Published in final edited form as:

Neuroreport. 2008 July 16; 19(11): 1101–1105. doi:10.1097/WNR.0b013e3283060a9d.

Cortical and subcortical contributions to the attentive processing of speech

Thomas A. Christensen, Sharon M. Antonucci, Julie L. Lockwood, Megan Kittleson, and Elena Plante

Department of Speech, Language and Hearing Sciences, The University of Arizona, Tucson, Arizona, USA

Abstract

The neuroanatomical correlates of attentive listening were investigated with functional magnetic resonance imaging and an attention task in which listeners responded only to words that combined two specific attributes of voice and semantic content. This task was performed under two different attentive listening conditions: (i) diotically, with words presented sequentially, and (ii) dichotically, with male and female voices presented simultaneously but segregated to different ears. For both conditions, functional magnetic resonance imaging revealed bihemispheric but right-lateralized activity patterns in mid-prefrontal, anterior cingulate, and inferior parietal areas, as well as significant anterior insular and subcortical activation. Manipulating attentional demands under different listening conditions revealed an important role for right anterior insula, striatum, and thalamus in the regulation of attentive listening to spoken language.

Keywords

attention; dichotic listening; frontoparietal; functional magnetic resonance imaging; language; striatum; thalamus

Introduction

Listening attentively to speech engages different neural resources relative to passive listening [1–3]. In addition to activity changes in primary and secondary auditory cortices, frontal and parietal resources are also recruited by attention, and a number of human brain imaging studies are converging on a model that involves a distributed network of frontal and parietal components that contribute to attentional control [4–6]. For example, a recent functional magnetic resonance imaging (fMRI) study suggested that prefrontal activation in language attention tasks reflected executive control, whereas parietal activation varied as a function of multimodal perceptual processing [7]. However, other studies showed that changes in frontoparietal activity also occur with auditory stimuli alone [8].

Demands on selective attention can be increased through the use of a dichotic listening task relative to diotic listening, but results using this contrast have varied. Some imaging studies that used consonant–vowel (CV) syllables or continuous speech as stimuli reported seeing increased activity in right precuneus/inferior parietal lobule [1,9–12], whereas other studies of selective attention to CV syllables and whole words reported no significant parietal

activation in either hemisphere [3,13,14]. An earlier imaging study involving attentive diotic listening to nouns also did not reveal any significant increase in parietal cortex activity [15], whereas a more recent study using whole sentences reported right inferior parietal activity associated with voice recognition but not semantic processing [11]. It therefore remains unclear whether the observed variations in frontal and parietal activity were driven by directing attention to a particular sensory modality (visual or auditory) or, instead, to a particular class of stimulus (e.g. simple tones vs. CV syllables vs. whole words) [8,16].

To address this issue, we analyzed fMRI responses for diotic versus dichotic listening conditions to assess the neuroanatomical substrates for increased demands on attention to spoken words. In contrast to earlier studies, we designed auditory selective attention tasks in which the type of attentional control was varied while the focus of attention remained constant. We hypothesized that our word-comprehension tasks (unlike passive listening) would activate prefrontal executive resources, but that the pattern of activity would index the different modes of attentional control involved in diotic versus dichotic stimulus presentation. Furthermore, if inferior parietal activation reflects modality selection [7], we would predict no significant differences in parietal activation between these two auditory tasks. On the other hand, if inferior parietal resources are involved in switching between different types of attentional control, independent of modality [12], we would expect to see differential activity in this cortical region.

Methods

Participants

Five men and nine women (one left-handed) between the ages of 18–49 years were studied. None had a history of speech, language, or other neurological disorders, and all were screened for contraindications with MRI scanning. The study was approved by the University of Arizona Institutional Review Board, and all the participants read and signed consent forms.

Stimuli

Auditory stimuli were single one-syllable to three-syllable English nouns presented in a block design. Half were recorded in a male voice and half in a female voice (two distinct speaker choices), and half were names of animals and half were names of foods (two distinct semantic choices). Both speakers used a standard American English dialect, and the digital stimuli (16-bit, 44.1-kHz sampling rate) were edited using Sound Forge software (Sony Media Software, Inc., New York City, New York, USA) so that their perceived loudness was approximately equal across trials. Words had durations of 197–1204 ms (mean $\pm$ SD=555 $\pm$ 173 ms) and the range of interstimulus intervals was 314–1339 (871 $\pm$ 157 ms) to accommodate variations in word length. Stimuli were played through MRI-compatible headphones (Resonance Technology, Inc., Northridge, California, USA).

