Distinct Brain Regions Associated With Syllable and Phoneme

Wai Ting Siok,^{1*} Zhen Jin,² Paul Fletcher,³ and Li Hai Tan⁴

¹*Department of Psychology, Stanford University, Stanford, California*

²*Beijing 306 Hospital, Beijing, People's Republic of China*

³*Division of Speech and Hearing Sciences, The University of Hong Kong, Hong Kong*

⁴*Joint Laboratories for Language and Cognitive Neuroscience,
The University of Hong Kong, Hong Kong*

Abstract: The syllable and the phoneme are two important units in the phonological structure of speech sounds. In the brain mapping literature, it remains unsolved as to whether there are separate brain regions mediating the processing of syllables and phonemes. Using functional magnetic resonance imaging (fMRI), we investigated the neural substrate of these phonological units with Chinese subjects. Results revealed that the left middle frontal cortex contributes to syllabic processing, whereas the left inferior prefrontal gyri contributes to phonemic processing. This pattern of findings offers compelling evidence for distinct cortical areas relevant to the representation of syllables and phonemes. *Hum. Brain Mapping* 18: 201–207, 2003. © 2003 Wiley-Liss, Inc.

Key words: fMRI; phonology; syllables and phonemes; phonological structure of speech sounds; neural basis for phonological processing

INTRODUCTION

In speech and linguistic sciences, it has been established that the phonological structure of speech sounds is hierarchically organized. The syllable is a basic functional unit and enjoys a particular top-of-the-hierarchy status. The phoneme is a minimal speech unit, represented by alphabetic letters in English. Between the syllable and the phoneme, there

exists an intermediate layer of representation corresponding to the intrasyllabic units termed onsets and rimes. Previous cognitive research showed that the accessibility of a phonological unit is determined by its position within the hierarchy, with syllable processing being easier than the processing of phonemes and other intrasyllabic units [Goswami and Bryant, 1990; Treiman, 1992].

Are there distinct cortical areas associated with the processing of syllables and phonemes? While recent brain imaging research with positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) has yielded important insights into the brain mechanisms underlying phonological processing in general, this important question remains far from clear. Studies with English and other alphabetic languages found that the superior temporal sulcus and/or the left inferior frontal cortex subserve phonological knowledge of language stimuli in syllable dis-

Contract grant sponsor: Hong Kong Research Grants Council; Contract grant number: HKU 7133/01H.

*Correspondence to: Wai Ting Siok, Ph.D., Department of Psychology, Stanford University, Jordan Hall Bldg. 420, Stanford CA 94305. E-mail: wtsiok@psych.stanford.edu

Received 30 August 2002; Accepted 12 October 2002

DOI 10.1002/hbm.10094

Figure 1.

Examples of experimental stimuli used in the experiment. **a:** The subject judged whether the two Chinese characters were homophones. Both characters are pronounced /mi2/. **b:** The subject judged whether the two Chinese characters shared the same initial consonant. The character above the fixation crosshair is pronounced /gan1/, and the character below the fixation crosshair is pronounced /gou1/. The two characters exposed in each pair had the same tone but shared no visual nor semantic similarity.

crimination [Gabrieli et al., 1998; Poldrack et al., 1999; Price et al., 1997; Rosen, 1992; Wise et al., 2001], monosyllabic word perception [e.g., Binder et al., 2000; Fiez et al., 1996], rhyme judgments [Booth et al., 2002; Lurito et al., 2000; Poldrack et al., 2001; Pugh et al., 1996; Xu et al., 2001], and phoneme monitoring [Demonet et al., 1992; Gandour et al., 2000; Newman and Twieg, 2001; Thierry et al., 1999; Zatorre et al., 1992, 1996]. Nevertheless, these studies have not directly compared syllabic with phonemic processing. This is of importance since the nature of alphabetic systems means that phonemic processing is almost always involved in rhyme and syllable decision tasks [see Poldrack et al., 1999; Zatorre et al., 1996], although different experimental tasks may place differential weighting on syllabic and phonemic level processing. Such task differences may be a critical reason why different cortical areas for phonological processing have been implicated in different studies.

