowest for *High* repetition wordlists, which repeated all phonemes. Although we did not have specific hypotheses concerning repetition-enhancement, we report regions that showed opposite effects to suppression (low < medium < high repetition) that were obtained by simply applying negative t-score thresholds to the same group t-statistic maps. Group t-statistic maps were submitted to a voxel threshold of $p = 0.001$ and corrections for multiple comparisons were performed at the cluster level, with cluster $p = 0.05$ (corrected; extent threshold = 15).

3. Results

Based on repetition-suppression, we predicted that phonological processing activity during perception would decrease as phonological repetition increased, consistent with the pattern:

High < Medium < Low. In line with our main hypothesis, we found significant clusters that demonstrated these changes in bilateral superior temporal sulci, shown in Figure 2. Examination of estimated timecourses for both STS sites revealed that responses to speech were positive, with amplitudes that systematically lowered with each increase in phonological repetition. Several regions including the left anterior cingulate gyrus, left precuneus, and bilateral SMG demonstrated repetition-enhancement, with higher responses to more repetitive or predictable phonological contents. Table 2 summarizes the regions with significant repetition-suppression or repetition-enhancement effects.

4. Discussion and Conclusion

We found phonological repetition-suppression in bilateral STS, which supports the view that these regions are recruited for phonological processing during speech recognition. As phonological repetition within lists increased, there were significant decreases in activity along middle to posterior portions of STS. This was significant since previous neuroimaging speech experiments have located activity in the vicinity of this region, but gave conflicting accounts with regard to lateralization. While bilateral activity was observed, visual examination of the data suggest that a broader swath of the left STS was activated relative to the right STS. Although one may be tempted to interpret this as evidence for left dominant phonological processing, this is not necessarily the case. In fact, one could conceivably argue that the right hemisphere is more efficient in processing phonological forms and therefore yields less activation. Alternatively, individual anatomical differences across the left versus right STS may lead to differential averaging patterns across the two sites. For example if activations in the right STS tend to be more variable than in the left STS from one subject to the next, averaging across subjects will result in apparently “weaker” activation in the right STS. In short, for various reasons we must resist the temptation to interpret wider activation as more important functionally. Thus, our repetition-suppression finding converges with neuroimaging studies reporting bilateral phonological processing activity, and lesion evidence from patients with word deafness.

Repetition suppression was observed in regions outside of the STS. This could be indicative of a broader network involved in phonological processing although it is unclear what the functional role of these other areas are. For example, it is possible that attentional processes are modulated as a function of the degree of phonological repetition: more repetition renders the stimuli more predictable, which may reduce attentional load. Activations outside of STS therefore may reflect attentional mechanisms and/or predictive coding mechanisms (see for example: van Wassenhove, Grant & Poeppel, 2005; Sanders & Poeppel, 2006; Sabri, et al., 2008; Schiller, et al., 2009). Evidence that the STS activations reflect perceptual-phonological processes rather than purely some meta-phonological process (such as attention) come from the observation that perceptual deficits have been reported following bilateral damage to the STS region (word deafness) and from the previous observation that the STS is sensitive to lexical-phonological manipulations (Okada & Hickok, 2006). Thus converging evidence helps constrain our interpretation of the STS activation. We also found repetition-enhancement in regions that were consistent with previous auditory fMRI studies that observed repetition-enhancements (e.g. Bergerbest, Ghahremani, Gabrieli, 2004; Hasson, Nusbaum, Small, 2006; Horner & Henson, 2008).

A related fMRI experiment by Gagnepain et al. (2008) compared activity while subjects listened to previously studied versus novel recorded stimuli in a lexical decision task. They found interactions between lexicality and priming effects in subregions including bilateral STS sites, which demonstrated significantly lower responses to primed versus unprimed words. One potential limitation of priming each word by itself is that aspects other than phonological content are repeated (e.g. acoustic, syntactic, semantic) and may prime other processes that are

sensitive to repetition. Instead of priming items through repeated presentations, our experiment manipulated the extent of phonological repetition and thereby avoided repetition in other unrelated aspects of speech in 2/3 conditions. Thus, our results clarify the findings of Gagnepain et al. (2008), and add support to the conclusion that bilateral STS sites are sensitive to phonological repetition during spoken word perception.

