

# Selective attention to lexical tones recruits left dorsal frontoparietal network

Xiaojian Li, Jack Gandour,<sup>CA</sup> Thomas Talavage,<sup>1</sup> Donald Wong,<sup>2</sup> Mario Dzemidzic,<sup>3</sup> Mark Lowe<sup>3</sup> and Yunxia Tong

Department of Audiology and Speech Sciences, Purdue University, 1353 Heavilon Hall, 500 Oval Drive, West Lafayette, IN 47907-2038; <sup>1</sup>School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN; <sup>2</sup>Department of Anatomy and Cell Biology; <sup>3</sup>Department of Radiology, Indiana University School of Medicine, Indianapolis, IN, USA

<sup>CA</sup>Corresponding Author: gandour@purdue.edu

Received 21 August 2003; accepted 28 August 2003

DOI: 10.1097/01.wnr.0000097045.56589.3f

Phonological processing activates a posterior superior region of inferior prefrontal cortex, but questions still remain about the relationship between phonology and this particular region. In this fMRI experiment, subjects were asked to match an intrasyllabic unit (Chinese tones) in an experimental condition vs. whole syllables in a control condition. The only difference between conditions is mediated by focus of attention, either to a subpart (i.e. tone) of

the syllable or to the whole syllable itself. Phonetic extraction of Chinese tones reveals a dorsal frontoparietal network in the LH that engages selective attention and internal guidance, two mediational components that are not restricted to phonological processing. *NeuroReport* 14:2263–2266 © 2003 Lippincott Williams & Wilkins.

**Key words:** Auditory perception; Attention; Chinese; fMRI; Immediate recall; Language; Phonetics; Pitch perception; Speech perception; Working memory

## INTRODUCTION

Phonetic extraction from auditory input has been demonstrated to activate a posterior superior subregion of the inferior prefrontal cortex [1]. However, questions remain about its functional role. Does its activation reflect phonological processing and/or its working memory in the auditory modality, or instead reflect other mediational, task-specific processes that transcend the cognitive domain? This same subregion has been observed to be activated in response to judgments about semantic [2,3], syntactic [4], or even spatial [5,6] categories, in addition to phonological judgments via visual input [3,7,8]. Thus, its role is neither material specific for phonology, nor process specific for vocal rehearsal. We therefore assume that phonetic extraction is mediated by a network of widely distributed cortical regions for more general functions.

The aim of this experiment is to identify the neural network underlying the phonetic extraction of suprasegmental sub-syllabic units, i.e. Chinese tones (e.g. ma<sup>1</sup> 'mother', ma<sup>2</sup> 'hemp', ma<sup>3</sup> 'horse', ma<sup>4</sup> 'scold'), in the auditory modality. In this experiment, tonal units are matched to each other in an experimental condition *vs* a control condition in which whole syllables are matched to each other. In both conditions, auditory stimuli and task procedures are kept constant. Articulatory recoding from text to speech is not required. Semantics, memory load and verbal rehearsal are negligible. In so doing, we are able to segregate

primarily the extraction of sub-syllabic units from whole syllables, mediated by selective attention, which is guided either to a subpart of the syllable or to the whole syllable itself.

## MATERIALS AND METHODS

**Subjects:** Twelve adult native speakers of Mandarin Chinese (six male; six female) from mainland China, ranging in age from 23 to 32 years, participated in this study. All subjects were strongly right-handed as measured by the Edinburgh Handedness Inventory [9] and exhibited normal hearing sensitivity at 0.5, 1, 2, and 4 kHz. All subjects gave informed consent in compliance with a protocol approved by the Institutional Review Board of Indiana University Purdue University Indianapolis and Clarian.

