◆ **Human Brain Mapping 34:3182–3192 (2013)** ◆

Right and Left Perisylvian Cortex and Left Inferior Frontal Cortex Mediate Sentence-Level Rhyme Detection in Spoken Language as Revealed by Sparse fMRI

Martina A. Hurschler,[1,2]* Franziskus Liem,[1,2] Lutz Jäncke,[1,2,3] and Martin Meyer[2,3,4]

[1] *Division of Neuropsychology, Institute of Psychology, University of Zurich, Zurich, Switzerland*
[2] *Institute of Psychology, Neuroplasticity and Learning in the Healthy Aging Brain (HAB LAB), University of Zurich, Zurich, Switzerland*
[3] *International Normal Aging and Plasticity Imaging Center, University of Zurich, Zurich, Switzerland*
[4] *Center for Integrative Human Physiology, University of Zurich, Zurich, Switzerland*

◆

Abstract: In this study, we used functional magnetic resonance imaging to investigate the neural basis of auditory rhyme processing at the sentence level in healthy adults. In an explicit rhyme detection task, participants were required to decide whether the ending syllable of a metrically spoken pseudosentence rhymed or not. Participants performing this task revealed bilateral activation in posterior–superior temporal gyri with a much more extended cluster of activation in the right hemisphere. These findings suggest that the right hemisphere primarily supports suprasegmental tasks, such as the segmentation of speech into syllables; thus, our findings are in line with the "asymmetric sampling in time" model suggested by Poeppel ([2003]: Speech Commun 41:245–255). The direct contrast between rhymed and nonrhymed trials revealed a stronger BOLD response for rhymed trials in the frontal operculum and the anterior insula of the left hemisphere. Our results suggest an involvement of these frontal regions not only in articulatory rehearsal processes, but especially in the detection of a matching syllable, as well as in the execution of rhyme judgment. *Hum Brain Mapp 34:3182–3192, 2013.* © 2012 Wiley Periodicals, Inc.

Key words: rhyme detection; functional lateralization; auditory fMRI; frontal operculum; anterior insula; perisylvian cortex; asymmetric sampling in time; phonological judgment

◆

Contract grant sponsor: Swiss National Science Foundation; Contract grant number: 320030-120661; Contract grant sponsor: Fonds zur Forderung des akademischen Nachwuchses (FAN); Contract grant sponsor: Zurcher Universitatsverein (ZUNIV).

*Correspondence to: Martina A. Hurschler, Institute of Psychology, Neuroplasticity and Learning in the Healthy Aging Brain (HAB LAB), University of Zurich, Sumatrastrasse 30, CH-8006 Zurich, Switzerland. E-mail: m.hurschler@psychologie.uzh.ch

Received for publication 20 August 2011; Revised 27 April 2012; Accepted 1 May 2012

DOI: 10.1002/hbm.22134

INTRODUCTION

The ability to detect rhyme is considered to be one of the earliest developing and most simple phonological awareness skills [Coch et al., 2011]. The sensitivity to spoken rhyme has previously been linked to the development of different language functions, such as, reading and spelling. Nevertheless, barely any neuroimaging studies about the neural correlates of auditory rhyme processing exist today.

Young children appear to appreciate rhyme [Bryant et al., 1989], and there is evidence that they are able to fulfill rhyme detection tasks as early as 3-year-old [Stanovich et al., 1984]. Hence, children seem to ascertain rhyme in spoken language before they have reached the ability to detect phonetic segments. This observation is consistent with the linguistic status hypothesis, which maintains that syllables have an advantage over intrasyllabic units and that intrasyllabic units, in turn, have an advantage over individual phonemes [Treiman, 1985].

Numerous behavioral longitudinal and crosscultural studies have been able to show that preschool experiences with auditory rhyme detection have a significant effect on later success in learning to read and write [Bryant et al., 1989]. Both sensitivity to spoken rhyme and measures for memory span are related to vocabulary development in preschoolers [Avons et al., 1998].

With respect to the neural correlates of auditory rhyme processing, evidence is currently sparse. Speech perception relies on mechanisms of time-resolution at a time scale level of milliseconds. The predominance of the left perisylvian region for most domains within speech processing is an evidenced fact in neuroscientific research [e.g., Friederici, 2011; Narain et al., 2003; Price, 2000; Vigneau et al., 2006]. Following the traditional model of language, the majority of colleagues, who do research in aphasia, emphasize the superior and cardinal role of the left hemisphere. Clinical literature has often reported sensory aphasic problems resulting from left temporal lobe lesions [e.g., Kuest and Karbe, 2002; Stefanatos, 2008; Turner et al., 1996]. This left perisylvian region is the site for both elemental functions, such as, phonetic processing, and higher purposes, namely, syntactic and semantic detection. However, gradually mounting evidence obtained from neuroimaging studies in non brain-damaged individuals proposes that the contribution of the right hemisphere to the processing of speech perception must not be underestimated [Jung-Beeman, 2005; Meyer, 2008; Poeppel and Hickok, 2004; Shalom and Poeppel, 2008; Stowe et al., 2005; Vigneau et al., 2011].

In the current study we investigate the neural signatures of auditory rhyme processing at the sentence level because we believe that learning more about this issue will contribute to the topic of functional lateralization in speech processing. This assumption is based on the very nature of different processes that are involved in the performance of an auditory rhyme detection task, such as, the automatic registration of phonological input, the processing of phonemic segmentation, the retention of information in the articulatory loop, the comparison of critical word-ending sounds, and both decision making and response provision [Baddeley et al., 1984]. As regards the suprasegmental processes, which form the basis of rhyme detection, one might predict a right-lateralized activation in the posterior–superior temporal gyrus (pSTG) as suggested by the "asymmetric sampling in time" (AST) hypothesis proposed by Poeppel [2003]. According to this framework, auditory information is preferentially integrated in differential temporal windows by the nonprimary auditory fields residing in the two hemispheres. While the left hemisphere is suggested to be specialized for the perception of rapidly changing acoustic cues (~40 Hz), this model predicts a better adaption of the right auditory cortex for slowly changing acoustic modulations (~4 Hz).

In support of the "AST"-hypothesis, different studies were able to demonstrate that the right supratemporal plane is especially amenable to slow acoustic modulations in speech [e.g., Hesling et al., 2005; Ischebeck et al., 2008; Plante et al., 2002; Zhang et al., 2010]. In particular, activation in the posterior supratemporal region of the right hemisphere was associated with speech melody processing [Gandour et al., 2004; Meyer et al., 2002, 2004] and explicit processing of speech rhythm [Geiser et al., 2008].

