Investigation of relation between speech perception and production based on EEG source reconstruction

Guancheng Li[*], Jianwu Dang[*,†],Gaoyan Zhang[*], Zhilei Liu[*], Hongcui Wang[*]
[*]The Tianjin Key Laboratory of Cognitive Computing and Application, Tianjin University, Tianjin, China
E-mail: guanchengli@tju.edu.cn, zhanggaoyan2013@163.com, zhileiliu@tju.edu.cn, hcwang@tju.edu.cn Tel: +86-22-27407757
[†]Japan Advanced Institute of Science and Technology, Japan
E-mail: jdang@jaist.ac.jp Tel: +81-07-61511235

Abstract— **Mirror neuron system has been investigated using the functional magnetic resonance imaging (fMRI) technique. Activation of the Broca's area and the premotor cortex (PMC), which related with speech production, were observed during speech perception, and seems to be a mirror. However, it is not clear how the mirror neurons function between speech production and perception. This study attempts to investigate the functions of the mirror neurons by utilizing the high temporal resolution of the Electroencephalography (EEG) system. The participants watched Chinese material from screen then heard the material reading from an earphone, finally made a judgement about the consistency of the two stimuli. The high-density EEG signal under source reconstruction revealed that the Wernicke's area activated before the Broca's area and PMC during the speech perception tasks. Results are also consistent with the mirror neuron system: the speech production related regions are working during the speech perception tasks.**

I. INTRODUCTION

The mirror neurons have been defined as the neurons that activated both when an animal acts and when the animal observes the same action performed by anothers [1]. Such neurons have been directly observed in primate species [2]. The Broca's area and the premotor cortex (PMC) activated in both speech perception and production tasks demonstrated that there exist mirror neurons in human brain [3].

Functional magnetic resonance imaging (fMRI) with a higher spatial resolution has been used to investigate brain functions. Besides PMC/Broca's areas, previous fMRI studies suggested that superior temporal gyrus/sulcus (STG/S) also involved in the speech cognition tasks [4]. However, due to the lower temporal resolution, it is difficult to obtain quantitative evaluation for such relations. To fully understand brain functions in speech production and perception, we should investigate the functions in both spatial and temporal aspects to exclude the potential factors such as task difficulty, greater attentional demand, languages and parts of speech [4]. Therefore, an Electroencephalography (EEG) experiment with high temporal resolution is benefit for this purpose.

The experiments which prove the mirror property of the neurons mostly use the syllables or alphabetic words as the stimuli. Tan has proved that the Chinese words include abundant information from its pictographic characters works in different pattern to some degree compare with the alphabetic words in his experiments[4], [5]. In this preliminary study, we use the Chinese words as the stimuli and mainly focus on the temporal relation of activations between the related areas. To ensure the spatial resolution, 128-electrode high density cap was used for the data collection. In order to induce stronger neural activity, verbs were chose as the stimuli material and they were demonstrated in multiple sensory channels [6], [7]. During the experiment, the participants were asked to finish a task, to ensure them attend conduct experiment under a great attentional condition, and low correct ratio data will be discarded. We obtained the results from the source reconstruction process and it produced us a continuous special change of the activated regions that contribute to explain the neural control of speech perception.

II. EEG EXPERIMENT

A. Participants

Sixteen subjects from Tianjin University (8 female, and 8 male subjects) aged from 21 to 28 (mean =23.9 years; standard deviation = 1.59 years) participated in this experiment. All subjects were native Chinese speakers with a normal or corrected vision, and without any history of speaking, hearing, or motor disorders. Every subject has undergone the Edinburgh Handedness Inventory [8] and was identified to be right-handed. The study was conducted in accordance with the Declaration of Helsinki [9]. All subjects gave their written informed consent and were paid for their participation.

B. Procedure

The subjects were asked to finish tasks in this experiment and in each trial there were following six steps: (1) one trial started with a white cross (+) lasting for 200ms to 300ms in the center of the monitor with black background (see Fig. 1); (2) then a text with two Chinese characters was displayed on the monitor for 600ms; (3) keeping the monitor black for 400ms; (4) presenting reading sound via earphone, and the onset time point of this step would be labeled when recording signal; (5) the subject was asked to finish a task after (4); (6) finally, a 800ms to 1000ms blank for ending one this trial and moving to the next one.