Experimental procedures

Demands on selective attention were manipulated while keeping both the verbal content and task instructions identical. Both listening conditions required participants to attend to a stream of verbal stimuli and select, via button press, those words that corresponded to a specific combination of target speaker (female) plus a target semantic category (animals). Cognitive demand was manipulated by using diotic presentation for one scan and a modified dichotic paradigm for the other. For diotic listening, a dual decision based on integrating voice and semantic information was required. For dichotic listening, in contrast, participants were required to attend to one ear (alternated with each block to control for lateralization effects) while ignoring or suppressing input from the other. In addition, each dichotic

stimulus block was preceded by a verbal instruction in the target channel that forced attention to the appropriate ear just before the start of the stimulus block. In other words, once attention was directed appropriately, the selection of semantic information, but not voice information, was required to make a decision.

The ratio of targets (animal word + female voice) to non-targets was 1: 3 in both tasks. In the dichotic task (two stimuli simultaneously), the number of targets was doubled to maintain a target/nontarget ratio of 1: 3. Also, words in the two auditory channels were edited so that their onsets were synchronized, and the number of syllables and their syllable-level stress patterns were matched. To improve attentional focus, participants were also instructed to remember the target items and were told that their memory for these words would be tested later (these results will be the subject of a subsequent report).

For both attention conditions, each stimulus block was alternated with a block of the same words played in reverse, and these two types of blocks were repeated six times per scan. Reversed speech produces sublexical nonsense stimuli that are often used as a control for primary acoustic processing, but reverse playback also conveys some phonetic information [17]. The contrast between forward and reverse blocks therefore corrected for the primary acoustic and speech-like processing stages that precede semantic processing of intelligible speech.

Functional magnetic resonance imaging scan parameters and analysis

All scans were acquired with a 3.0 T GE Signa VH/i scanner (General Electric Medical Systems, Madison, Wisconsin, USA) equipped with a quad-head RF coil. Twenty-six contiguous axial slices had an in-plane resolution of $3.44 \times 3.44 \times 5$ mm and covered the entire brain. Two functional scans (counterbalanced) were acquired using a spiral in/out pulse sequence (TR=2.3 s, TE=30 ms, flip angle=90°). High-resolution SPGR images were obtained in the sagittal plane (TR=30 ms, TE=minimum, flip angle=30°) and used for localization and coregistration of functional data across participants after transformation into Talairach space.

Image processing and statistical analyses were carried out with the AFNI suite of programs (<http://afni.nimh.nih.gov/afni/>). Analysis of fMRI images followed standard procedures for preprocessing data, including spatial smoothing (4 mm FWHM) and statistical parametric modeling of hemodynamic responses. Data were normalized and hemodynamic models were developed and offset by 0 to 4 s in 0.5-s increments to account for interparticipant and intra-participant variations in the onset of the blood oxygenation level-dependent response. An iterative regression procedure was used to fit the different models to the signal obtained from each participant and the results from the best-fitting model were retained for further analysis.

Group activation was analyzed with a two-factor analysis of variance, and a Monte Carlo simulation was used to correct for multiple comparisons (at $P=0.01$, minimum cluster size=4 voxels). For those clusters above this threshold, T-scores are reported in Table 1. Peak activation loci were determined by mapping Talairach coordinates. If these locations were ambiguous, the activation maps were overlaid on the individual participants' SPGR scans and reexamined on a participant-by-participant basis.

Results

Using a subjective five-point rating scale, participants ($n=13$) judged the dichotic listening task to be more challenging than the diotic listening task [2.46 ± 0.85 vs. 1.7 ± 0.69 ;

$F(1,24)=5.99$, $P=0.02$]. In contrast, no significant difference in response accuracy was found [$F(1,24)=0.85$, $P=0.37$], indicating that the level of difficulty was similar for the two tasks.

Regions of activation for the two attention conditions are summarized in Table 1. Overall, the diotic listening condition resulted in greater blood oxygenation level-dependent signal amplitudes as well as a greater number of activated brain regions (blue map in Fig. 1). The strongest signals were localized bilaterally in dorsal anterior cingulate cortex and anterior insula (AIC) (Table 1). In other areas, activation was asymmetric across the two hemispheres. Temporal lobe activity in Heschl's gyrus, posterior superior temporal gyrus and sulcus, and fusiform gyrus was strongly left lateralized (Table 1; Fig. 1b). Regional frontal activation was greatest in the anterior portion of the inferior frontal sulcus [Brodmann areas (BA) 9/10] in the right hemisphere, whereas peak activation in the left hemisphere was centered more posteriorly over the inferior frontal junction (BA 6/9). In addition, diotic attention specifically engaged the right frontopolar cortex (Table 1), whereas the dichotic condition alone engaged the left superior frontal gyrus (Fig. 1b).