According to Poeppel [2003], the AST model permits different predictions regarding the lateralization of different speech perception tasks. One such prediction states that "phonetic phenomena occurring at the level of syllables should be more strongly driven by right hemisphere mechanisms" [Poeppel, 2003, p 251]. The problem with investigating this assumption is that syllables always contain their phonemic constituents [Poeppel, 2003]. Therefore, an insightful experiment should disentangle selective processing of syllables from the more general processing of their constituent phonemes. This reasoning has found some support by a dichotic listening study that showed increased rightward lateralization when the focus of the task emphasized syllabicity instead of the phonemic structure of the stimuli [Meinschaefer et al., 1999].

We believe, that akin to speech meter, rhymes serve as structural devices. Geiser et al. [2008] have previously investigated the neural correlates of explicit rhythm processing in spoken sentences by using German pseudosentences spoken in either an isochronous, or a conversational rhythm. In the explicit task, subjects had to judge, whether the heard pseudosentence was "isochronous" or "nonisochronous" (rhythm task) that is whether the sentence had a metrical structure or not. In the implicit condition, unattended rhythm processing was measured, while participants had to decide, whether the sentence they heard was a question or a statement (prosody task). One particular result that they provided is increased rightward lateralization in temporal and frontal regions associated with explicit processing of speech rhythm. Interestingly, they did not find this right lateralized temporal activation in the implicit stimulus-driven processing condition. The observed difference in activation between implicit and explicit condition is in line with previous auditory functional imaging studies that were able to demonstrate task-dependent modulation of auditory cortical areas involved in speech processing [Noesselt et al., 2003; Poeppel et al., 1996; Scheich et al., 2007; Tervaniemi and Hugdahl, 2003]. The task used in our study resembles the explicit task used in the study by Geiser et al. [2008] insofar as the focus of subjects' attention is explicitly set to suprasegmental analysis. Based on the aforementioned findings, we hypothesize that an explicit rhyme detection task at the sentence level should be associated with increased involvement of the right perisylvian cortex.

With respect to the direct comparison between rhymed and nonrhymed stimuli we have to consider cognitive demands that may be involved. To accurately perform a rhyme detection task, the phonetic information should not only be segmented into syllables; indeed, it should also be memorized until the critical phoneme is encountered. The distance between the two relevant phonemes involves working memory (WM), as one item must be kept active until it can be compared with a second phonetic element. According to Baddeley's influential model, verbal memory is thought to be divided by a subvocal rehearsal system and a phonological store. While the phonological store is suggested to hold auditory/verbal information for a very short period of time, articulatory rehearsal is a more active process that retains the information in the phonological store [Baldo and Dronkers, 2006]. It has been previously argued that rhyme judgments engage both of these processes [Baddeley et al., 1984]. Several PET and functional magnetic resonance imaging (fMRI) studies that used 2-back or 3-back tasks to investigate WM found activation in the left IFG [mostly in the opercular part, corresponding FOP; see Rogalsky and Hickok, 2011; Tzourio-Mazoyer et al., 2002], which was related to articulatory rehearsal. In addition, it has been proposed that the left IPL subserves the phonological store [e.g., Paulesu et al., 1993].

Contrary to most of the previous studies about rhyme processing, we used pseudosentences instead of real word stimuli. Therefore, we are able to rule out possible confounds brought about by obvious semantic processing. To control for WM load, the pseudosentences were spoken metrically. This enables the span between the end rhymes to remain constant. To direct the participants' attention to the phonology stimuli's last syllable, all of the pseudosentences were spoken in the same isochronous rhythm.

As previously mentioned, explicit rhyme detection at the sentence level has not yet been investigated with fMRI methodology. Based on the predictions of the AST-hypothesis, as well as findings from the aforementioned studies pertaining to prosody and speech meter, we predict that the rhyme detection task per se should be related to enhanced supratemporal recruitment of the right auditory-related cortex. Because of the cognitive demands of the task used, we also expect the recruitment of areas related to the phonological loop of the WM, such as, the left inferior parietal lobe and the (left) frontal operculum.

Since our approach investigates hemispheric lateralization in processing acoustic suprasyllabic spoken language, we further explore the division of labor between the right and the left auditory-related cortex. The goal of this study is to investigate neural signatures of auditory rhyme detection at the sentence level. This should not only enhance the understanding of the neural processes underlying the detection of rhyme in rhymed (metrical) sentences, but also the relationship between slowly changing acoustic modulations and right auditory-related cortex functions in general.

METHODS

Subjects

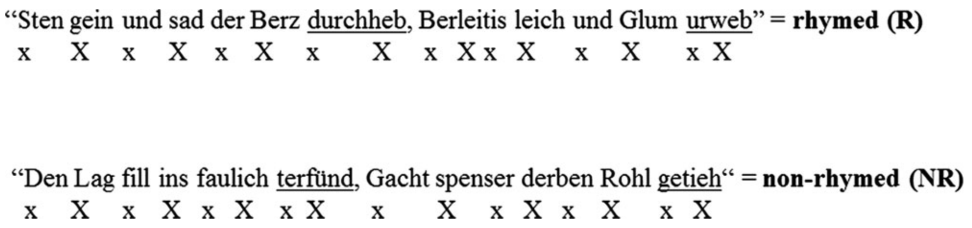
A total of 22 healthy subjects (11 females) aged 19–31 years (mean = 23.5, SD = 3.6) participated in this study. According to the Annett-Handedness-Questionnaire (AHQ) [Annett, 1970], all subjects were consistently right-handed. They were native speakers of (Swiss) German with no history of neurological, major medical, psychiatric, or hearing disorders. All subjects gave written consent in accordance with procedures approved by the local Ethics Committee. Subjects were paid for their participation.

Stimuli

Stimuli material comprised a total of 72 pseudosentences containing phonotactically legal pseudowords. Our stimuli resemble so-called "jabberwocky" sentences used in prior studies [e.g., Friederici et al., 2000; Hahne and Jescheniak, 2001], in that, they contain some real German function words. In contrast with typical jabberwocky sentences, they display a regular meter and do not contain systematic morphological markers, to minimize semantic and syntactic associations. Rhymed and nonrhymed sentences were matched based on the amount of function words they contained.

The last syllable of the stimuli either rhymed (*R*) or did not rhyme (NR) with the last syllable of the first part of the sentence (see Fig. 1). The pseudosentences were metrically spoken by a trained female speaker and consisted of a verse form, which means that sentences followed a regular meter (eight iambs per sentence). As a result, each pseudosentence contained 16 syllables and the sentences consisted of a mean of 10.4 pseudowords (SD ± 1.4).

All stimulus items were normalized in amplitude to 70% of the loudest signal in a stimulus item. All pseudosentences were analyzed by the means of PRAAT speech editor [Boersma, 2001]. Stimuli were balanced with respect to mean intensity, and the length of all stimuli was set to exactly 6 s.