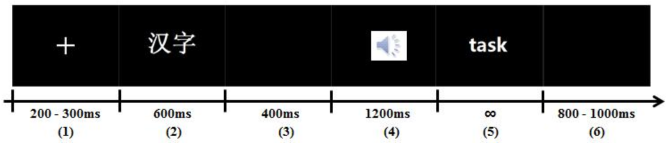

Fig. 1 A schematic of the experimental procedure.

The subjects were asked to judge the coincidence of characters in (2) and reading sound in (4) by pressing keyboard. They were asked to give response as quick as they can, which ensures that subjects maintain a high attention during previous steps. Based on the correct ratio, we discarded the data of the subjects if their attention level is not higher than 90%.

Before the experiment, they were asked to practice several times for better understanding the experiment. After number of trials, the subjects were allowed to take a rest for relax during the experiment.

C. Material

The auditory stimuli consisted of real speech sounds, which were read by a male announcer. The auditory stimuli signals were recorded in formalized into a single channel of 16 bit and sampled at 44.1 kHz with duration of 900ms, and were presented by a binaural earphone. The visual stimuli were Chinese characters displayed on the center of a 14inch monitor.

Material used in this study consisted of Chinese words and non-words. The Chinese words were composed of disyllabic verbs, and the non-words were made up by the juxtaposition of two legal characters, but had no meaning.

All kinds of the stimuli were matched in visual frequency and complexity: the Chinese words had a mean frequency of occurrence more than 14 per million according to the Modern Chinese Frequency Dictionary [10], while the average occurrence of one hundred words was 21.14 per million. To control the difficulties, the mean number of strokes for Chinese words was 18.03, and 18.34 for non-words.

Four conditions with 25 trials for each were included in the study. One hundred trials were carried out in a random order. As shown in Table I, the four conditions consist of two factors with two levels for each: coincidence of two stimuli in one trial and the material type in each trial. One trial was only used one kind of material. This means the subject would never watch non-word at (2) and hear word at (4), vice verse.

TABLE I
FOUR CONDITIONS IN EXPERIMENT

Condition	Word type	Coincidence
1	verb	true
2	verb	false
3	non-words	true
4	non-words	false

The ANOVA analysis has been conducted to ensure that four conditions of data without significant difference on the occurrence and stroke numbers. Since there was no occurrence for the non-words groups, so the ANOVA analysis for occurrence just has been conducted on the Chinese words and $F(1, 98) = 0.096$, $p > 0.1$. For the stroke numbers, $F(3, 166) = 0.116$, $p > 0.1$.

D. EEG Data Collection

The Neuroscan Synamps system with an electrode cap with 128 sintered Ag/AgCl electrodes was used for EEG signal recording. Six additional channels have been used for recording vertical electrooculography (VEOG) and horizontal electrooculography (HEOG). The VEOG includes two electrodes on the upside (VEOU) and downside (VEOL) of the left eye, and two symmetrical electrodes on the right eye (85 on the upside and 84 on the downside), while the HEOG includes two electrodes on the outside of both eyes (HEOL and HEOR). The impedance of each electrode was kept below 5 kΩ. The EEG signal was recorded with a sampling rate of 1000 Hz and processed using a band-pass filter with a range of 0.01–100 Hz.

E. Data Analysis

There were three steps for the data analysis part: signal filtering, preprocessing and source reconstruction.

In the signal filtering step, the EEG data was subjected to band-pass filtering (0.1-30Hz). The interval of signals for target stimulus to be calculated was 200ms before to 1000ms after the stimulus onset. The interval from -200ms to 0ms was taken as the baseline. After this step, we obtained four averaged signal files under four conditions from each subject.