**Stimuli:** A list of three Chinese monosyllables followed by a monosyllabic probe made up a sequence for each trial. All stimuli were produced by a native male speaker from Beijing. In matching sequences for the tone task (T), only the last item in the list shared a tone in common with its probe (Table 1). No other phonetic units in any of the syllables matched those in the probe. In non-matching sequences for T, no phonetic units (i.e. onset, rhyme, tone) matched those in the probe. In matching sequences for the syllable task (S), only the last item in the list matched the probe. In

**Table 1.** Sample auditory stimuli for tone and syllable matching.

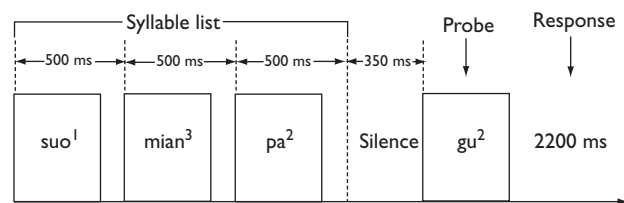
Phonetic Unit	Syllable list		Probe	Match	
Tone	rou <sup>2</sup>	bei <sup>3</sup>	duo <sup>4</sup>	na <sup>4</sup>	yes
	bo <sup>2</sup>	ka <sup>3</sup>	tu <sup>1</sup>	meng <sup>4</sup>	no
Syllable	rou <sup>2</sup>	bei <sup>3</sup>	<b>duo<sup>4</sup></b>	<b>na<sup>4</sup></b>	no
	bo <sup>2</sup>	ka <sup>3</sup>	<b>tu<sup>1</sup></b>	<b>tu<sup>1</sup></b>	yes

Chinese syllables are written in pinyin transcription. Target phonetic units are in bold type. Tones are indicated in superscript.

non-matching sequences for S, there were no matching syllables. Matching and non-matching sequences were equal in number. The T and S tasks shared the same syllable sequences except that the last items were made identical to their probes in matching sequences for S. Within a sequence, no adjacent syllables formed a bisyllabic word. No easily confusable pairs of consonants (e.g., /ch, c/, /ch, sh/, /s, z/) or rhymes (e.g., /an, ang/) were used. No adjacent syllables were subject to tone sandhi effects. Thirty-two sequences of a syllable list and probe were constructed for T and S, respectively.

**Tasks:** The scanning sequence consisted of two tasks (T, S) presented in blocked format (36s) alternating with 18s rest intervals. The order of presentation of T and S was counterbalanced across subjects. Half of the subjects started with T, the other half with S. Total duration of the sequence was 7.5 min. There were four blocks per task. Each block contained eight 4.5s trials, four matching and four non-matching, in random order. The timing sequence of a trial consisted of the syllable list and probe (2300 ms) and the response interval (2200 ms; Fig. 1). Instructions were delivered to subjects via headphones near the end of the rest intervals immediately preceding each task (in Chinese): last tone for T and last syllable for S. Subjects were asked to ignore the first two items in the syllable list and make a matching judgment on the last item in comparison to the probe according to the task (T or S). They responded by pressing a mouse button (left = yes, right = no). Presentation time for the last item and probe was within the spans for immediate memory [10] and attention [11]. Both the last item and probe were preceded by syllables, thus minimizing the possibility of an unequal prefix effect [12]. Prior to the actual scanning session, subjects were trained to a high level of accuracy (90%) on all experimental tasks using stimuli different from those used during image acquisition.

**Imaging protocol:** Subjects lay supine in a 1.5T Sigma GE LX Horizon scanner equipped with birdcage transmit-receive radiofrequency head coils. Gradient-echo EPI images were acquired with the following parameters: TR/TE 2.25s/50 ms; matrix 64 × 64; flip angle (FA) 90°; 24 × 24 cm FOV; 125 kHz receiver bandwidth. Sixteen contiguous 7.5 mm axial slices without an interslice gap were used to image the entire cerebrum. The effects of head motion were minimized by using a head-neck pad and dental bite bar. The raw fMRI data were Hamming-filtered, spatially, to increase the BOLD contrast-to-noise ratio with only a small loss of spatial resolution [13]. Prior to functional imaging scans, high-resolution, anatomic images were acquired in 124 contiguous axial slices using a 3D spoiled-Grass



**Fig. 1.** Structure and timing sequence of a trial, consisting of a syllable list, probe, and response interval. Sample trial illustrates a match for tone between the last item and the probe.