"Sten gein und sad der Berz durchheb, Berleitis leich und Glum urweb" = **rhymed (R)**

x X x X x X x X x Xx X x X x X

"Den Lag fill ins faulich terfünd, Gacht spenser derben Rohl getieh" = **non-rhymed (NR)**

x X x X x X x X x X x X x X x X

Figure 1.

Examples of pseudosentences. Underlined are the pseudowords, which had to be compared.

Task/Procedure

Each participant read instructions to the experiment, gave their written consent, and completed the Annett-Handedness-Questionnaire. During scanning, the room lights were dimmed and a fixation cross was projected, via a forward projection system, onto a translucent screen placed at the supine position at the end of the magnet's gurney. Subjects viewed the screen through a mirror attached to the head coil. Stimuli were presented using Presentation® software (Version 0.70, www.neurobs.com). The stimulus presentation was synchronized with the data acquisition by employing a 5 V TTL trigger pulse. We used an MR-compatible piezoelectric auditory stimulation system that is incorporated into standard Philips headphones for binaural stimulus delivery.

Subjects were instructed to decide as quickly and as accurately as possible whether the pseudosentences that they were presented with rhymed or not. They indicated their response by pressing a button on the response box with either their right index finger, or with their right middle finger. Additionally, a total of 10 null events were created to be a baseline condition and were randomly included in the time course of the experiment. During the empty trials, subjects were instructed to press a random button. In one run, a total of 82 trials (36 rhymed pseudosentences, 36 nonrhymed pseudosentences, and 10 empty trials) were presented. A fixation cross was presented for 500 ms prior to each stimulus presentation. The task in the scanner lasted 20 min 30 s.

Data Acquisition

The functional imaging study was performed on a Philips 3T Achieva whole-body MR unit (Philips Medical System, Best, The Netherlands) equipped with an eight-channeled Philips SENSE head coil. To acquire data, a clustered sparse temporal acquisition technique was used. This scheme combines the principles of a sparse temporal acquisition with a clustered acquisition [Liem et al., 2012; Schmidt et al., 2008; Zaehle et al., 2007]. That way, the stimuli were binaurally presented in an interval devoid of auditory scanner noise. Three consecutive volumes were collected, to cover the peak of the event-related hemodynamic signal (see Fig. 2).

Functional time series were collected from 16 transverse slices covering the entire perisylvian cortex with a spatial resolution of $2.7 \times 2.7 \times 4$ mm^3 by using a Sensitivity Encoded (SENSE) [Pruessmann et al., 1999], single-shot, gradient-echo planar sequence (acquisition matrix 80×80 voxels, SENSE accelerator factor $R = 2$, FOV = 220 mm, TE = 35 ms). The volumes were acquired with an acquisition time of 1,000 ms each, a flip angle = 68°, and a 12 s intercluster interval was employed; as a result, one trial lasted 15 s. Furthermore, a standard 3D T1-weighted volume for anatomical reference was collected with a gradient echo sequence with a $0.94 \times 0.94 \times 1$ mm^3 spatial resolution (160 axial slices, acquisition matrix 256×256 voxels, FOV = 240×240 mm, repetition time [TR] = 8.17 ms, flip angle = 8°).

Data Analysis

Behavioral data analysis and ROI statistics were performed by using SPSS Statistics 19.0 (SPSS Inc.).

Behavioral data

During the experiment in the scanner, behavioral performance data on the rhyme detection task were collected. Data (reaction time and accuracy) were corrected for outliers (>2 SD above or below mean value). A repeated-measures t-test was performed to identify significant differences between the conditions.

fMRI analysis

Artifact elimination and image analysis was performed by using MATLAB 7.4 (Mathworks, Natick, MA) and the

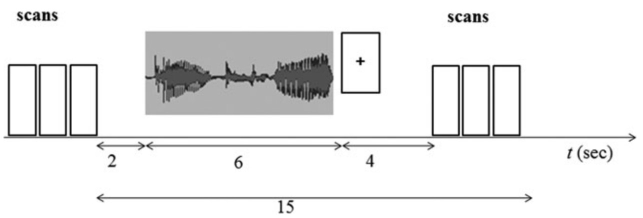


Figure 2.

Acquisition scheme. Depicted are the three time points of acquisition and the stimulus presentation in one trial.

TABLE I. Brain areas showing significant increases for rhymed and nonrhymed condition relative to baseline

	Left hemisphere					Right hemisphere				
Condition/region	*T* score	Voxels	*x*	*y*	*z*	*T* score	Voxels	*x*	*y*	*z*
Rhyme > rest										
Superior temporal gyrus						13.34	322	44	−14	2
	11.94	102	−50	−22	12					
	8.29	47	−48	−2	6					
Total amount of voxels		149					322			
Nonrhyme > rest										
Superior temporal gyrus						12.82	349	62	−16	−2
	12.08	104	−50	−22	12					
	8.52	22	−48	−2	6					
	7.89	40	−52	−12	−4					
Total amount of voxels		166					349			

Note: x, y, z = MNI coordinates of local maxima. Voxels = number of voxels at $P < 0.05$ after family-wise correction for multiple comparisons across the whole brain.

SPM5 software package (Institute of Neurology, London, UK; http://www.fil.ion.ucl.ac.uk). To account for movement artifacts, all volumes were realigned to the first volume, normalized into standard stereotactic space (voxel size $2 \times 2 \times 2$ mm^3, template provided by the Montreal Neurological Institute), and smoothed using a Gaussian kernel with a 6-mm full-width-at-half-maximum that increased the signal-to-noise ratio of the images. Due to the low number of sampling points, a boxcar function (first order, window length = 3 s) was modeled for each trial. In addition, two regressors of no interest were included, to account for the T1-decay along the three volumes [Liem et al., 2012; Zaehle et al., 2007]. The resulting contrast images from each of the first level fixed-effects analysis were entered into one-sample t-tests (df = 21); thereby, permitting inferences about condition effects across subjects [Friston et al., 1999]. Unless otherwise indicated, regions reported showed significant effects of $P < 0.05$ and were FWE corrected.

Post-hoc region of interest analyses

To statistically test for asymmetry in cluster size of temporal activation, cluster sizes in the right and the left STG at the single-subject level ($P < 0.001$, unc.) were extracted via an in-house-tool and subjected to a 2×2 repeated-measures ANOVA with the factors *condition* and *hemisphere*, followed by paired t-tests with the cluster extent in the right and the left STG for both conditions.

