Brain Mechanisms for Processing Phonemic and Allophonic Contrasts in Speech Sounds

Seiya FUNATSU, Satoshi IMAIZUMI, Akira HASHIZUME, Kaoru KURISU

1 INTRODUCTION

Acoustically discriminative variations in speech sounds, for instance, phones with harmonic structures versus phones without harmonic structures, usually evoke different percepts, such as voiced versus unvoiced consonants. However, some discriminative variations, such as Japanese voiced vowels versus devoiced vowels, evoke phonemically identical percepts. Interestingly, most Japanese speakers do not notice, or even cannot detect whether vowels are devoiced, although they have the ability to discriminate voiced versus unvoiced/devoiced vowels when isolated from their contexts. The present study aimed at uncovering the brain mechanisms responsible for this interesting phenomenon by analyzing magnetic mismatch fields (MMF) measured using magnetoencephalography (MEG).

The questions addressed in this paper are the following. 1) Is the allophonic brain response the same as the phonemic brain response? 2) If the allophonic response differs from the phonemic one, how is it different?

In Japanese, especially in Tokyo dialect, the high vowels [i] and [ɯ] tend to be pronounced as devoiced vowels [i̥] and [ɯ̥] with a high probability when surrounded by voiceless consonants. These allophonic variations do not affect word meaning. For instance, [sukijaki] and [sɯ̥kijaki] both mean /sukijaki/, although their acoustical and perceptual qualities are different.

There are many studies on the Japanese vowel devoicing with respect to phonetics, phonology and physiology. In these studies, however, there are many inconsistent results. In terms of acoustic studies, for instance, re the influence of pitch accent on devoicing, unaccented [i] and [ɯ] are more devoiced than accented [i] and [ɯ] [1, 2]. However, Maekawa [3] reported that the probability of devoicing does not relate to pitch accent. Moreover, as, for the consonantal environment, vowels preceded by fricatives are easier to devoice than those preceded by stops and affricates [1, 3]. On the other hand, Kuriyagawa and Sawashima showed that the probability of devoicing does not depend upon the types of consonants [2]. Fujimoto [4] in a photoelectric glottographic study reported that the vowel /i/ between fricatives was less devoiced than when the vowel /i/ was preceded or followed by a stop or an affricate. In the most recent study using a large-scale Japanese corpus (The Corpus of Spontaneous Japanese: CSJ), Maekawa and Kikuchi [5] concluded that in the case of vowel /u/, if a fricative is followed by a stop or affricate, vowel devoicing occurs most frequently (devoicing rate 97.5 % (stop case), 95.1 % (affricate case)), and in the case when an affricate is followed by a fricative, vowel devoicing is least likely to occur (48.1 %). With regard to speech rate, it has been confirmed that speech rate affects devoicing e.g. a fast speech rate is more likely to be associated with devoicing than a slow rate [1, 2, 6-8]. Concerning dialects, Kinki dialect (Osaka, Kyoto dialects), in contrast with Tokyo dialect, has a low devoicing rate, for instance, /kike/ is devoiced 97 % in Tokyo, but 42 % in Kinki [9]. In addition, the activity of laryngeal muscles during production of devoiced vowels was investigated using electromyograph and fiberscope [10-12]. Most of the investigations about vowel devoicing are about production, not perception or detection. A perception study by Cutler, Otake, and McQueen [13] showed that in nonsense words vowel devoicing makes speech processing difficult for Japanese speakers, but in meaningful words it does not make it difficult. Yet even fewer studies have investigated vowel devoicing in the field of brain sciences.

In the human body, brain activities generate electric currents in neurocytes in the brain. In addition, the electric currents in neurocytes induce a magnetic field around the head. We are thus able to determine the regions of brain activation by analyzing the source of this magnetic field. A magnetoencephalogram is obtainable by the measurement of the magnetic fields generated in the brain. The advanced features of MEG are as follows: (1) the information about higher order brain functions, cognition, and memory, can be obtained, (2) the measurement using MEG is a completely non-invasive technique, (3) the spatial localization of the current source can be estimated to within an order of a few millimeters, and (4) MEG has a 1 millisecond temporal resolution, making it far superior to other such devices.