In the preprocessing step, the signals under same condition from different subjects were averaged again. Then the signals from 128 channels were stacked as butterfly plots for searching the interval with strong neural actions. The Mean Global Field Power (MGFP) were calculated, the peak interval indicated the strong neural action. Before the Principal Component Analysis (PCA) and Independent Component Analysis (ICA) processing, noise estimation on the interval from -200ms to 0ms was taken to make it possible to transform data from their μV / fT units to Signal Noise Ratios (SNRs). The signal at each condition was decomposed into signal and noise subspaces using the PCA decomposition to remove the noise. After PCA, the ICA was conducted to extract the dominant EEG pattern. We applied the PCA/ICA in Curry 6.0, and the first two or three largest principal components values with the SNR higher than 1.0 dB were suggested to be useful for source reconstruction.

In the source reconstruction part, the Boundary Element Model (BEM) consisting of three compartments was built. The compartments used were skin, skull and brain tissue. The method used to conduct the source reconstruction was a Current Density Reconstructions (CDR) based solution, called Standardized Low Resolution Electromagnetic Tomography (sLORETA), which was widely used in source reconstruction.

III. RESULTS AND DISCUSSION

A. Behavioral Results

We have collected the response latency of each subject after finishing the task, and calculated the correct ratio. The averaged behavioral data collected from the 12 subjects is shown in Table II (Four subjects' data has been discarded, the electrode detached during experiment for two of them and the rest two was discarded for low attention level).

TABLE II
BEHAVIORAL DATA UNDER FOUR CONDITIONS

Condition	Response Latency (ms)	Correct Ratio
1	382.58	0.992
2	451.11	0.997
3	402.58	0.99
4	453.46	1

The average response latencies indicate that, the coincide answers (Conditions 1 and 3) takes less time than the inconsistent ones, and the word conditions (Conditions 1 and 2) takes less time slightly than the non-word conditions. The correct ratio implies that the task difficulty has been kept on a compatible level, and all the subjects have finished the task with great concentration, it verifies the reliability of the signal. Moreover, the standard deviation of the correct ratio is about 0.005, which indicates that the difficulty under the four conditions have no significant difference.

B. Current Density Reconstructions based Results

The peak of MGFP is found to take place in the interval from 100ms to 200ms under all four conditions. Accordingly the current density reconstructions (CDR) based analysis is carried out in the interval.

The CDR results are shown in Table III and Fig. 2. Table III lists the areas' functional region name with mass activity at dominant hemisphere (left) for sLORETA solutions. Fig. 2 (A) shows the left view of source activity at each time point for condition 2 as projected to a representation of the cortical surface, the color bar shows the F-distribute value indicates the strength of the activation. Fig. 2 (B) shows the detailed source activity under condition 2 from 121ms to 130ms.