RESULTS

Behavioral Data

Individual mean reaction times (RT), as well as accuracy scores were distributed normally in both the *R* and the NR conditions (Kolmogorov-Smirnov one-sample test: $d = 0.153$, $P > 0.20$, and $d = 0.162$, $P > 0.20$) and were compared using a parametric two-sample t-test. Concerning RT no significant difference between *R* and NR conditions was revealed (mean ± SD = 635.1 ± 190.66 and 598.9 ± 167.015, respectively, $t = 1.214$, df = 21). On the contrary, accuracy was significantly lower in the *R* condition, as compared with the NR condition (92.4 ± 2.6% and 97.8 ± 1.25%, respectively; $t = 5.232$, $P < 0.001$, df = 21).

Imaging Data

Whole-head analysis

Rhyme detection task. In a first step of analysis, main effects for the rhyme detection task were investigated. Therefore, rhymed (*R*) and nonrhymed (NR) conditions were separately contrasted to the baseline (fixation cross and random button press). Table I and Figure 3 present regions that reveal significant supra-threshold BOLD-activation for each of the two experimental conditions, as compared with the empty trials. In both conditions a bilateral superior temporal fMRI pattern could be observed and exhibited a more expanded cluster of significant activation ($P < 0.05$, FWE corrected) in the right, as compared with the left hemisphere. Notably, the peak activation in the right auditory-related cortex of the posterior temporal lobe was more anteriorially and medially situated in the *R* (44-14-12), than in the NR condition (62-16-2).

To statistically test for this rightward temporal lateralization in cluster size for both contrasts (*R* > rest, NR > rest) for each subject's statistic map (first-level contrast), left and right cluster sizes within the superior temporal gyrus were extracted and subjected to a paired sample t-test. As depicted in Figure 4, temporal cluster size was significantly larger in the right, than the left hemisphere in the *R* condition ($t = 6.513$, $P < 0.001$, df = 21). This was also the case for the NR condition ($t = 5.029$, $P < 0.001$, df = 21).

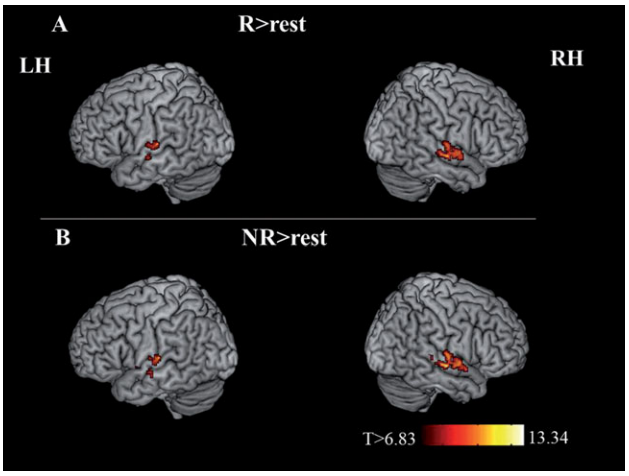


Figure 3.
Brain areas showing significantly greater activation during the processing of (A) rhymed and (B) nonrhymed condition compared with rest. Each cluster is thresholded at $P < 0.05$, FWE corrected with a spatial extent minimum of 20 contiguous voxels per cluster. The corresponding cortical regions, cluster sizes, peak *T*-values and MNI coordinates can be found in Table I.

Rhymed vs. nonrhymed pseudosentences. The direct contrast between both conditions (Table II, Fig. 5) revealed increased BOLD-responses in the anterior insula and the deep opercular portion of the inferior frontal gyrus of the left hemisphere for rhymed, as compared with the nonrhymed pseudosentences ($P < 0.05$ FWE corrected at cluster level, $k > 25$). Since the expected effects in the direct contrasts are smaller than in the contrasts versus rest, we adopted the more liberal approach of clusterwise FWE correction, to not miss effects. The reverse contrast at the same threshold did not reveal any significantly different activation patterns between the NR and the *R* condition.

DISCUSSION

In the current study, we investigated the neural basis of rhyme detection in healthy adults with a particular focus on lateralized processing. At the behavioral level, we did not find a significant difference in reaction times between rhymed and nonrhymed conditions. This finding is consistent with studies using visually presented rhyming words [Khateb et al., 2000, 2007; Rayman and Zaidel, 1991; Rugg and Barrett, 1987]. The significantly increased error rate for rhymed as compared with nonrhymed sentences, was also evident in previous studies [Rayman and Zaidel, 1991; Rugg, 1984; Rugg and Barrett, 1987]. We assume that subjects showed a bias towards negative responses, when they were not completely sure of the answer. This may be due to the speed-demands placed upon them (caused by the instruction to “respond as quickly and accurately as possible”) [Khateb et al., 2007].

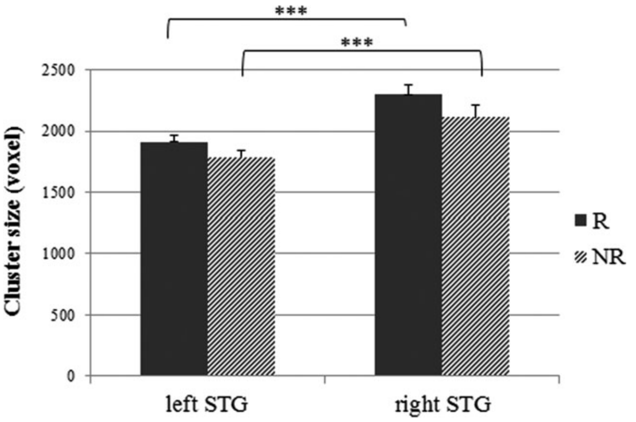


Figure 4.
Size of activated clusters in bilateral superior temporal gyrus (STG). Mean value of each subjects’ ($n = 22$) cluster extent in *R* > rest and NR > rest contrasts (***$P < 0.001$).

The assumption that cortical fields in the right temporal lobe along the superior temporal gyrus and sulcus play an essential role in the analysis of the speech signal continues to receive ever-increasing support [Boemio et al., 2005; Hickok, 2001; Lattner et al., 2005; Meyer et al., 2002, 2004; Vigneau et al., 2011]. The right lateralized activation was observed while subjects were performing a rhyme detection task at the sentence level. This result buttresses the results of previous studies, which have investigated the auditory processing of slowly changing cues, namely, prosody and speech meter [Geiser et al., 2008; Meyer et al., 2002; Zhang et al., 2010]. According to the AST hypothesis, the auditory-related cortex of the right hemisphere is more inclined to process slowly changing acoustic cues [Meyer, 2008; Poeppel, 2003; Zatorre and Gandour, 2008]. We posit that the right lateralized

TABLE II. Brain areas showing significant increases for rhymed compared with nonrhymed trials

Condition/region	H	*T* score	Voxels	*x*	*y*	*z*
Rhyme > nonrhyme						
Inferior frontal gyrus, opercular part	L	7.49	40	–52	14	0
Anterior insula	L	6.05	40	–28	24	6

Note: *x*, *y*, *z* = MNI coordinates of local maxima. H = hemisphere, L = left, voxels = number of voxels. *T* scores and cluster size are reported if they are significant at $P < 0.05$ after family-wise correction for multiple comparisons at cluster level ($k > 25$).