Written Chinese offers a unique opportunity in disentangling different types of phonological processes. As a logographic system, written Chinese uses characters that have a square shape as a basic writing unit. The Chinese character's phonology is defined at the monosyllabic level, with no parts in the character corresponding to phonological segments such as phonemes. Thus, each character has a monosyllabic pronunciation, and the computation of visual form into phonological codes based on letter-to-sound or grapheme-to-phoneme correspondence rules is impossible in Chinese [Perfetti and Tan, 1998]. However, the monosyllabic pronunciation of the character is segmentable indeed. For instance, the character 他 is pronounced /ta1/ (the numeral following the syllable

refers to tone), meaning *he*. The syllabic unit /ta/ can be decomposed into /t/ and /a/, two phonemes. Because in this character, no componential unit sounds as /t/ or /a/, the character's syllabic level phonology is always activated as a whole, without involvement of phonemic units. Phonological segmentation of Chinese characters is possible, if that is required in an explicit task; but fine-grained phonological units such as phonemes have to be activated following syllabic level phonological computation.

In this fMRI study, we required native Chinese speakers to perform two phonological tasks in order to ascertain whether separate cortical regions are responsible for syllabic and phonemic processing. In a homophone decision task, subjects judged whether two Chinese characters (i.e., single-character words) exposed synchronously were homophones. In Chinese, homophone judgments are made based exclusively on syllable level information, without involvement of phonemic information. A second phonological task was used to examine phonemic processing. This was done by getting subjects to decide whether a pair of characters exposed synchronously shared the same beginning consonant. To perform this task, subjects had to first access the two characters' syllable level phonological information, then segment syllabic units into initial phonemes. To control for the effects of word familiarity on cognitive and neuronal activities [Fiez et al., 1999], the stimuli used in our experiment were matched across frequencies of occurrences in the two tasks. In addition, the two characters exposed in each pair shared no visuo-orthographic similarity. This meant that subjects had to make their decisions based on a character's phonological attributes rather than on their visual properties. Each of the two phonological tasks alternated with a visual control task in which the subject made a judgment on the font size of two characters. Figure 1 shows examples of stimuli used in our study.

SUBJECTS AND METHODS

Subjects

Eleven male volunteers participated. All subjects were native Chinese (Mandarin) speakers, ranging in age from 19 to 21 years (mean age 20.6 years, SD 1.3 years). They were strongly right handed as judged by the handedness inventory devised by Snyder and Harris [1993]. Informed consent was obtained from each subject prior to testing.

Apparatus and procedure

MRI experiments were performed with a 2 T GE/Elscint Prestige MRI scanner at Beijing 306 Hospital. Visual stimuli were presented to subjects by way of a projector onto a translucent screen. Subjects viewed the stimuli through a mirror attached to the head coil. A T_2^* -weighted gradient-echo EPI sequence was used for fMRI scans, with the slice thickness = 6 mm, in-plane resolution = 2.9×2.9 mm, TR/TE/flip angle = 2,000 msec/45 msec/90 degrees. Eighteen contiguous axial slices parallel to the AC-PC line covering the whole brain were acquired while the subject performed one of the three tasks. The high resolution ($2 \times 2 \times 2$ mm³) anatomical brain MR images were acquired using a T_1 -weighted, 3-D gradient-echo sequence.

Each phonological condition consisted of 30 pairs of Chinese characters that were either homophones or shared the same initial consonant and 30 pairs of Chinese characters that were unrelated (i.e., shared no orthographic features, phonology or meaning). The baseline condition consisted of 54 pairs of unrelated Chinese characters. The individual characters in 27 pairs had the same font size; in the other 27 pairs, the individual characters were of a different size. All characters were commonly used and had a frequency of occurrences greater than 20 per million according to the *Modern Chinese Frequency Dictionary* [1986]. Visual complexity of the characters was matched across all conditions.