Both attention tasks also elicited inferior parietal activity (BA 40), but only in the right hemisphere. Importantly, however, the location of the activation maxima differed for the two attention conditions. In the diotic condition, parietal activity was centered over deep intraparietal sulcus near precuneus. In contrast, the dichotic attention condition engaged an area within the dorsal supramarginal gyrus (Fig. 1a).

Both attentive listening conditions also recruited significant subcortical activity that was readily visible at 3.0T. For diotic attention (Fig. 2a), activation maxima were localized to bilateral AIC and to thalamic nuclei that corresponded most closely to the area of the ventral-anterior and ventral-lateral nuclei, but activity also spread posteriorly to the bilateral medial-dorsal nuclei. In the striatum, peak activity was lateralized to right caudate, with a smaller activity focus in left putamen (Table 1). For dichotic listening, however, suprathreshold activation in the striatum and thalamus was seen only in the right hemisphere (Fig. 2b). This lateralization occurred despite the fact that the direction of attention was shifted between ears from block to block during dichotic listening. Unlike the cortical activation patterns, a significant difference in subcortical activation was seen across conditions, with greater activity elicited in right AIC, striatum (mostly caudate body and tail), and thalamus during diotic listening (Fig. 2c; $P=0.01$, corrected).

Discussion

In contrast to an earlier study [18], we found that dichotic listening, which involves suppression/inhibition of one stimulus while attending simultaneously to another, did not elicit significantly greater frontal activation when compared with diotic listening. We note, however, that our diotic task is more complex in that it involves word recognition (vs. listening to tones or CV syllables), and it requires a 'dual decision' based on integration of voice and semantic content. Thus, our activations probably index different cognitive operations that are involved in completing this more demanding word-recognition task. We also note that in our cued dichotic-listening condition, once attention was directed to the ear in which the targets would be heard, the listener's response was not dependent on the analysis of voice content. Thus, although we cannot rule out that contralateral suppression/inhibition accounts for some reduced activity in dichotic attention conditions, we believe that the increased cognitive demands required of the diotic task probably raised the overall level of middle prefrontal activation as compared with dichotic listening.

Recent evidence points to an important role for middle prefrontal cortex in the inhibitory gating of sensory processes and the short-term maintenance of attention [19]. In rats,

damage to this region leads to an attention deficit and expression of impulsive behavior, and injection of glutamate-receptor antagonists into middle prefrontal cortex leads to both cognitive and behavioral deficits resembling the effects of prefrontal lesions on executive functioning [20]. These data emphasize the importance of middle prefrontal regions in attentional modulation, and our results using solely auditory stimulation provide new evidence that attentional control over auditory language processing is not limited to areas of prefrontal cortex that are traditionally associated with either motor aphasia (Broca's area) or visual attention (frontal eye fields) [6]. They instead provide fresh support for the hypothesis that a combination of resources in multiple prefrontal areas is engaged bilaterally during the attentional control of language processing.

In accordance with prior studies [1,9–12], activity in right inferior parietal cortex was also associated with attention to our verbal stimuli, but the intensity of the activation did not simply increase during dichotic listening. Instead, there was a subtle shift in activation maxima from intraparietal sulcus during diotic attention to supramarginal gyrus under dichotic listening conditions. Activation spreading across inferior parietal cortex is consistent with the idea that this area serves as an attentional focus for verbal materials [21]. One fMRI study of bimodal (auditory and visual) language processing concluded that inferior parietal activation was associated mainly with the increased demands of dividing attention between modalities [7]. In contrast, our results using only auditory stimuli demonstrate a role in selective attention for multiple areas of the right inferior parietal cortex that does not involve multimodal control [10].

Our results also provide new evidence supporting the role of emotion-processing centers and subcortical structures in attentive listening [8]. Consistent with this idea, the AIC and dorsal anterior cingulate cortex (right lateralized), both components of a so-called 'salience network' [22], were activated under both of our attention conditions. AIC receives a pronounced afferent projection from the thalamus [23], and our contrast analysis showed that increased insular, thalamic, and striatal involvement under diotic attention conditions was localized to the right hemisphere. This result is in agreement with a previous PET study that examined attention to tones versus visual stimuli [24], and is also consistent with the hypothesis that thalamic nuclei regulate cortical interactions between language and attention [25].