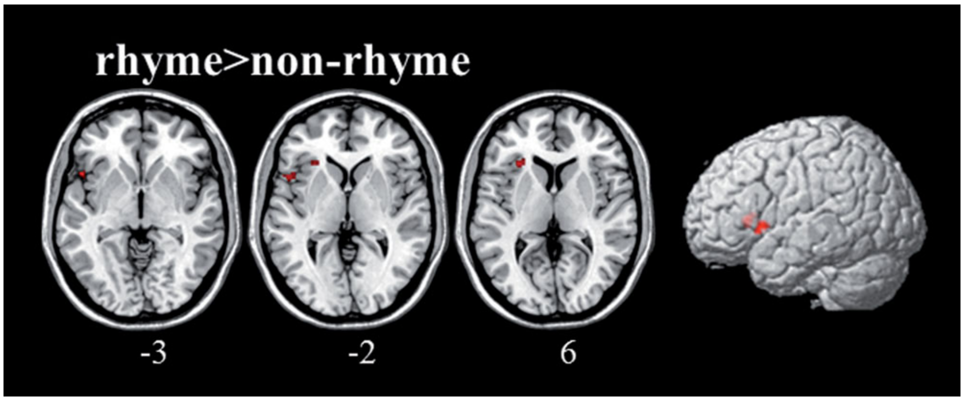


Figure 5.

Brain areas showing significantly greater activation during the processing of rhymed compared with nonrhymed pseudosentences. Each cluster is thresholded at $P < 0.05$, FWE-corrected at cluster level ($k > 25$). The corresponding cortical regions, cluster sizes, peak *T*-values and MNI coordinates can be found in Table II. Figures are displayed in neurological convention.

activation elicited during the explicit rhyme detection task complies with the predictions of this AST framework.

Akin to prosody and especially speech meter, rhymes serve as structural devices. Indeed, the segmentation of spoken sentences into single syllables is a suprasegmental computation, which relies on the analysis within larger time windows (~250 ms). The fact that we found this lateralized activation in cluster-size irrespectively of the condition and task performance provides support to the hypothesis of a task-dependent, top-down modulation of lateralization effects in parts of the auditory-related cortex that may be preferentially sensitive to suprasegmental acoustic aspects speech and music [Brechmann and Scheich, 2005; Tervaniemi and Hugdahl, 2003]. Geiser et al. [2008] found a similar right lateralization for speech rhythm perception only in an explicit, task-driven processing condition, which implies that areas of the right (and left) STG are partly modulated by task demand [Poeppel et al., 1996].

The direct contrast between rhymed and nonrhymed trials demonstrated increased BOLD response in the left hemisphere for rhymed pseudosentences in the opercular part of the IFG and the anterior insula. The finding of increased rhyme related fronto-opercular activation is of specific interest, since rhyming targets should have been phonologically primed and would therefore require less processing than nonrhyming targets [Coch et al., 2008]. However, a closer look at the literature pertaining to priming in auditory modality reveals a wide diversity of results. The best candidates for comparison to this study are experiments that used sequentially presented primes and targets in the auditory modality. The most consistent findings in such studies are reduced activation for related targets in the bilateral IFG, as well as in the bilateral superior temporal gyrus [Orfanidou et al., 2006; Vaden et al., 2010]. Notably, studies that did report priming effects in the IFG [Bergerbest et al., 2004; Orfanidou et al., 2006; Thiel et al., 2005] did not require explicit judgments between the prime and target word, as was the case in this study.

To our knowledge, this is the first fMRI study that directly compares rhymed to nonrhymed pseudosentences. A small number of fMRI studies implementing an explicit rhyme detection task compared BOLD response associated with a rhyme detection task to other tasks. But the stimuli employed in these studies were visually presented (therefore involving grapho-phonemic conversion) and included words and/or pseudo words [e.g., Cousin et al., 2007], or single syllables [Sweet et al., 2008]; thus, they obviously did not include direct contrasts between rhymed and nonrhymed sentences.

Therefore, we cannot rely upon these studies when attempting to elucidate the differences involved in auditory processing of rhymed versus nonrhymed items at the sentence-level.

Incidentally, various EEG investigations of the auditory modality have produced an electrophysiological rhyming effect for spoken word pairs. This effect is usually observed when a pair of words is presented and subjects are requested to make a phonemically based judgment and, it is typically expressed by a more negative bilateral posterior response for nonrhyming than for rhyming targets [Rugg, 1984]. Elsewhere, various researchers have demonstrated a reversal of this effect at lateral sites, that is, rhyming targets produced more negative responses than nonrhyming targets [Coch et al., 2005; Khateb et al., 2007]. In such an ERP study that included a rhyme-detection task with words, Coch et al. [2005] found a rhyming effect with a frontal leftward asymmetry in children and adults. They used a simple prime-target auditory rhyming paradigm with nonword stimuli (e.g., nin-rin and ked-voo). Interestingly, they found a more negative response to nonrhyming targets over posterior sites and an increased negativity to rhyming targets at lateral anterior sites. Subsequently, a visual rhyme-detection study conducted by Katheb et al. [2007] reported a specific left lateralized negativity for rhymed versus nonrhymed targets. Their estimated source localization indicated the major difference between rhyming and nonrhyming words as being positioned in predominantly left frontal and temporal areas. The fact that the rhyming effect can also be found when target words are spoken in a different voice than primes suggests that this effect is an index of phonological processing instead of a physical-acoustic mismatch [Praamstra and Stegeman, 1993]. However, due to the inverse problem and the limited spatial resolution of the EEG technique, the informative value of EEG studies for the present work is quite limited and comparisons must be interpreted with caution.

In our study, we found a significant signal increase in the left frontal operculum and the left anterior insula during the rhymed trials as compared with the nonrhymed trials; this finding was absent during the reverse contrast (NR > *R*). The left inferior frontal gyrus (LIFG) has been shown to be related to a myriad of functions in speech processing [e.g., Davis et al., 2008; Lindenberg et al., 2007; Meyer and Jancke, 2006]. Activation in the LIFG has been previously associated with segmentation processes or sublexical distinctions in different speech perception tasks [see Poeppel and Hickok, 2004] and a variety of syntactic and semantic operations [Hagoort, 2005; Shalom and Poeppel, 2008]. Nevertheless, there is currently no consensus with regards to the contribution that the LIFG makes to language processing [Friederici, 2011; Hickok, 2009]. Besides unspecific, modality independent involvement in different language tasks, this region has been suggested to reflect aspects of articulatory rehearsal [Meyer et al., 2004], discrimination of subtle temporal acoustic cues during speech and nonspeech [Zaehle et al., 2008], as well as auditory search [Giraud et al., 2004].