A block design was used, with three blocks of each of the two phonological conditions and six blocks of the baseline condition. Each phonological block consisted of an instruction and 20 stimulus pairs. Equal number of pairs were the same and different. Each baseline block consisted of an instruction and nine stimulus pairs. In each trial, a pair of characters was exposed synchronously for 1,000 msec, one above and one below a fixation crosshair, followed by a 1,000-msec blank time. Subjects indicated a positive response by pressing the key corresponding to the index finger of their right (dominant) hand. They were asked to perform the tasks as quickly and accurately as possible.

Data analysis

SPM99 software (Wellcome Department of Cognitive Neurology, UK) was used for image preprocessing and statistical analyses [Friston et al., 1995]. The functional images were realigned to the first functional volume using sinc interpolation. They were

then spatially normalized to an EPI template based on the ICBM152 stereotactic space, an approximation of canonical space [Talairach and Tournoux, 1988]. Images were resampled into $2 \times 2 \times 2$ mm cubic voxels and spatially smoothed with an isotropic Gaussian kernel (6 mm full width at half-maximum). Statistical analyses were calculated on the smoothed data using a delayed boxcar design with a 6-sec delay from onset of block to account for the transit effects of hemodynamic responses. For each subject, adjusted mean images were created for each condition after removing global signal and low-frequency covariates. Activation maps were calculated by comparing images acquired during each task state (homophone and initial consonant judgements) with those acquired during the control state (font size judgement), using a students' group *t*-test. Activations that fell within clusters of 20 or more contiguous voxels exceeding the corrected statistical threshold of $P < 0.05$ were considered significant. Conjunction analysis was performed to find out common areas activated by both tasks. Direct comparisons of initial-consonant with homophone judgement were made to locate cortical regions that were involved in phonemic processing. We constrained the analysis to two structurally defined region-of-interest (ROI): left frontal cortex and left temporal cortex, regions reported to be involved in phonological processing in alphabetic languages. The small-volume statistical correction of Worsley et al. [1996] was applied to the analysis.

RESULTS

The statistical parametric maps averaged across subjects for the two phonological conditions (homophone and initial-consonant judgments) compared to the font size judgments are shown in Figure 2. For homophone judgments that required only syllable level phonology, peak activation was located in the left middle frontal cortex (BA 9), with an activation volume of 351 voxels (Table I). Left superior frontal gyrus (BA 8) and cingulate gyrus (BA 32) were also provoked. For initial-consonant judgments that demanded both syllabic as well as phonemic processing, left middle frontal cortex at BA 9 was active, but peak activation in this task was shifted inferiorly to the left inferior frontal cortex (BAs 44, 45, 46), with an activation volume of 941 voxels. Initial-consonant decisions also provoked cortical activity at left medial frontal gyrus (BA 6/8/9).

We performed conjunction analysis across the homophone and initial-consonant judgment scans [Price and Friston, 1997] using SPM99 with a height thresh-

(a) Homophone Judgment – Baseline

(b) Initial Consonant Judgment – Baseline

(c) Conjunction Analysis

Figure 2.

Functional maps. Averaged brain activations involved in the homophone and initial consonant judgment tasks. Normalized activation brain maps averaged across 11 subjects demonstrate the statistically significant activations ($P < 0.05$, corrected and at least 20 contiguous voxels) in homophone vs. font size judgment (a),

initial consonant vs. font size judgment (b), and conjunction analysis across the two experimental scans (c). Planes are axial sections, labeled with the height (mm) relative to the bicommissural line. The left side corresponds to the left hemisphere of the brain.

old of $P = 0.05$ (corrected). The conjunction results indicated that the left middle frontal gyrus mediated both homophone and initial-consonant decision (Fig. 2c).