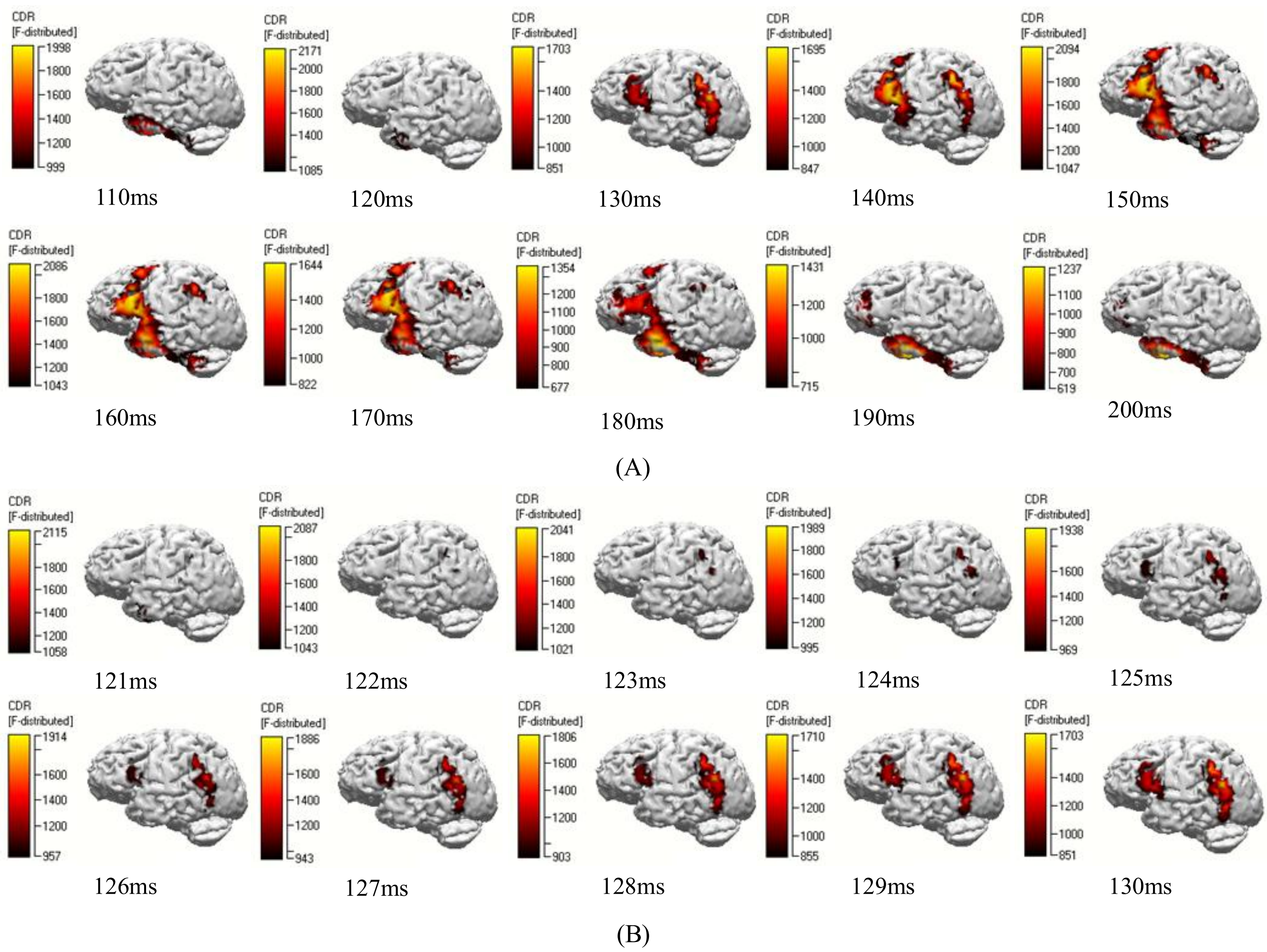

Fig. 2 A time sequence of the left view under condition 2.

TABLE III
ACTIVATED AREAS UNDER EACH CONDITION

(A) 110-150ms

Conditions \ Time Point(ms)	110	120	130	140	150
2			MTC, AG, SG, PO, PT, MFG	PMC, MTC, STC, AG, SG, PO, PT, MFG	PMC, MTC, STC, SG, PO, PT, MFG
3	ITC, MTC, STC	ITC, MTC, STC, OG	ITC, MTC, STC, SG	ITC, MTC ,STC, AG, SG	ITC, MTC, ITC, SG
4	ITC, MTC, STC	ITC, MTC, STC	ITC, MTC, STC, PO, PT	PFC, ITC, MTC, STC, PO, PT	PFC, ITC, MTC, STC, PO, PT

(B) 160-200ms

Conditions \ Time Point(ms)	160	170	180	190	200
2	PMC, MTC, STC, SG, PO, PT, MFG	PMC, MTC, STC, SG, PO, PT, MFG	PMC, MTC ,STC, PO, PT, MFG	PFC, MFG	
3	ITC, MTC, STC, SG	ITC, MTC, STC, SG	ITC, MTC, STC, SG	ITC, MTC, STC, SG	ITC, MTC, STC
4	ITC, MTC, STC, PO, PT	ITC, MTC, STC, PO, PT	ITC, MTC, STC, PO, PT, MFG	OG, ITC, MTC, STC, PO, PT, MFG	OG, ITC

Note.
The order in each cell consists with the Brodmann's Area number.
The functional regions' names in the table are list here:

MFG = Middle Frontal Gyrus, PFC = Prefrontal Cortex,
PMC = Premotor Cortex, PO = Pars Opercularis,
PT = Pars Triangularis, SG = Supramarginal Gyrus,
AG = Angular Gyrus, ITC = Inferior Temporal Cortex,
MTC = Middle Temporal Cortex, STC = Superior Temporal Cortex,
OG = Occipital Gyrus.