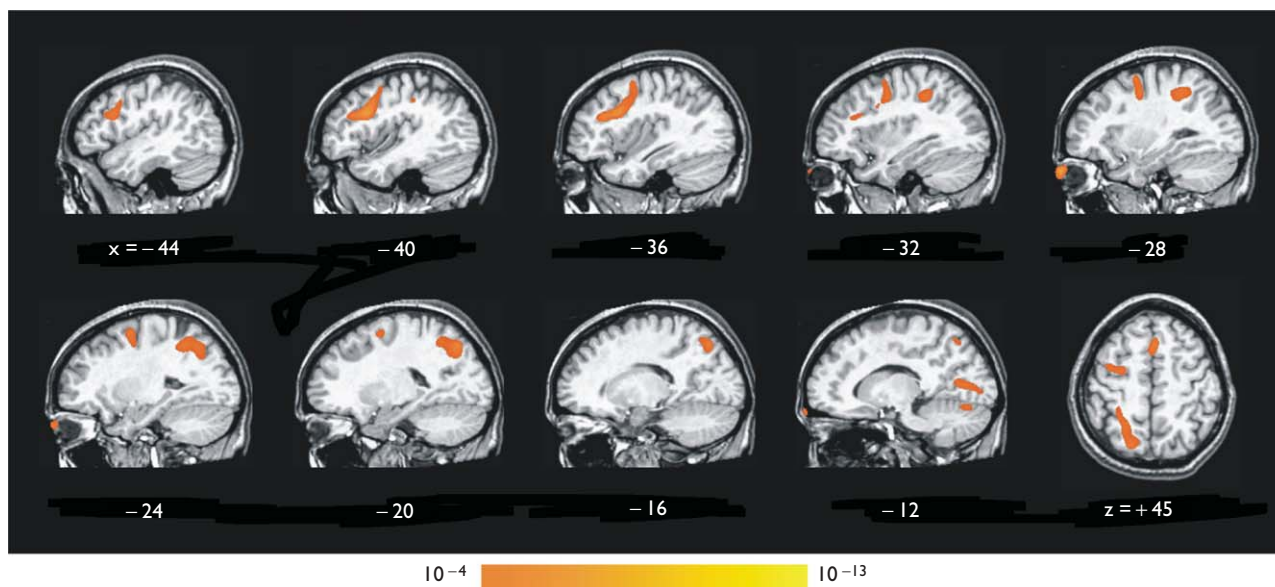
sequence (slice thickness 1.1–1.2 mm; TR/TE 35 ms/8 ms; 1 excitation; 30° FA; matrix 256 × 128; FOV 24 × 24 cm; receiver bandwidth 32 kHz) for purposes of anatomic localization and co-registration.

**Imaging analysis:** Group analysis was conducted with the AFNI software package [14]. First, for every subject, the volumes acquired in the functional imaging scan were rigid-body motion-corrected to the fourth volume acquired. The fourth volume was chosen as the reference because it was acquired after steady-state magnetization had been achieved, but still sufficiently early to provide good registration with respect to the high-resolution anatomic images. Further, this reference volume was acquired prior to the onset of the initial task block, thus avoiding task-related effects present in subsequent functional volumes. Second, the data were detrended by performing the least-squares fit to a second-order polynomial to the time-series for each voxel. Signal changes in the time series were analyzed using the AFNI deconvolution option, allowing hemodynamic responses to the two tasks (T, S) to be deconvolved from the baseline (rest) and directly compared to one another. The brain activation map of interest (T vs S) was constructed by computing Student's *t*-statistic values for each voxel in the imaging volume based upon a comparison of the voxel time-course with the hemodynamic response, predicted by the AFNI Waver function, to both the T and S tasks (activation favoring the T task was assigned a positive *t*-statistic value).