Direct comparisons of the two phonological conditions using small volume corrections indicated that activity for initial-consonant judgment was larger than that for homophone judgment in left inferior frontal gyrus (BA 44/45/46, cluster size = 465 voxels, corrected cluster level $P < 0.05$). No significant difference was found in left temporal cortex.

DISCUSSION

In our fMRI study, subjects were required to segment two Chinese words' monosyllabic phonology so

that they could decide whether they have the same initial consonant. Compared to homophone judgments based solely on monosyllabic phonology without involvement of phonemic coding, initial consonant identity judgments peaked in the left inferior frontal gyrus. This result suggests that phonemic processing of word initial consonants is mediated by the left inferior frontal cortex. This finding is in harmony with that obtained from brain mapping studies with English. For example, Zatorre et al. [1992, 1996] asked subjects to discriminate final consonants in auditorily exposed lexical and non-lexical CVC items (e.g., *fat-tid*). Compared to a pitch-discrimination task, the phoneme discrimination task provoked left inferofrontal activation. Interestingly, this area was not active when subjects were required to passively listen to the same

TABLE I. Regions of significant activation relative to baseline condition

Regions activated	Homophone vs. font size judgment						Initial consonant vs. font size judgment					
	No. of voxels	BA	Coordinates			Z score	No. of voxels	BA	Coordinates			Z score
			<i>x</i>	<i>y</i>	<i>z</i>				<i>x</i>	<i>y</i>	<i>z</i>	
L inferior frontal gyrus							676	45/46	−50	30	8	8.10
							265	44/45	−48	11	18	5.87
L middle frontal gyrus	351	9	−51	15	34	7.39	476	9	−50	15	27	6.68
	58	6	−32	7	59	6.59	62	8/9	−46	12	42	5.69
L superior frontal gyrus	21	8	−2	22	54	5.19						
L medial frontal gyrus							140	6/8	−2	22	43	5.61
							79	9	−2	39	33	5.27
Cingulate gyrus	121	32	−4	31	30	5.90						

stimuli. Zatorre et al. proposed that the dorsal portion of the left inferior frontal region mediates explicit fine-grained phonemic analysis. Because this sort of phonemic analysis demands more working memory resources compared to the analysis of larger phonological units, it has been proposed that left inferior frontal region is more important for effortful processes such as retrieval, manipulation, or selection of phonological representations [Fiez et al., 1999]. Our and others' neuroimaging results have lent strong support to these proposals [see also Booth et al., 2002; Gandour et al., 2000; Lurito et al., 2000; Newman and Twieg, 2001; Xu et al., 2001].

More importantly, we discovered that when our experimental tasks required syllabic level phonological information, the left middle frontal cortex was strongly activated. The left middle frontal cortex even peaked in our homophone decision task that demanded syllabic phonology without participation of any phonemic units. Also notable is that the left inferior frontal gyrus failed to show above-threshold activity in homophone decisions that did not call for phonemic contribution. These findings provide compelling evidence that indicates that the left middle frontal cortex is responsible for syllables rather than phonemes.

In the literature, the contribution of the left middle frontal cortex to phonological computation of English words (exposed auditorily or visually) has not been well identified. Two published fMRI studies, however, are highly relevant here. Poldrack et al. [1999] asked subjects to perform a syllable counting task on printed nonwords, a semantic decision task on words (abstract/concrete), and a perceptual decision task (uppercase/lowercase decision). Compared to the perceptual decision, syllable counting activated the dorsal portion of inferior frontal cortex bilaterally.

Relative to the semantic task, however, the syllable counting task elicited peak activation of the left middle frontal cortex at BA 9 and 46. In a rhyme judgment task, Xu et al. [2001] found that rhyming words and pseudowords revealed strong activation in the left posterior mid and inferior prefrontal gyri (BA 9 and Broca's area) and left inferior occipital-temporal junction. The involvement of the left middle frontal cortex in these two studies indicated that their subjects might tend to activate the syllable level phonology of English words and nonwords and strategically rely less on phonemic coding carried out by letter-sound conversions.