Fig. 2 (A) demonstrated an integrated activation regions' moving sequence under condition 2 (Subject watched verbs at (2) and heard different verbs at step (4)). At 130ms, we observed that the Wernicke's area near the Supramarginal Gyrus (SG) and the Broca's area at Brodmann 44 and 45 (BA 44 & 45) activated. Then activation near SG receded quickly after 140ms while the actvation near Broca's area expanded unward to the premotor cortex (PMC) and downward frontal part of the temporal gyrus including the Superior Temporal Cortex (STC), the Middle Temporal Cortex (MTC) and the Inferior Temporal Cortex (ITC). And after 180ms, activation at the frontal part of the brain quickly disappeared.

The sequence which is shown in Fig. 2 (A) consists with the hypothesis of the dual-stream model [9]. The dual-stream model considers that there are two routes working during speech perception: the dorsal route from the Wernicke's area, Broca's area and end the PMC; the ventral route from the Wernicke's area though the post temporal gyrus to frontal temporal gyrus. The dorsal route contributes to extract the phonological information and help with articulation or speech production, while the ventral route contributes to extract the meaning for understanding.

It is clear that the PMC/Broca's area working during speech perception, just like mirror neuron from the results. However, how does the phonological information flow through the dorsal route is not proved. So we conducted a detailed source reconstruction from 120ms to 130ms under condition 2 (See Fig. 2 (B)). It is shown that the Wernicke's area activated before the Broca's area which indicates that the phonological information comes from the Wernicke's area goes to the Broca's area.

For the condition 1 (Subject watched verbs at (2) and heard same verbs at (4)), we found the PMC/Broca's area did not activate as expected. We consider this may due to the visual stimuli (2) and the task design. The subjects have been watched the Chinese characters before listening. The Chinese characters originate from pictographic characters, which include not only semantic information but also phonological information. Previous studies have found that people understand the Chinese characters through two paths: the "orthographical information - semantic information" path, and the "orthographical information - phonological information - semantic information" path [12], [13]. Under condition 1, the subjects extracted the phonological information at (2), and when listening to the sound at (4), the matched phonological information may not rouse the mirror neuron again in a short time, so we observed it under condition 2 because (2) and (4) give different phonological information.

For the condition 3 (Subject watched non-words at (2) and heard same non-words at (4)), we also found that the PMC/Broca's area did not activate while these areas activated under condition 4 (Subject watched non-words at (2) and heard different non-words at (4)) without the activation of the Wernicke's area. For this result, we hold that under these two conditions, the non-words which lack of integrated semantic information may not activate the whole route like the verbs on account of that these non-words do not meet the pronunciation of the words used in usual. However, we still found that the Angular Gyrus (AG) and the Supramarginal Gyrus (SG) around the Wernicke's area activated under condition 3. This may cause by the single character in the non-words still works respectively. The corresponded stimulus at (4) strengthened phonological information for each character. And the dorsal route stopped here because this phonological information is irregular. At the same time, the stimulus at (2) did not strengthen phonological information at (4) under condition 4, so there is no massive activation around the Wernicke's area. The areas around the Broca's area like Prefrontal Cortex (PFC), Pars Opercularis (PO), Pars Triangularis (PT) activated. This may cause by the single characters in non-words at (2) roused ventral route partially.

A theory about the ventral network holds that the ventral route includes four major sub routes: the uncinate fasciculus (UF), the inferior fronto-occipital fasciculus (IFOF), the middle longitudinal fasciculus (MdLF) and the inferior longitudinal fasciculus (ILF)) [14]. The condition 3 seems consist with the MdLF which is considered to be the semantic and phonological language networks [15], [16], the matched characters and speech sound may go through this route. Moreover, the activated regions under condition 4 seems consist with the UF, a hook-shaped tract. It was proved to be existed in primates [17], [18]. Previous studies have proved that the illness on UF may cause some degree of semantic processing impairment. Other researchers found intraoperative stimulation of UF will cause semantic errors in picture naming so they believe that UF related to the ventral language pathway. Therefore the unmatched information at (2) and (4) may be likely to rouse this route.