Using AFNI software, the resulting statistical volumes were transformed to 1 mm isotropic voxels in the Talairach [15] coordinate system, enabling the combination of data across subjects. The Talairach-converted *t*-statistic maps (one for each subject) were blurred using a 5.4 mm FWHM Gaussian kernel, and combined into a single average (group) dataset. Significantly activated voxels ( $p < 0.001$ ) located within a radius of 1 mm were grouped into clusters, with a minimum volume of 325 ml used as the cluster threshold ( $p < 0.0001$ ) based on simulations conducted using the AFNI AlphaSim program.

## RESULTS

**Behavioral data:** Mean ( $\pm$  s.d.) percentage response accuracy was not significantly different between the T ( $98.0 \pm 4.1$ ) and S ( $98.8 \pm 2.0$ ) tasks ( $t(16)_{\text{Satterthwaite}} = 0.57$ , n.s.). Mean response time was significantly ( $t(22) = -3.44$ ,  $p = 0.0024$ ) slower for the T than the S task (T  $568 \pm 197$  ms; S:  $344 \pm 111$  ms).



**Fig. 2.** Averaged fMRI activation maps for the contrast between Chinese tones and syllables. A series of sagittal sections, from  $x = -44$  to  $-12$ , show thresholded activation data superimposed on a representative brain in stereotaxic space [16]. Activation is seen in the left middle frontal gyrus ( $x = -44$  to  $-20$ ) and left superior parietal lobule ( $x = -32$  to  $-12$ ). An axial section ( $z = +47$ , bottom row right) shows that activity in these dorsolateral frontal and parietal areas was predominantly left-sided. Stereotaxic  $x$  coordinates refer to medial-lateral position relative to midline (negative = left; positive = right);  $z$  coordinates refer to superior-inferior position relative to the commissural line (positive = superior).

**fMRI data:** Peak activation is located in the left inferior precentral sulcus (PCS) in close proximity to its junction with the inferior frontal sulcus (IFS) (Fig. 2,  $x = -40$ ). This discrete focus extends rostrally along the course of the IFS ( $x = -36$ ) and dorsally into the posterior middle frontal gyrus (MFG) and precentral gyrus ( $x = -36$ ). A peak focus of activation is observed in the posterior superior parietal lobule (SPL). This focus extends rostrally throughout the SPL into close proximity to the intraparietal sulcus ( $x = -24$ ). Peak activation is also observed in the medial frontal gyrus, i.e. supplementary motor area (pre-SMA), about 22 mm rostral to the anterior commissure ( $z = 45$ ). Further details (t-values, volume, Talairach coordinates, Brodmann areas) about these frontoparietal foci as well as others are listed in Table 2.

## DISCUSSION

Phonetic extraction of Chinese tones elicits activity in a left-sided dorsal frontoparietal, attention network, including a dorsolateral subregion of posterior prefrontal cortex in the vicinity of the PCS/IFS junction, a medial frontal pre-SMA region, and distinct regions within the SPL. This network, in full or in part, has been observed in previous studies of phonological processing in the auditory modality of both segmental [1,16] and suprasegmental [17] units.

Activity in the PCS/IFS junction area is observed in tasks where subjects are required to selectively attend to particular characteristics of a stimulus regardless of sensory modality or linguistic category (e.g., phonology, semantics, syntax). In this study, subjects focused on Chinese tonal phonemes via auditory input while ignoring other phonetic characteristics of a syllable. Using auditory stimuli, other studies have also demonstrated activity in the PCS/IFS area when subjects are asked to focus on segmental phonemes [1] or on syllables that occur in specific serial positions in a list

**Table 2.** Brain areas showing significant activation for the direct comparison between tone and syllable tasks.

	BA	x	y	z	Peak t-value	Extent (ml)
Frontal						
L MFG	8/9/44	-39	13	26	5.79	5.96
M mdFG	6	3	22	45	4.75	1.17
Parietal						
L SPL	7	-19	-69	44	4.88	4.47
Occipital						
L LG	17/18	-8	-81	6	4.81	2.50
Cerebellum						
M vermis	1	-65	-26		5.77	7.72

BA = Brodmann area; L = left, M = medial, MFG = middle frontal gyrus, mdFG = medial frontal gyrus, SPL = superior parietal lobule, LG = lingual gyrus. Coordinates ( $x$ ,  $y$ ,  $z$ ) correspond to the Talairach atlas.