Nonetheless, the central role of the left middle frontal cortex in language processing has been repeatedly demonstrated in fMRI studies with Chinese. In particular, peak activity of the left middle frontal area has been obtained in word generation [Tan et al., 2000], semantic judgment [Tan et al., 2001], homophone judgment (compared to fixation) [Tan et al., 2001], and rhyme decision (see Tan et al., pp. 158–166, this issue). Our present finding obtained with our Beijing MRI scanner converges on past results from the San Antonio MRI scanner. Elsewhere, we have hypothesized that the extremely strong activation seen in left middle frontal area in reading Chinese is associated with the coordination of the phonological processing of the Chinese logographs that was explicitly required by the experimental task and the elaborated analyses of the visual-spatial locations of the strokes in a logograph that was demanded by its unique square configuration (see Tan et al., 2003). In other words, the cortical organization of phonological knowledge of written words reflects associations with visual features and associations with linguistic attributes (e.g., orthographic). Our present findings further indicated

that the monosyllabic nature of Chinese results in the strong activation of left middle frontal cortex.

Our finding may also be general across languages. Combined with the findings of Poldrack et al. [1999] and Xu et al. [2001], we may conclude that the left middle frontal cortex contributes to the processing of syllables, while the left inferior frontal gyrus mediates the processing of phonemes and phonological segmentation.

CONCLUSIONS

The syllable as a basic functional unit enjoys a particular linguistic status in the phonological structure of speech sounds. The syllabic level representation in the cognitive network is important both for language comprehension and language production. For instance, it has been hypothesized that syllable-level phonological encoding is an important stage in speech production, which then allows a phonetic plan to be assembled [Levelt, 1993]. Despite the significance of the syllable in language cognition and cognitive neuroscience of language, its neural basis remains unknown. Our fMRI study reported in this article has generated important results of human cortical regions associated with the processing of syllable, which are distinct from brain areas associated with the processing of phonemes. In particular, we have found that the left middle frontal cortex mediates syllabic, rather than phonemic, level phonological processing. Phonemic computation and phonological segmentation, on the other hand, are subserved by the left inferior prefrontal gyri. Whether these findings from the phonological processing of Chinese words are generalizable to English and other languages merits further investigation.

ACKNOWLEDGMENTS

We are grateful to Dr. Benjamin Xu for his helpful comments.