Phonological processing activity in bilateral superior temporal regions is consistent with reports of the bilateral temporal stroke pattern often associated with word deafness (Stefanatos, 2008). The characteristic of word deafness is that hearing is generally spared (ie. patients have approximately normal pure tone thresholds), while sufferers can no longer understand spoken words. However, reports often find complex auditory deficits not revealed by audiogram accompanying the inability to perceive words, demonstrating that phonological processing from an auditory speech source involves computations that are not speech-selective (Griffiths, Rees & Green, 1999; Poeppel, 2001; Stefanatos, 2008).

A limitation of the current study was that we were focused exclusively on phonological processing activity during spoken word recognition. Future studies might use a similar phonological repetition manipulation to refine the extent of functional overlap between acoustic and phonological processes in cortex. Bilateral middle STS revealed phonological sensitivity in our study, but these areas also decreased activity as acoustic distortions were modulated over heard words in a recent fMRI study by Obleser et al. (2008). Manipulating phonemes the way that we did requires corresponding acoustic differences – so a demonstration of secondary acoustic processing in the same vicinity indicates that our observations may reflect processing of complex auditory information⁵. Investigating the degree of functional overlap between acoustic and phonological processing within subjects is an important next step in clarifying our understanding of bilateral STS activity and the extent that phonological and acoustic processes are dissociable in cortical activity.

In conclusion, this experiment added new evidence that bilateral superior temporal sulci are recruited for phonological processing in speech, using phonological repetition-suppression. Increased phonological repetition resulted in successive activity decreases in bilateral STS, a result that converges with some previous neuroimaging speech experiments and the bilateral lesion pattern associated with word deafness. Instead of assuming functional disconnections between speech and nonspeech auditory processes, we suggest that the relationships between them will be an important new avenue in understanding phonological processes in perception, and language in the brain.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was funded by NIH Grant D 003681. Thank you to Stephen Wilson and Kai Okada for analysis advice, and Emily Grossman for design suggestions. We also would like to acknowledge the helpful feedback from our reviewers, which greatly improved the manuscript.

⁵Although there has been an extensive debate in the cognitive science literature, there is little evidence that phonological processes in perception are not entirely auditory in nature.