Previous studies were able to show that subvocal rehearsal processes are essentially mediated by parts of the LIFG [Paulesu et al., 1993]. The posterior–dorsal aspect of the LIFG (corresponding to the opercular part) might be preferentially engaged in phonology-related, sublexical processes [Burton et al., 2000; Zurowski et al., 2002]. This region is commonly suggested to be one part of the phonological loop in the Baddeley model [Paulesu et al., 1993; Smith and Jonides, 1999], and there is evidence that it mediates phonological rehearsal. Hemodynamic changes in the opercular frontal inferior region have been previously associated with making phonological judgments [Demonet et al., 1992; Poldrack et al., 1999; Zatorre et al., 1992].

Since this study used pseudosentences, subjects could not build up expectations about the following words. Instead, they were required to maintain the critical segment from the first part of the sentence in their mind for 3 s until they heard the second critical segment, after which they made their decision by pressing a button box. Thus, it is clear that phonological rehearsal is needed, to detect rhyme; therefore, the involvement of inferior frontal regions is not surprising. The subjects in this study did not know whether the sentence that they were listening to rhymed or not until they heard the last syllable. Therefore, this result cannot be explained by WM load per se; instead it is linked to the different outcomes resulting from the comparison between the syllables.

As suggested by Rogalsky and Hickok [2011], parts of the frontal operculum corresponding to regions in which we noted differences are essential for the integration of information that is maintained via articulatory rehearsal processes or decision-level processes, or both. The fact that we found activation in this region when we made a direct comparison between the rhymed versus the nonrhymed sentences bolsters the notion that the opercular portion of the LIFG plays a role in various decision-processes involved in a task that relies on phonological WM. This interpretation also fits with results of previous studies, which found that the LIFG is involved in a adverse listening condition with enhanced demands on response selection [Binder et al., 2004; Giraud et al., 2004; Vaden et al., 2010; Zekveld et al., 2006].

The direct comparison of rhymed with nonrhymed trials also revealed increased BOLD response in the left anterior insula. This region has previously been associated with diverse functions [Mutschler et al., 2009]. Sharing extensive connections with different structures in temporal, frontal, and parietal cortices, the insula is perfectly situated for the task of integrating different sensory modalities. Previous research has identified the anterior insula as a key player in general processes of cognitive control [Cole and Schneider, 2007; Dosenbach et al., 2007]. The anterior insula also seems to play a role in perception at each of the sensory modalities [Sterzer and Kleinschmidt, 2010]. Besides its involvement in subvocal rehearsal processes during WM activation, the left insula supports coordination processes in the complex articulatory programs that are needed during pseudoword processing [Ackermann and Riecker, 2004; Dronkers et al., 2004]. Dyslectic children show less activation than typically developing children in bilateral insulae during an auditory rhyme-detection task with words and pseudowords [Steinbrink et al., 2009]. Furthermore, there is evidence that the left anterior insula is also involved in the phonological recognition of words [Bamiou et al., 2003]. Thus, our findings provide further evidence that the insula is involved in the auditory-motor network [Mutschler et al., 2009]. However, our experimental design does not permit further discussion pertaining to the left anterior insula activation that we found.

The finding of significant differences in left frontal brain regions, which are associated with rhyme perception, coincides with results from the EEG studies discussed above. To reiterate, the aforementioned EEG studies produced significant differences for the direct contrasts between rhymed and nonrhymed stimuli. Due to the limited temporal resolution of fMRI technique, it is not possible to clearly link activation to a particular step of processing during the rhyme judgments. The stimuli used in both conditions did not contain syntactic or semantic information, and they did not differ in terms of intelligibility. Therefore, our finding that the reported left frontal brain activations were significant for the direct contrast level of analysis between rhymed and nonrhymed pseudosentences implies that these regions may not only be involved in articulatory rehearsal processes, but are also enmeshed in the last step of the analysis, namely, the detection of phonological matching.

Even though WM load was theoretically identical in both conditions, we nevertheless, must consider that task difficulty may have contributed to the difference in brain activation between the conditions. It has previously been shown that activation of the LIFG can be modulated by task-difficulty [Zekveld et al., 2006]. Since this is the first fMRI study that investigates auditory rhyme detection in an explicit paradigm at the sentence level, follow-up studies with more conditions that pose different cognitive demands should be introduced. Future research of this sort will prove helpful in disentangling brain responses that are associated with specific processes involved in auditory rhyme recognition.

CONCLUSION

We composed a rhyme detection task with pseudosentences to investigate the neural correlates of rhyme perception in healthy adults. Subjects in this study were requested to decide whether the last syllable of the pseudosentences rhymed or not. We found a task-related right-lateralized pattern of activation in the superior temporal lobe. This result implies that explicit rhyme processing at the sentence level—like prosody or meter in speech [Geiser et al., 2008; Meyer et al., 2002]—essentially relies on the processing in longer time windows wherefore the right temporal cortex has been proposed to be specialized [Poeppel, 2003]. Direct comparisons between rhymed and nonrhymed pseudosentences showed increased activation for the correctly recognized rhymed trials in left fronto-opercular areas (deep frontal operculum and adjoining anterior insula). These regions have been previously linked to processes of phonological WM and articulatory rehearsal.

ACKNOWLEDGMENTS

We express our gratitude to Sarah McCourt Meyer and two anonymous reviewers for helpful comments on an earlier draft.

REFERENCES

Ackermann H, Riecker A (2004): The contribution of the insula to motor aspects of speech production: A review and a hypothesis. Brain Lang 89:320–328.

Annett M (1970): A classification of hand preference by association analysis. Br J Psychol 61:303–321.

Avons S, Wragg C, Cupples L, Lovegrove W (1998): Measures of phonological short-term memory and their relationship to vocabulary development. Appl Psycholing 19:583–601.

Baddeley A, Lewis V, Valler G (1984): Exploring the articulatory loop. Q J Exp Psychol 36A:467–478.

Baldo J, Dronkers N (2006): The role of inferior parietal and inferior frontal cortex in working memory. Neuropsychology 20:529–538.

Bamiou D, Musiek F, Luxon L (2003): The insula (Island of Reil) and its role in auditory processing. Literature review. Brain Res Rev 42:143–154.

Bergerbest D, Ghahremani D, Gabrieli J (2004): Neural correlates of auditory repetition priming: Reduced fMRI activation in the auditory cortex. J Cogn Neurosci 16:966–977.