Conclusion

These results provide new insight into the effects of different listening demands on cortical and subcortical involvement during top-down, attention-modulated processing of normal human speech. Distinct frontoparietal networks are selectively recruited to satisfy different attentional demands, and our data point to an important role for right anterior insula, anterior cingulate, striatum, and thalamus in the regulation of attentive listening to spoken language.

Acknowledgments

The authors are grateful to Scott Squire for expert technical assistance. This work was supported by a grant from the Arizona Alzheimer's Research Center to E. Plante, and an NIH-NIDCD Career Development Award (1 K01 DC008812) to T. Christensen.

References

1. Alho K, Vorobyev VA. Brain activity during selective listening to natural speech. *Front Biosci.* 2007; 12:3167–3176. [PubMed: 17485291]
2. Price CJ. The anatomy of language: contributions from functional neuroimaging. *J Anat.* 2000; 197:335–359. [PubMed: 11117622]

3. Hugdahl K, Thomsen T, Ersland L, Rimol LM, Niemi J. The effects of attention on speech perception: an fMRI study. *Brain Lang.* 2003; 85:37–48. [PubMed: 12681347]
4. LaBerge D. Sustained attention and apical dendrite activity in recurrent circuits. *Brain Res Rev.* 2005; 50:86–99. [PubMed: 15921761]
5. Raz A, Buhle J. Typologies of attentional networks. *Nat Rev Neurosci.* 2006; 7:367–379. [PubMed: 16760917]
6. Corbetta M, Shulman G. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci.* 2002; 3:201–215. [PubMed: 11994752]
7. Shaywitz BA, Shaywitz SE, Pugh KR, Fulbright RK, Skudlarski P, Mencl WE, et al. The functional neural architecture of components of attention in language-processing tasks. *Neuroimage.* 2001; 13:601–612. [PubMed: 11305889]
8. Jäncke L, Specht K, Shah JN, Hugdahl K. Focused attention in a simple dichotic listening task: an fMRI experiment. *Cogn Brain Res.* 2003; 16:257–266.
9. Lipschutz B, Kolinsky R, Damhaut P, Wikler D, Goldman S. Attention-dependent changes of activation and connectivity in dichotic listening. *Neuroimage.* 2002; 17:643–656. [PubMed: 12377140]
10. Pugh KR, Shaywitz BA, Shaywitz SE, Fulbright RK, Byrd D, Skudlarski P, et al. Auditory selective attention: an fMRI investigation. *Neuroimage.* 1996; 4:159–173. [PubMed: 9345506]
11. von Kriegstein K, Eger E, Kleinschmidt A, Giraud AL. Modulation of neural responses to speech by directing attention to voices or verbal content. *Cogn Brain Res.* 2003; 17:48–55.
12. Hugdahl K, Law I, Kyllingsbaek S, Brønnick K, Gade A, Paulson O. Effects of attention on dichotic listening: an ¹⁵O-PET study. *Hum Brain Map.* 2000; 10:87–97.
13. Thomsen T, Rimol L, Ersland L, Hugdahl K. Dichotic listening reveals functional specificity in prefrontal cortex: an fMRI study. *Neuroimage.* 2004; 21:211–218. [PubMed: 14741658]
14. Jäncke L, Buchanan TW, Lutz K, Shah NJ. Focused and nonfocused attention in verbal and emotional dichotic listening: an fMRI study. *Brain Lang.* 2001; 78:349–363. [PubMed: 11703062]
15. Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME. Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature.* 1988; 331:585–589. [PubMed: 3277066]
16. Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci.* 2007; 8:393–402. [PubMed: 17431404]
17. Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Springer JA, Kaufman JN, et al. Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex.* 2000; 10:512–528. [PubMed: 10847601]
18. Jäncke L, Shah NJ. Does dichotic listening probe temporal lobe functions? *Neurol.* 2002; 58:736–743.
19. Mears RP, Klein AC, Cromwell HC. Auditory inhibitory gating in medial prefrontal cortex: single unit and local field potential analysis. *Neurosci.* 2006; 141:47–65.
20. Carli M, Baviera M, Invernizzi RW, Balducci C. Dissociable contribution of 5-HT_{1A} and 5-HT_{2A} receptors in the medial prefrontal cortex to different aspects of executive control such as impulsivity and compulsive perseveration in rats. *Neuropsychopharmacology.* 2006; 31:757–767. [PubMed: 16192987]
21. Chein JM, Ravizza SM, Fiez JA. Using neuroimaging to evaluate models of working memory and their implications for language processing. *J Neurolinguist.* 2003; 16:315–339.
22. Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, et al. Dissociable intrinsic connectivity networks for salience processing and executive control. *J Neurosci.* 2007; 27:2349–2356. [PubMed: 17329432]
23. Dupont S, Boullieret V, Hasboun D, Semah F, Baulac M. Functional anatomy of the insula: new insights from imaging. *Surg Radiol Anat.* 2003; 25:113–119. [PubMed: 12819943]
24. Frith CD, Friston KJ. The role of the thalamus in ‘top down’ modulation of attention to sound. *Neuroimage.* 1996; 4:210–215. [PubMed: 9345511]
25. Radanovic M, Azambuja M, Mansur L, Porto C, Scaff M. Thalamus and language: interface with attention, memory and executive functions. *Arq Neuro-Psiquiatr.* 2003; 61:34–42.