References

- Aleman A, Formisano E, Koppenhagen H, Hagoort P, de Haan EHF, Kahn RS. The functional neuroanatomy of metrical stress evaluation of perceived and imagined spoken words. *Cerebral Cortex* 2005;15(2):221–228. [PubMed: 15269107]
- Ashtari M, Lencz T, Zuffante P, Bilder R, Clarke T, Diamond A, et al. Left middle temporal gyrus activation during a phonemic discrimination task. *Neuroreport* 2004;15(3):389–393. [PubMed: 15094489]
- Baddeley AD. The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology* 1966;18:302–309. [PubMed: 5956072]
- Benson RR, Richardson M, Whalen DH, Lai S. Phonetic processing areas revealed by sinewave speech and acoustically similar non-speech. *Neuroimage* 2006;31:342–353. [PubMed: 16530428]
- Benson RR, Whalen DH, Richardson M, Swainson B, Clark VP, Lai S, et al. Parametrically dissociating speech and nonspeech perception in the brain using fMRI. *Brain and Language* 2001;78(3):364–396. [PubMed: 11703063]
- Bergerbest D, Ghahremani DG, Gabrieli JDE. Neural correlates of auditory repetition priming: Reduced fMRI activation in the auditory cortex. *Journal of Cognitive Neuroscience* 2004;16:966–977. [PubMed: 15298784]
- Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Springer JA, Kaufman JN, et al. Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex* 2000;10(5):512–528. [PubMed: 10847601]
- Binder JR, Rao SM, Hammeke TA, Yetkin FZ, Jesmanowicz A, Bandettini PA, et al. Functional magnetic-resonance-imaging of human auditory-cortex. *Annals of Neurology* 1994;35(6):662–672. [PubMed: 8210222]
- Crinion JT, Lambon-Ralph MA, Warburton EA, Howard D, Wise RJS. Temporal lobe regions engaged during normal speech comprehension. *Brain* 2003;126:1193–1201. [PubMed: 12690058]
- Dehaene-Lambertz G, Pallier C, Serniclaes W, Sprenger-Charolles L, Jobert A, Dehaene S. Neural correlates of switching from auditory to speech perception. *Neuroimage* 2005;24(1):21–33. [PubMed: 15588593]
- Dehaene-Lambertz G, Dehaene S, Anton JL, Campagne A, Ciuciu P, Dehaene GP, et al. Functional segregation of cortical language areas by sentence repetition. *Human Brain Mapping* 2006;27(5):360–371. [PubMed: 16565949]
- Démonet JF, Chollet F, Ramsay S, Cardebat D, Nespoulous JL, Wise R, Rascol A, Frackowiak R. The anatomy of phonological and semantic processing in normal subjects. *Brain* 1992;115:1753–1768.
- Dufour S, Peereman R. Lexical competition in phonological priming: Assessing the role of phonological match and mismatch lengths between primes and targets. *Memory & Cognition* 2003;31(8):1271–1283.
- Eimas PD, Corbit JD. Selective adaptation of linguistic feature detectors. *Cognitive Psychology* 1973;4:99–109.
- Formisano E, De Martino F, Bonte M, Goebel R. "Who" is saying "what"? Brain-based decoding of human voice and speech. *Science* 2008;322:970–969. [PubMed: 18988858]
- Friston KJ, Penny WD, Glaser DE. Conjunction revisited. *Neuroimage* 2005;25(3):661–667. [PubMed: 15808967]
- Gagnepain P, Chetelat G, Landeau B, Dayan J, Eustache F, Lebreton K. Spoken word memory traces within the human auditory cortex revealed by repetition priming and functional magnetic resonance imaging. *Journal of Neuroscience* 2008;28(20):5281–5289. [PubMed: 18480284]
- Giraud AL, Kell C, Thierfelder C, Sterzer P, Russ MO, Preibisch C, et al. Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. *Cerebral Cortex* 2004;14(3):247–255. [PubMed: 14754865]
- Graves WW, Grabowski TJ, Mehta S, Gupta P. The left posterior superior temporal gyrus participates specifically in accessing lexical phonology. *Journal of Cognitive Neuroscience* 2008;20(9):1698–1710. [PubMed: 18345989]
- Griffiths TD, Rees A, Green GGR. Disorders of human complex sound processing. *Neurocase* 1999;5(5):365–378.