Binder J, Liebenthal E, Possing E, Medler D, Ward B (2004): Neural correlates of sensory and decision processes in auditory object identification. Nat Neurosci 7:295–301.

Boemio A, Fromm S, Braun A, Poeppel D (2005): Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat Neurosci 8:389–395.

Boersma P (2001): Praat, a system for doing phonetics by computer. Glot Int 5:341–345.

Brechmann A, Scheich H (2005): Hemispheric shifts of sound representation in auditory cortex with conceptual listening. Cereb Cortex 15:578–587.

Brett M, Anton J, Velabregue R, Polin J (2002): Region of interest analysis using an SPM tollbox. NeuroImage 16:1140–1141.

Bryant P, Bradley L, Maclean M, Crossland J (1989): Nursery rhymes, phonological skills and reading. J Child Lang 16:407–428.

Burton M, Small S, Blumstein S (2000): The role of segmentation in phonological processing: An fMRI investigation. J Cogn Neurosci 12:679–690.

Coch D, Grossi G, Skendzel W, Neville H (2005): ERP nonword rhyming effects in children and adults. J Cogn Neurosci 17:168–182.

Coch D, Hart T, Mitra P (2008): Three kinds of rhymes: An ERP study. Brain Lang 104:230–243.

Coch D, Mitra P, George E, Berger N (2011): Letters rhyme: Electrophysiological evidence from children and adults. Dev Neuropsychol 36:302–318.

Cole M, Schneider W (2007): The cognitive control network: Integrated cortical regions with dissociable functions. NeuroImage 37:343–360.

Cousin E, Peyrin C, Pichat C, Lamalle L, Le Bas JF, Baciu M (2007): Functional MRI approach for assessing hemispheric predominance of regions activated by a phonological and a semantic task. Eur J Radiol 63:274–285.

Davis C, Kleinman J, Newhart M, Gingis L, Pawlak M, Hillis A (2008): Speech and language functions that require a functioning Broca's area. Brain Lang 105:50–58.

Demonet J, Chollet F, Ramsay S, Cardebat D, Nespoulous J, Wise R, Rascol A, Frackowiak R (1992): The anatomy of phonological and semantic processing in normal subjects. Brain 115:1753–1768.

Dosenbach N, Fair D, Miezin F, Cohen A, Wenger K, Dosenbach R, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE (2007): Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci USA 104:11073–11078.

Dronkers N, Ogar J, Willock S, Wilkins D (2004): Confirming the role of the insula in coordinating complex but not simple articulatory movements. Brain Lang 91:23–24.

Friederici A (2011): The brain basis of language processing: From structure to function. Physiol Rev 91:1357–1392.

Friederici A, Meyer M, von Cramon DY (2000): Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. Brain Lang 75:289–300.

Friston K, Zarahn E, Josephs O, Henson R, Dale A (1999): Stochastic designs in event-related fmri. NeuroImage 10:607–619.

Gandour J, Tong Y, Wong D, Talavage T, Dzemidzic M, Xu YS, Li XJ, Lowe M (2004): Hemispheric roles in the perception of speech prosody. NeuroImage 23:344–357.

Geiser E, Zaehle T, Jancke L, Meyer M (2008): The neural correlate of speech rhythm as evidenced by metrical speech processing. J Cogn Neurosci 20:541–552.

Giraud A, Kell C, Thierfelder C, Sterzer P, Russ M, Preibisch C, Kleinschmidt A (2004): Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. Cereb Cortex 14:247–255.

Hagoort P (2005): On Broca, brain, and binding: A new framework. Trends Cogn Sci 9:416–423.

Hahne A, Jescheniak J (2001): What's left if the jabberwock gets the semantics? An ERP investigation into semantic and syntactic processes during auditory sentence comprehension. Brain Res Cogn Brain Res 11:199–212.

Hesling I, Clement S, Bordessoules M, Allard M (2005): Cerebral mechanisms of prosodic integration: Evidence from connected speech. NeuroImage 24:937–947.

Hickok G (2001): Functional anatomy of speech perception and speech production: Psycholinguistic implications. J Psycholing Res 30:225–235.

Hickok G (2009): The functional neuroanatomy of language. Phys Life Rev 6:121–143.

Ischebeck A, Friederici A, Alter K (2008): Processing prosodic boundaries in natural and hummed speech: An fMRI study. Cereb Cortex 18:541–552.

Jung-Beeman M (2005): Bilateral brain processes for comprehending natural language. Trends Cogn Sci 9:512–518.

Khateb A, Pegna AJ, Landis T, Michel CM, Brunet D, Seghier ML, Annoni JM (2007): Rhyme processing in the brain: An ERP mapping study. Int J Psychophysiol 63:240–250.

Khateb A, Pegna A, Michel C, Custodi M, Landis T, Annoni J (2000): Semantic category and rhyming processing in the left and right cerebral hemisphere. Laterality 5:35–53.

Kuest J, Karbe H (2002): Cortical activation studies in aphasia. Curr Neurol Neurosci Rep 2:511–515.

Lattner S, Meyer M, Friederici A (2005): Voice perception: Sex, pitch, and the right hemisphere. Hum Brain Mapp 24:11–20.

Liem F, Lutz K, Luechinger R, Jancke L, Meyer M (2012): Reducing the interval between volume acquisitions improves "sparse" scanning protocols in event-related auditory fMRI. Brain Topogr 25:182–193.

Lindenberg R, Fangerau H, Seitz R (2007): "Broca's area" as a collective term? Brain Lang 102:22–29.

Meinschaefer J, Hausmann M, Gunturkun O (1999): Laterality effects in the processing of syllable structure. Brain Lang 70:287–293.

Meyer M (2008): Functions of the left and right posterior temporal lobes during segmental and suprasegmental speech perception. Z Neuropsychol 19:101–115.

Meyer M, Alter K, Friederici A, Lohmann G, von Cramon DY (2002): FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. Hum Brain Mapp 17:73–88.

Meyer M, Jancke L. (2006): Involvement of the left and right frontal operculum in speech and nonspeech perception and production. In Grodzinsky Y, Amunts K, editors. Broca's Region. New York: Oxford University Press. pp 218–241.

Meyer M, Steinhauer K, Alter K, Friederici A, von Cramon DY (2004): Brain activity varies with modulation of dynamic pitch variance in sentence melody. Brain Lang 89:277–289.

Mutschler I, Wieckhorst B, Kowalevski S, Derix J, Wentlandt J, Schulze-Bonhage A, Ball T (2009): Functional organization of the human anterior insular cortex. Neurosci Lett 457:66–70.