IV. CONCLUSION

In the course of this study, it was visually apparent that the PMC and Broca's area participated in the speech perception procedure, which provided concrete evidence to the mirror neuron appeared in speech perception. The result from condition 2 also demonstrated that the dual-stream model activated at the interval of 100ms to 200ms. These high temporal resolution results will be contribute to illustrate the working mechanism of brain during the speech perception and production procedure. To achieve totally understanding of the working mechanism during speech perception and production procedure, more accurate data should be collected, and we will do more further research about this topic.

ACKNOWLEDGMENT

The research is supported partially by the National Basic Research Program of China (No. 2013CB329301), and the National Natural Science Foundation of China (No. 61303109). The study is supported partially by JSPS KAKENHI Grant (22500150).

REFERENCES

[1] Rizzolatti G., Craighero L., "The mirror-neuron system," Annu. Rev. Neurosci, vol. 27, pp. 169-192, 2004.

[2] Rizzolatti G., Fadiga L, Fogassi L., et al. "Resonance behaviors and mirror neurons," Archives italiennes de biologie, vol. 137(2), pp. 85-100, 1999.

[3] Callan D., Callan A., Gamez M., et al. "Premotor cortex mediates perceptual performance," Neuroimage, vol. 51(2), pp. 844-858, 2010.

[4] Callan D., Jones J. A., Callan A., "Multisensory and modality specific processing of visual speech in different regions of the premotor cortex," Frontiers in psychology, May 2014.

[5] Tan L H, Laird A R, Li K, et al. "Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta analysis" Human brain mapping, 2005, 25(1), pp. 83-91.

[6] Yang J, Tan L H, Li P. "Lexical representation of nouns and verbs in the late bilingual brain," Journal of Neurolinguistics, 2011, 24(6), pp. 674-682.

[7] Li P., Jin Z., Tan L. H., "Neural representations of nouns and verbs in Chinese: an fMRI study," NeuroImage, vol. 21(4), pp. 1533-1541, 2004.

[8] Caplan B., Mendoza J. E., Edinburgh Handedness Inventory, Springer New York, pp. 928-928, 2011.

[9] World Medical Association General Assembly, "World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects," Journal international de bioéthique, 2004.

[10] Wang, H., Modern Chinese frequency dictionary, Beijing Language Institute, 1986.

[11] Hickok G., Poeppel D., "The cortical organization of speech processing," Nature Reviews Neuroscience, vol. 8(5), pp. 393-402, 2007.

[12] Aaronson, D., Ferres, S., "A model for coding lexical categories during reading," Journal of Experimental Psychology: Human Perception and Performance, vol. 9, pp. 700–725.

[13] Liu B., Jin Z., Qing Z., et al., "The processing of phonological, orthographical, and lexical information of Chinese characters in sentence contexts: an ERP study," Brain research, vol. 1372, pp. 81-91, 2011.

[14] Bajada C. J., Ralph M. A. L., Cloutman L. L., "Transport for Language South of the Sylvian Fissure: The routes and history of the main tracts and stations in the ventral language network," Cortex, 2015.

[15] Makris N., Pandya D. N., "The extreme capsule in humans and rethinking of the language circuitry," Brain Structure and Function, vol. 213(3), pp. 343-358, 2009.

[16] De Champfleur, N. M., Maldonado, I. L., Moritz-Gasser, et al., "Middle longitudinal fasciculus delineation within language pathways: a diffusion tensor imaging study in human," European journal of radiology, vol. 82(1), pp. 151-157, 2013.

[17] Dejerine, Joseph Jules, Anatomie des centres nerveux. Rueff, 1895.

[18] Schmahmann, J. D., and Pandya, D, Fiber pathways of the brain. Oxford University Press, 2009.