- Grill-Spector K, Henson R, Martin A. Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences* 2006;10(1):14–23. [PubMed: 16321563]
- Hasson U, Skippe JI, Nusbaum HC, Small SL. Abstract coding of audiovisual speech: Beyond sensory representation. *Neuron* 2007;56(6):1116–1126. [PubMed: 18093531]
- Hasson U, Nusbaum HC, Small SL. Repetition suppression for spoken sentences and the effect of task demands. *Journal of Cognitive Neuroscience* 2006;18(12):2013–2029. [PubMed: 17129188]
- Hickok G, Poeppel D. The cortical organization of speech processing. *Nature Review Neuroscience* 2007;8:393–402.
- Horner AJ, Henson RN. Priming, response learning and repetition suppression. *Neuropsychologia* 2008;46(7):1979–1991. [PubMed: 18328508]
- Howard D, Patterson K, Wise R, Brown WD, Friston K, Weiller C, Frackowiak R. The cortical localization of the lexicons. *Brain* 1992;115:1769–1782. [PubMed: 1486460]
- Klein D, Zatorre RJ, Chen JK, Milner B, Crane J, Belin P, Bouffard M. Bilingual brain organization: A functional magnetic resonance adaptation study. *Neuroimage* 2006;31(1):366–375. [PubMed: 16460968]
- Kucera, Henry; Francis, W. Nelson Providence, Brown University Press; 1967. Computational analysis of present-day American English.
- Liebenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA. Neural substrates of phonemic perception. *Cerebral Cortex* 2005;15(10):1621–1631. [PubMed: 15703256]
- Macey PM, Macey KE, Kumar R, Harper RM. A method for removal of global effects from fMRI time series. *Neuroimage* 2003;22:360–366. [PubMed: 15110027]
- Mazoyer BM, Tzourio N, Frak V, Syrota A, Murayama N, Levrier O, et al. The cortical representation of speech. *Journal of Cognitive Neuroscience* 1993;5(4):467–479.
- Mellet E, Tzourio N, Crivello F, Joliot M, Denis M, Mazoyer B. Functional anatomy of spatial mental imagery generated from verbal instructions. *Journal of Neuroscience* 1996;16(20):6504–6512. [PubMed: 8815928]
- Meyer M, Zysset S, von Cramon DY, Alter K. Distinct fMRI responses to laughter, speech, and sounds along the human peri-sylvian cortex. *Cognitive Brain Research* 2005;24(2):291–306. [PubMed: 15993767]
- Mottonen R, Calvert GA, Jaaskelainen IP, Matthews PM, Thesen T, Tuomainen J, et al. Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *Neuroimage* 2006;30(2):563–569. [PubMed: 16275021]
- Narain C, Scott SK, Wise RJS, Rosen S, Leff A, Iversen SD, et al. Defining a left-lateralized response specific to intelligible speech using fMRI. *Cerebral Cortex* 2003;13(12):1362–1368. [PubMed: 14615301]
- Obleser J, Eisner F, Kotz SA. Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *Journal of neuroscience* 2008;28(32):8116–8124. [PubMed: 18685036]
- Okada K, Hickok G. Identification of lexical-phonological networks in the superior temporal sulcus using functional magnetic resonance imaging. *Neuroreport* 2006;17(12):1293–1296. [PubMed: 16951572]
- Pakhomov, S. MNI Space Utility. 2006. [Database.] Retrieved from Positron Emission Tomography Lab of the Institute of the Human Brain website, October 2008: http://www.ihb.spb.ru/~pet_lab/MSU/MSUMain.html
- Patterson RD, Johnsrude IS. Functional imaging of the auditory processing applied to speech sounds. *Philosophical Transactions of the Royal Society B-Biological Sciences* 2008;363(1493):1023–1035.
- Poeppel D, Guillemin A, Thompson J, Fritz J, Bavelier D, Braun AR. Auditory lexical decision, categorical perception, and FM direction discrimination differentially engage left and right auditory cortex. *Neuropsychologia* 2004;42(2):183–200. [PubMed: 14644105]
- Poeppel D. Pure word deafness and the bilateral processing of the speech code. *Cognitive Science* 2001;25(5):679–693.
- Poeppel D, Idsardi WJ, van Wassenhove V. Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B-Biological Sciences* 2008;363(1493):1071–1086.