Narain C, Scott SK, Wise RJ, Rosen S, Leff A, Iversen SD, Matthews PM (2003): Defining a left-lateralized response specific to intelligible speech using fMRI. Cereb Cortex 13:1362–1368.

Noesselt T, Shah N, Jancke L (2003): Top-down and bottom-up modulation of language related areas—An fMRI study. Bmc Neurosci 4:ARTN 13.

Orfanidou E, Marslen-Wilson WD, Davis M (2006): Neural response suppression predicts repetition priming of spoken words and pseudowords. J Cogn Neurosci 18:1237–1252.

Paulesu E, Frith C, Frackowiak R (1993): The neural correlates of the verbal component of working memory. Nature 362:342–345.

Plante E, Creusere M, Sabin C (2002): Dissociating sentential prosody from sentence processing: Activation interacts with task demands. NeuroImage 17:401–410.

Poeppel D (2003): The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. Speech Commun 41:245–255.

Poeppel D, Hickok G (2004): Towards a new functional anatomy of language. Cognition 92:1–12.

Poeppel D, Yellin E, Phillips C, Roberts TP, Rowley HA, Wexler K, Marantz A (1996): Task-induced asymmetry of the auditory evoked M100 neuromagnetic field elicited by speech sounds. Cognitive Brain Res 4:231–242.

Poldrack R, Wagner A, Prull M, Desmond J, Glover G, Gabrieli J (1999): Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. NeuroImage 10:15–35.

Praamstra P, Stegeman D (1993): Phonological effects on the auditory N400 event-related brain potential. Brain Res Cogn Brain Res 1:73–86.

Price CJ (2000): The anatomy of language: Contributions from functional neuroimaging. J Anat 197:335–359.

Pruessmann K, Weiger M, Scheidegger M, Boesiger P (1999): Sense: Sensitivity encoding for fast MRI. Magn Reson Med 42:952–962.

Rayman J, Zaidel E (1991): Rhyming and the right hemisphere. Brain Lang 40:89–105.

Rogalsky C, Hickok G (2011): The role of Broca's area in sentence comprehension. J Cogn Neurosci 23:1664–1680.

Rugg MD (1984): Event-related potentials in phonological matching tasks. Brain Lang 23:225–240.

Rugg M, Barrett S (1987): Event-related potentials and the interaction between orthographic and phonological information in a rhyme-judgment task. Brain Lang 32:336–361.

Scheich H, Brechmann A, Brosch M, Budinger E, Ohl F (2007): The cognitive auditory cortex: Task-specificity of stimulus representations. Hearing Res 229:213–224.

Schmidt C, Zaehle T, Meyer M, Geiser E, Boesiger P, Jancke L (2008): Silent and continuous fMRI scanning differentially modulate activation in an auditory language comprehension task. Hum Brain Mapp 29:46–56.

Shalom DB, Poeppel D (2008): Functional anatomic models of language: Assembling the pieces. Neuroscientist 14:119–127.

Smith EE, Jonides J (1999): Storage and executive processes in the frontal lobes. Science 283:1657–1661.

Stanovich K, Cunningham A, Cramer B (1984): Assessing phonological awareness in kindergarten-children—Issues of task comparability. J Exp Child Psychol 38:175–190.

Stefanatos G (2008): Speech perceived through a damaged temporal window: Lessons from word deafness and aphasia. Semin Speech Lang 29:239–252.

Steinbrink C, Ackermann H, Lachmann T, Riecker A (2009): Contribution of the anterior insula to temporal auditory processing deficits in developmental dyslexia. Hum Brain Mapp 30:2401–2411.

Sterzer P, Kleinschmidt A (2010): Anterior insula activations in perceptual paradigms: Often observed but barely understood. Brain Struct Funct 214:611–622.

Stowe LA, Haverkort M, Zwarts F (2005): Rethinking the neurological basis of language. Lingua 115:997–1042.

Sweet LH, Paskavitz J, Haley A, Gunstad J, Mulligan R, Nyalakanti P, Cohen RA (2008): Imaging phonological similarity effects on verbal working memory. Neuropsychologia 46:1114–1123.

Tervaniemi M, Hugdahl K (2003): Lateralization of auditory-cortex functions. Brain Res Brain Res Rev 43:231–246.

Thiel A, Haupt WF, Habedank B, Winhuisen L, Herholz K, Kessler J, Markowitsch HJ, Heiss WD (2005): Neuroimaging-guided rTMS of the left inferior frontal gyrus interferes with repetition priming. NeuroImage 25:815–823.

Treiman R (1985): Onsets and rimes as units of spoken syllables: Evidence from children. J Exp Child Psychol 39:161–181.

Turner R, Kenyon L, Trojanowski J, Gonatas N, Grossman M (1996): Clinical, neuroimaging, and pathologic features of progressive nonfluent aphasia. Ann Neurol 39:166–173.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002): Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15:273–289.

Vaden K, Muftuler L, Hickok G (2010): Phonological repetition-suppression in bilateral superior temporal sulci. NeuroImage 49:1018–1023.

Vigneau M, Beaucousin V, Herve P, Duffau H, Crivello F, Houde O, Mazoyer B, Tzourio-Mazoyer N (2006): Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. NeuroImage 30:1414–1432.

Vigneau M, Beaucousin V, Herve P, Jobard G, Petit L, Crivello F, Mellet E, Zago L, Mazoyer B, Tzourio-Mazoyer N (2011): What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? Insights from a meta-analysis. NeuroImage 54:577–593.

Zaehle T, Geiser E, Alter K, Jancke L, Meyer M (2008): Segmental processing in the human auditory dorsal stream. Brain Res 1220:179–190.

Zaehle T, Schmidt C, Meyer M, Baumann S, Baltes C, Boesiger P, Jancke L (2007): Comparison of “silent” clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. NeuroImage 37:1195–1204.

Zatorre R, Evans A, Meyer E, Gjedde A (1992): Lateralization of phonetic and pitch discrimination in speech processing. Science 256:846–849.

Zatorre R, Gandour J (2008): Neural specializations for speech and pitch: Moving beyond the dichotomies. Philos Trans R Soc Lond B Biol Sci 363:1087–1104.

Zekveld A, Heslenfeld D, Festen J, Schoonhoven R (2006): Top-down and bottom-up processes in speech comprehension. NeuroImage 32:1826–1836.

Zhang L, Shu H, Zhou F, Wang X, Li P (2010): Common and distinct neural substrates for the perception of speech rhythm and intonation. Hum Brain Mapp 31:1106–1116.

Zurowski B, Gostomzyk J, Gron G, Weller R, Schirrmeister H, Neumeier B, Spitzer M, Reske SN, Walter H (2002): Dissociating a common working memory network from different neural substrates of phonological and spatial stimulus processing. NeuroImage 15:45–57.