REFERENCES

- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET (2000): Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex* 10:512–528.
- Booth JR, Burman DD, Meyer JR, Gitelman DR, Parrish TB, Mesulam MM (2002): Modality independence of word comprehension. *Hum Brain Mapp* 16:251–261.
- Demonet JF, Chollet F, Ramsay S, Cardebat D, Nespoulous JL, Wise R, Rascol A, Frackowiak R (1992): The anatomy of phonological and semantic processing in normal subjects. *Brain* 115:1753–1768.
- Fiez JA, Raichle ME, Balota DA, Tallal P, Petersen SE (1996): PET activation of posterior temporal regions during auditory word presentation and verb generation. *Cereb Cortex* 6:1–10.
- Fiez, J. A., Balota, D. A., Raichle, M. E., Petersen, S. E (1999): Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron* 24:205–218.
- Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak, RSJ (1995): Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp* 2:189–210.
- Gabrieli JDE, Brewer JB, Poldrack RA (1998): Images of medial temporal lobe functions in human learning and memory. *Neurobiol Learn Mem* 70:275–283.
- Gandour J, Wong D, Hsieh L, Weinzaepfel B, Van Lancker D, Hutchins GD (2000): A crosslinguistic PET study of tone perception. *J Cogn Neurosci* 12:207–222.
- Goswami U, Bryant P (1990): Phonological skills and learning to read. London: Erlbaum.
- Levelt WJ (1993): Timing in speech production with special reference to word form encoding. In: Tallal P, Galaburda AM, editors. *Temporal information processing in the nervous system: Special reference to dyslexia and dysphasia*. New York: Academy of Sciences; p 283–295.
- Lurito JT, Kareken DA, Lowe MJ, Chen SH, Mathews VP (2000): Comparison of rhyming and word generation with fMRI. *Hum Brain Mapp* 10:99–106.
- Modern Chinese Frequency Dictionary (1986): Beijing, China: Beijing Language Institute.
- Newman SD, Twieg D (2001): Differences in auditory processing of words and pseudowords: an fMRI study. *Hum Brain Mapp* 14:39–47.
- Perfetti CA, Tan LH (1998): The time course of graphic, phonological, and semantic activation in visual Chinese character identification. *J Exp Psychol Learn Mem Cogn* 24:101–118.
- Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JE (1999): Functional specialization for semantic and phonological processing in the left inferior frontal cortex. *Neuroimage* 10:15–35.
- Poldrack RA, Temple E, Protopapas A, Nagarajan S, Tallal P, Merzenich M, Gabrieli JD (2001): Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI. *J Cogn Neurosci* 13:687–697.
- Price CJ, Friston KJ (1997): Cognitive function: a new approach to brain activation experiments. *Neuroimage* 5:261–270.
- Price CJ, Moore CJ, Humphreys, GW, Wise, RJS (1997): Segregating semantic from phonological processes during reading. *J Cogn Neurosci* 9:727–733.
- Pugh KR, Shaywitz BA, Shaywitz SE, Constable RT, Skudlarski P, Fulbright RK, Bronen RA, Shankweiler DP, Katz L, Fletcher JM, Gore JC (1996): Cerebral organization of component processes in reading. *Brain* 119:1221–1238.
- Rosen S (1992): Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci* 336: 367–373.
- Snyder PJ, Harris LJ (1993): Handedness, sex and familiar sinistrality effects on spatial tasks. *Cortex* 29:115–134.
- Talairach J, Tournoux P (1988): Co-planar stereotaxic atlas of the human brain. New York: Theime Medical.
- Tan LH, Spinks JA, Gao JH, Liu HL, Perfetti CA, Xiong J, Stofer KA, Pu Y, Liu Y, Fox PT (2000): Brain activation in the processing of Chinese characters and words: a functional MRI study. *Hum Brain Mapp* 10:16–27.

- Tan LH, Liu HL, Perfetti CA, Spinks JA, Fox PT, Gao JH (2001): The neural system underlying Chinese logograph reading. *Neuro-Image* 13:836–846.
- Tan LH, Spinks JA, Feng C-M, Siok WT, Perfetti CA, Xiong J, Fox PT, Gao J-H (2003): Neural systems of second language reading are shaped by native language. *Hum Brain Mapp* 18:158–166.
- Thierry G, Boulanouar K, Kherif F, Ranjeva JP, Demonet JF (1999): Temporal sorting of neural components underlying phonological processing. *Neuroreport* 10:2599–2603.
- Treiman R (1992): The role of intrasyllabic units in learning to read and spell. In: Gough PB, Ehri LC, editors. *Reading acquisition*. New Jersey: Erlbaum; p 65–106.
- Wise RJ, Scott SK, Blank SC, Mummery CJ, Murphy K, Warburton EA (2001): Separate neural subsystems within ‘Wernicke’s area’. *Brain* 124:83–95.
- Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC (1996): A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Mapp* 4:58–83.
- Xu B, Grafman J, Gaillard WD, Ishii K, Vega-Bermudez F, Pietrini P, Reeves-Tyer P, DiCamillo P, Theodore W (2001): Conjoint and extended neural networks for the computation of speech codes: The neural basis of selective impairment in reading words and pseudowords. *Cereb Cortex* 3:267–277.
- Zatorre RJ, Evans AC, Meyer E, Gjedde A (1992): Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256:846–849.
- Zatorre RJ, Meyer E, Gjedde A, Evans AC (1996): PET studies of phonetic processing of speech: Review, replication, and reanalysis. *Cereb Cortex* 6:21–30.