- 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, Wise RJS, Warburton EA, Moore CJ, Howard D, Patterson K, Frackowiak RSJ, Friston KJ. Hearing and saying. The functional neuro-anatomy of auditory word processing. *Brain* 1996;119:919–931. [PubMed: 8673502]
- Price C, Thierry G, Griffiths T. Speech-specific auditory processing: where is it? *Trends in Cognitive Sciences* 2005;9:271–276. [PubMed: 15925805]
- Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience* 2009;12(6):718–724.
- Rimol, LM.; Specht, K.; Weis, S.; Savoy, R.; Hugdahl, K. Processing of sub-syllabic speech units in the posterior temporal lobe: An fMRI study; *Neuroimage*; 2005. p. 1059-1067.
- Rissman J, Eliassen JC, Blumstein SE. An event-related fMRI investigation of implicit semantic priming. *Journal of Cognitive Neuroscience* 2003;15(8):1160–1175. [PubMed: 14709234]
- Romaya, J. Cogent 2000, Version 1.25. [Software.] This experiment was realised using Cogent 2000 developed by the Cogent 2000 team at the FIL and the ICN and Cogent Graphics developed by John Romaya at the LON at the Wellcome Department of Imaging Neuroscience. 2003. Retrieved from the Laboratory of Neurobiology webpage, June 2007: <http://www.vislab.ucl.ac.uk/Cogent>
- Rorden, C. MRICRON, Version 1. 2008. [Software.] Retrieved from Chris Rorden's homepage, June 2008: <http://www.sph.sc.edu/cmd/rorden>
- Sabri M, Binder JR, Desai R, Medler DA, Leitel MD, Liebenthal E. Attentional and linguistic interactions in speech perception. *NeuroImage* 2008;39(3):1444–1456. [PubMed: 17996463]
- Samuel AG. Lexical activation produces potent phonemic percepts. *Cognitive psychology* 1997;32:97–127. [PubMed: 9095679]
- Sanders LD, Poeppel D. Local and global auditory processing: behavioral and ERP evidence. *Neuropsychologia* 2006;45:1172–1186. [PubMed: 17113115]
- Schiller NO, Horemans I, Ganushchak L, Koester D. Event-related brain potentials during the monitoring of speech errors. *NeuroImage* 2009;44:520–530. [PubMed: 18930155]
- Smith EE, Jonides J, Marshuetz C, Koeppel RA. Components of verbal working memory: Evidence from neuroimaging. *Proceedings of the National Academy of Sciences of the United States of America* 1998;95(3):876–882. [PubMed: 9448254]
- Sperling G. Phonemic model of short-term auditory memory. *Proceedings of the American Psychological Association* 1968;3:63–34.
- Sperling G, Speelman RG. The effect of sound stimuli on short-term memory. *Bell Telephone Laboratories Technical Memorandum*. 1967
- Stefanatos GA. Speech perceived through a damaged temporal window: lessons from word deafness and aphasia. *Seminars in speech and language* 2008;29(3):239–252. [PubMed: 18720320]
- Sumner M, Samuel AG. Lexical inhibition and sublexical facilitation are surprisingly long lasting. *Journal of Experimental Psychology-Learning Memory and Cognition* 2007;33(4):769–790.
- Thierry G, Giraud A, Price C. Hemispheric dissociation in access to the human semantic system. *Neuron* 2003;38:499–506. [PubMed: 12741995]
- Thorn ASC, Frankish CR. Long-term knowledge effects on serial recall of nonwords are not exclusively lexical. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2005;31(4):729–735.
- Uppenkamp S, Johnsrude IS, Norris D, Marslen-Wilson W, Patterson RD. Locating the initial stages of speech-sound processing in human temporal cortex. *Neuroimage* 2006;31(3):1284–1296. [PubMed: 16504540]
- Vaden, KI.; Hickok, GS.; Halpin, HR. Irvine Phonotactic Online Dictionary. 2005. [Data file]. Available from www.iphod.com.
- van Wassenhove V, Grant KW, Poeppel D. Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences of the United States of America* 2005;102(4):1181–1186. [PubMed: 15647358]

- Vitevitch MS, Luce PA. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language* 1999;40(3):374–408.
- Vitevitch MS. The influence of sublexical and lexical representations on the processing of spoken words in english. *Clinical Linguistics & Phonetics* 2003;17(6):487–499. [PubMed: 14564833]
- Vouloumanos A, Kiehl KA, Werker JF, Liddle PF. Detection of sounds in the auditory stream: event related fMRI evidence for differential activation to speech and nonspeech. *Journal of Cognitive Neuroscience* 2001;13(7):994–1005. [PubMed: 11595101]
- Whalen DH, Benson RR, Richardson M, Swainson B, Clark VP, Lai S, Mencl WE, Fulbright RK, Constable RT, Liberman AM. Differentiation of speech and nonspeech processing within primary auditory cortex. *Journal of the acoustical society of America* 2006;119:575–581. [PubMed: 16454311]
- Weide, Robert L. CMU Pronouncing Dictionary. 1994. [Data file]. Available from the Speech at CMU Web Page, <http://www.speech.cs.cmu.edu/cgi-bin/cmudict>
- Wilson M. MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. *Behavioural Research Methods, Instruments and Computers* 1988;20(1):6–11.
- Wilson SM, Iacoboni M. Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *Neuroimage* 2006;33(1):316–325. [PubMed: 16919478]
- Wise R, Chollet F, Hadar U, Friston K, Hoffner E, Frackowiak R. Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain* 1991;114:1803–1817. [PubMed: 1884179]
- Zatorre RJ, Evans AC, Meyer E, Gjedde A. Lateralization of phonetic and pitch discrimination in speech processing. *Science* 1992;256(5058):846–849. [PubMed: 1589767]
- Zatorre RJ, Gandour JT. Neural specializations for speech and pitch: moving beyond the dichotomies. *Philosophical Transactions of the Royal Society* 2008;363(1493):1087–1104.