Fig. 1.

Components of the bihemispheric 'attentive listening' networks ($n=14$; $P=0.01$, corrected). The diotic attention network is in blue, the partially overlapping dichotic attention network is in red. Seven principal regions of interest are labeled. (a) Right hemisphere angled forward, illustrating differential parietal activity for the two attention conditions. (b) Left hemisphere view showing differential posterior superior temporal gyrus (pSTG) activity for diotic attention and superior frontal gyrus (SFG) activity for dichotic attention. aIFS, anterior portion of the inferior frontal sulcus; dACC, dorsal anterior cingulate cortex; IPS, intraparietal sulcus; pIFS, posterior portion of the inferior frontal sulcus; SMG, supramarginal gyrus.

Fig. 2.

Cluster-based contrast maps of regional blood oxygenation level-dependent responses under the two attention conditions. (a) Diotic condition, (b) cued dichotic-listening condition (both relative to reversed speech), and (c) contrast of diotic > dichotic. Columns are representative slices taken at three levels through the inferior–superior axis (Z-coordinates in Talairach space are in mm). In (a) and (b), blue: $P=0.01$, uncorrected; yellow: $P=0.01$, corrected; and in (c), red: $P=0.01$, corrected. AIC, anterior insula; aIFS, anterior portion of the inferior frontal sulcus; Cb, caudate body; Ct, caudate tail; HG, Heschl’s gyrus; L, left; MD, medial-dorsal nuclei; pIFS, posterior portion of the inferior frontal sulcus; Put, putamen; R, right; STG, superior temporal gyrus; VA-VL, ventral-anterior and ventral-lateral nuclei.

Table 1

Brain activations during the two attentive listening tasks versus reversed speech ($n=14$; $P<0.01$, corrected; clusters >237 voxels in Talairach space)

	Diotic listening				Dichotic listening			
	x	y	z	$T_{\max}$	x	y	z	$T_{\max}$
Right hemisphere								
dACC	4	13	44	10.44	3	7	51	7.34
Anterior insula	31	20	9	9.33	31	20	7	8.66
Middle frontal gyrus								
Anterior IFS	31	36	26	5.40	31	46	24	6.54
Frontopolar (dorsal)	34	51	10	4.47	—	—	—	—
Frontopolar (ventral)	29	56	-9	4.10	—	—	—	—
Intraparietal sulcus	28	-46	39	4.13	—	—	—	—
Supramarginal gyrus	—	—	—	—	42	-37	40	3.84
Caudate	16	1	19	4.52	15	-2	19	4.99
Putamen/caudate	—	—	—	—	—	—	—	—
Thalamus	14	-9	13	7.26	14	-10	15	4.06
Left hemisphere								
dACC	-6	6	51	8.38	-6	6	51	7.24
Anterior insula	-36	18	8	9.69	-34	20	9	5.87
Middle frontal gyrus								
Posterior IFS/IFJ	-42	12	26	5.13	-43	-1	32	4.80
Superior frontal gyrus	—	—	—	—	-35	45	30	5.70
Fusiform gyrus	-32	-34	-19	3.83	-36	-36	-17	4.25
Heschl's gyrus	-44	25	15	4.65	—	—	—	—
pSTG/pSTS	-57	-39	10	4.54	—	—	—	—
Caudate	-11	0	19	4.10	—	—	—	—
Putamen/caudate	-23	12	16	4.18	—	—	—	—
Thalamus	-13	-6	12	6.31	—	—	—	—

Data represent mean values averaged across all participants. T -scores are at cluster maxima; x, y, z: Talairach coordinates.

dACC, dorsal anterior cingulate cortex; IFJ, inferior frontal junction; IFS, inferior frontal sulcus; pSTG, posterior superior temporal gyrus; pSTS, posterior superior temporal sulcus.