[18]. Similar findings have been reported for phonological tasks in the visual modality [3,7,8]. What is common to these phonological studies is the requirement that subjects direct their attention selectively to a syllable-internal phonological unit (consonant, vowel, tone) or selectively to the whole syllable itself. But this region is not dedicated to phonological categories exclusively. Tasks in which subjects are required to focus on semantic [2,3] or syntactic [4] categories, while ignoring others of the same type, similarly activates the PCS/IFS junction area.

This claim is corroborated by studies of selective attention using visuospatial tasks [6] and auditory/visual language tasks [5]. In addition, selective attention requires cue-related preparation for an expected target specified prior to the stimulus. In this study, subjects are likely to prepare to decode the pitch contours of a target syllable and to match the target to its following probe. Cue-related preparation yields activation of the left PCS/IFS junction area as well [6,19].

The left SPL is activated when subjects voluntarily (top-down *vs* stimulus-driven, bottom-up) orient their attention along the input processing stream [19–22]. In this study, subjects' attention is voluntarily engaged to the tone of the last syllable, disengaged, and then re-engaged to the following probe. The activity of SPL also reflects sustained engagement over a short time period [6,21,23], which possibly enables the to-be-matched target and probe to be paired for comparison.

Pre-SMA activity likely reflects subjects' internal guidance for extracting tones. Such guidance is necessary because the last syllable and its probe may differ in their segmental composition as well as their tone (e.g. duo<sup>4</sup> *vs* na<sup>4</sup>; tu<sup>1</sup> *vs* na<sup>4</sup>). Therefore, tone matching cannot be based simply on acoustic perception, as is the case for syllable matching (e.g., na<sup>4</sup> *vs* na<sup>4</sup>; duo<sup>4</sup> *vs* na<sup>4</sup>). This guidance may take the form of a schematized procedure which organizes processing into efficient steps, including subvocal decoding of tones, switching from syllables to sub-syllabic tones, and comparison of tones between the target and its probe. Activity in this region has been previously reported in phonological [7,17] and semantic [2] tasks. Moreover, it has been shown that free response within a category as compared to a specified response [24], as well as cue-related preparation [19], leads to an increase in pre-SMA activity.

The absence of significant brain activity in ventral aspects of the inferior frontal gyrus (IFG) or the inferior parietal lobule (IPL) is noteworthy [17]. In working memory studies, the former has been implicated in articulatory rehearsal, the latter in phonological storage [25]. Segmental (consonant or rhyme) matching on Chinese characters activates ventral aspects of the IFG [3,7,8]. However, this activation may reflect the conversion of visual text to speech. In this study, such conversion is unnecessary because the input modality is auditory, even though articulatory recoding from audition to speech may be necessary.

Our experimental paradigm keeps rehearsal and storage demands to a minimum in both tasks because there is no distracter between the target stimulus and its probe. This design feature makes it possible to segregate the effects of selective attention in phonological processing minus a working memory confound.

## CONCLUSIONS

Phonetic extraction of Chinese tones reveals a cortical network in the LH that recruits mediational processes related to selective attention and internal guidance. These processes apply to language as well as other domains. The pre-SMA retrieves a memory scheme for initiating preparation prior to each processing step in the task; the PCS/IFS junction area encodes the perceived signal, stores it for immediate matching, and transfers the matched code to the neighboring motor area for a response; the SPL marks and maintains the engagement of the activity in the PCS/IFS area so that the to-be-matched target and probe can be properly paired. In the language domain, these mediational processes make it possible for a

listener to extract linguistically significant parameters in his/her native language.