12.6 Second Acquisition Block

10.5 Second Acquisition Block

8.4 Second Acquisition Block

Figure 1. Jittered block trial design

The jittered block design presented wordlists in a fixed duration with a varied rest period, or inter-trial interval. This is an efficient and statistically powerful alternative to block and event related designs.

Figure 2. Significant repetition-suppression trend in left and right STS

Left and right STS demonstrated decreases in BOLD that were significantly correlated with increased phonological repetition. All voxels shown in red passed the $p < 0.001$ ($t = 3.69$, $df = 16$) and cluster extent correction ($p < 0.05$, extent = 15 voxels) thresholds.

Table 1**Experiment design and number of observations**

Summary of the data collected for each condition, over the course of each run and each experiment session. During each run, subjects were presented with each of the three main conditions with equal frequency, while catch trials appeared only twice per run.

Condition	Trials per Run	Total Trials
Low phonological repetition <i>ex. jug, knit, rage, hem</i>	9 trials, 45 volumes	72 trials, 360 volumes
Medium phonological repetition <i>ex. cab, calf, cat, cap</i>	9 trials, 45 volumes	72 trials, 360 volumes
High phonological repetition <i>ex. hip, hip, hip, hip</i>	9 trials, 45 volumes	72 trials, 360 volumes
Catch Trials <i>ex. hig, sheeve, tomb, batch</i>	2 trials, 10 volumes	16 trials, 80 volumes
Total:	29 trials, 145 volumes	232 trials, 1160 volumes

Table 2
Summary of phonological repetition effects on speech perception activity
Summary of regions with significant effects of phonological repetition or enhancement, with significantly increased or decreased responses to wordlists with increased repetition.

Region Description	x	Peak MNI coords y	z	Cluster extent	Cluster p (cor)	Peak t
<i>Repetition-Suppression</i>						
Supplementary Motor Area, Cingulate Gyrus (L)	-3	3	54	88	< 0.001	6.95
IFG, Sub-Gyral (L)	-27	27	0	15	0.029	5.37
IFG (R)	36	12	36	28	0.001	5.2
Middle Frontal Gyrus (R)	45	30	27	32	< 0.001	5.14
Cingulate Gyrus (R)	6	24	36	18	0.011	5.02
Sub-Gyral, Precentral Gyrus (L)	-36	12	21	33	< 0.001	5.81
Superior Temporal Sulcus (L)	-63	-30	3	43	< 0.001	6.25
Superior Temporal Sulcus (R)	45	-33	-3	16	0.021	4.88
Sub-Gyral, Superior Parietal Lobule, Precuneus (L)	-27	-60	45	23	0.003	4.7
Medial Dorsal Nucleus (R)	6	-15	3	16	0.021	4.47
<i>Repetition-Enhancement</i>						
Superior Frontal Gyrus, Medial Frontal Gyrus, Anterior Cingulate (L)	-3	60	12	691	< 0.001	9.65
IFG (R)	51	30	0	17	0.015	6.84
Cingulate Gyrus, Precuneus (L)	-3	-51	30	428	< 0.001	8.52
Angular Gyrus, SMG (L)	-48	-60	33	378	< 0.001	8.92
SMG (L)	-60	-30	36	34	< 0.001	8.24
SMG (R)	63	-27	33	152	< 0.001	6.73
Superior Temporal Gyrus (R)	48	0	-6	16	0.021	5.75
Middle Temporal Gyrus (R)	57	-48	15	54	< 0.001	5.62
Middle Temporal Gyrus (R)	45	-63	12	30	< 0.001	4.45
Sub-Gyral, Insula, Claustrum (R)	42	-12	-9	93	< 0.001	6.53

Group t-statistic maps were thresholded at $p < 0.001$ ($t = 3.69$, $df = 16$) with a cluster extent corrected $p < 0.05$ (extent = 15). MNI coordinates are given for each peak, with probabilistically defined cluster labels given by the MNI Space Utility (Pakhomov, 2006). Whole brain volume = 40009 voxels. Cluster extent is given in normalized MNI space.

Abbreviations: IFG = inferior frontal gyrus, SMG = Supramarginal Gyrus; (R) or (L) denotes right or left, respectively.