## REFERENCES

- Burton MW. The role of inferior frontal cortex in phonological processing. *Cogn Sci* 25, 695–709 (2001).
- Chee MW, Hon N, Lee HL and Soon CS. Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. *Neuroimage* 13, 1155–1163 (2001).
- Tan LH, Liu HL, Perfetti CA, Spinks JA, Fox PT and Gao JH. The neural system underlying Chinese logograph reading. *Neuroimage* 13, 836–846 (2001).
- Friederici AD, Opitz B and von Cramon DY. Segregating semantic and syntactic aspects of processing in the human brain: An fMRI investigation of different word types. *Cerebr Cortex* 10, 698–705 (2000).
- Shaywitz BA, Shaywitz SE, Pugh KR, Fullbright RK, Skudlarski P, Mencl WE *et al.* The functional neural architecture of components of attention in language-processing tasks. *Neuroimage* 13, 601–612 (2001).
- Shulman GL, d'Avossa G, Tansy AP and Corbetta M. Two attentional processes in the parietal lobe. *Cerebr Cortex* 12, 1124–1131 (2002).
- Siok WT, Jin Z, Fletcher P and Tan LH. Distinct brain regions associated with syllable and phoneme. *Hum Brain Mapp* 18, 201–207 (2003).
- Tan LH, Spinks JA, Feng CM, Siok WT, Perfetti CA, Xiong J *et al.* Neural systems of second language reading are shaped by native language. *Hum Brain Mapp* 18, 158–166 (2003).
- Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9, 97–113 (1971).
- Baddeley A. *Working Memory*. New York: Clarendon Press; 1986.
- Miller G. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 63, 81–97 (1956).
- Mills CB and Martin JG. Articulatory organization in digit perception and recall. *Percept Psychophys* 21, 463–468 (1977).
- Lowe MJ and Sorenson JA. Spatially filtering functional magnetic resonance imaging data. *Magn Reson Med* 37, 723–729 (1997).
- Cox RW. Afni: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29, 162–173 (1996).
- Talairach J and Tournoux P. *Co-planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System: An Approach to Cerebral Imaging*. New York: Thieme Medical Publishers; 1988.
- Zatorre RJ, Meyer E, Gjedde A and Evans AC. PET studies of phonetic processing of speech: review, replication, and reanalysis. *Cerebr Cortex* 6, 21–30 (1996).
- Gandour J, Wong D, Lowe M, Dziedzic M, Sathamnuwong N, Tong Y and Li X. A cross-linguistic fMRI study of spectral and temporal cues underlying phonological processing. *J Cogn Neurosci* 14, 1076–1087 (2002).
- Gelfand JR and Bookheimer SY. Dissociating neural mechanisms of temporal sequencing and processing phonemes. *Neuron* 38, 831–842 (2003).
- Brass M and von Cramon DY. The role of the frontal cortex in task preparation. *Cerebr Cortex* 12, 908–914 (2002).
- Corbetta M, Kincade JM, Ollinger JM, McAvoy MP and Shulman GL. Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neurosci* 3, 292–297 (2000).
- Corbetta M and Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nature Rev Neurosci* 3, 201–215 (2002).
- Mazoyer P, Wicker B and Fonlupt P. A neural network elicited by parametric manipulation of the attention load. *Neuroreport* 13, 2331–2334 (2002).
- Tata MS, Prime DJ, McDonald JJ and Ward LM. Transient spatial attention modulates distinct components of the auditory ERP. *Neuroreport* 12, 3679–3682 (2001).
- Crosson B, Sadek JR, Maron L, Gokcay D, Mohr CM, Auerbach EJ *et al.* Relative shift in activity from medial to lateral frontal cortex during internally versus externally guided word generation. *J Cogn Neurosci* 13, 272–283 (2001).
- Jonides J, Schumacher EH, Smith EE, Koeppel RA, Awh E, Reuter-Lorenz PA *et al.* The role of parietal cortex in verbal working memory. *J Neurosci* 18, 5026–5034 (1998).

Acknowledgements: We are grateful to T. Osborn and J. Zimmerman for their technical assistance in the MRI laboratory. M.L. is currently affiliated with the Cleveland Clinic Foundation, Mario Dziedzic with MDZ Consulting Inc. Supported by NIH research grant R01 DC04584-04 (J.G.); NIH postdoctoral traineeship (X.L.).