The MMF, which is the magnetic equivalent of electric mismatch negativity (MMN) in electroencephalography (EEG), is generated by an event-related auditory neural process which automatically registers the stimulus difference or change. These mismatch responses (MMF and MMN) are evoked in an oddball paradigm, in which one standard stimulus is presented as frequent (probability: 80-90%) and one deviant stimulus is presented as rare (10-20%). In this oddball paradigm, memory traces are developed by the standard stimulus and one can detect the deviation with reference to these traces. The mismatch responses are not evoked by the deviant stimulus itself [14].

Numerous experiments conducted to date using tones have confirmed that mismatch responses are elicited by acoustic differences between standard stimulus and deviant stimulus in terms of frequency [15-17], amplitude [18, 19, 17], and duration [20, 17, 21]. A MEG study of the frequency mismatch response elicited an MMF due to 30 Hz difference between a 1030 Hz deviant tone and a 1000 Hz standard tone [15]. Sams et al. [16] in an EEG study used a 1000 Hz tone as a standard stimulus, with 1004 Hz and 1008 Hz tones as deviant stimuli. The 1008 Hz tone deviant elicited a small MMN, whereas the 1004 Hz tone, a 4 Hz deviant near the threshold level of auditory frequency discrimination, did not. Moreover, they showed that the amplitude of MMN in the 1016 Hz was larger than that in the 1008 Hz. Re amplitude of stimuli, Näätänen et al. [19] used an 80 dB SPL tone as standard and 77, 70, and 50 dB SPL tones as deviants. They found that the larger the difference between the standard stimulus and deviant stimulus was, the larger the amplitude of MMN.

In the present study, we measured MMF using MEG. Identifying the response area (hemisphere) is relatively easy using MEG, but more difficult using EEG due to the heterogeneity of conductivity in bones of the cranium. If Japanese subjects detect the difference between voiced and devoiced vowels (allophonic), they elicit MMF; however, if they cannot detect the difference, they do not elicit MMF. In addition, we explored whether characteristics of the allophonic MMF are the same as those of the phonemic ones.

2 EXPERIMENT 1

2.1. Method

Stimuli

We selected stimuli which fulfilled the following requirements: (1) High devoicing rates [see 5]; (2) first consonant a stop or an affricate (easy to determine consonant onset); (3) meaningful words; and (4) same pitch accent type. Based on this, we selected stimuli, [tsɯta] (ivy), [tsɯ̥ta], and [ʧi̥ta] (place name), which were recorded from a Japanese male speaker aged 45, a speech scientist who was able to easily produce the voiced vowel in the devoicing context. He uttered [tsɯta], [tsɯ̥ta] and [ʧi̥ta] 10 times each and these were sampled at 11.025 kHz, digitized at 16 bits. This sampling rate was due to the limitation of the sound presentation system. Stimuli [tsɯta] and [tsɯ̥ta] were cut into [tsɯ] and [ta], and [tsɯ̥] and [ta]. Then these two [ta] were discarded. [tsɯ] was then aligned to 140 ms by trimming a few cycles from the vowel portion. Next, a new [ta] was uttered by the same speaker and was aligned to 200 ms by trimming. [tsɯ] and [ta], and [tsɯ̥] and [ta] were then linearly aligned to form a sequence of [tsɯta] and [tsɯ̥ta] with a total length of 420 ms, including a stop closure of 80 ms between the first and second syllables. Stimuli for phoneme discrimination, [ʧi̥ta], were made the same way. Figure 1 shows the spectrograms of these stimuli.

A. Allophonic condition. An oddball paradigm was adopted, in which a standard stimulus was presented at a high probability of 85% and a deviant stimulus was presented at a low probability of 15%. Deviants never occurred in immediate succession. The maximum number of standards between adjacent deviants was 19, and the minimum number was 2. The stimulus onset asynchrony (SOA) was 1 s. SOA is Inter-stimulus interval (ISI) plus stimulus duration. Two sessions, one with the standard-deviant pair [tsɯta] vs. [tsɯ̥ta], and one with the standard-deviant pair [tsɯ̥ta] vs. [tsɯta], were conducted with an interleave 5 min. rest in a counter-balanced order. As can be seen in the spectrograms in figure 1, the acoustical difference between [tsɯ̥ta] and [tsɯta] is that the vowel [ɯ̥] in [tsɯ̥ta] is devoiced and the fricative portion of [ts] is prolonged.

B. Phonemic condition. As in the allophonic condition, stimuli were presented using the oddball paradigm. Deviants never occurred in immediate succession. The SOA was 1 s. Two sessions, one with the standard-deviant pair [tsɯ̥ta] vs. [ʧi̥ta], and one with the standard-deviant pair

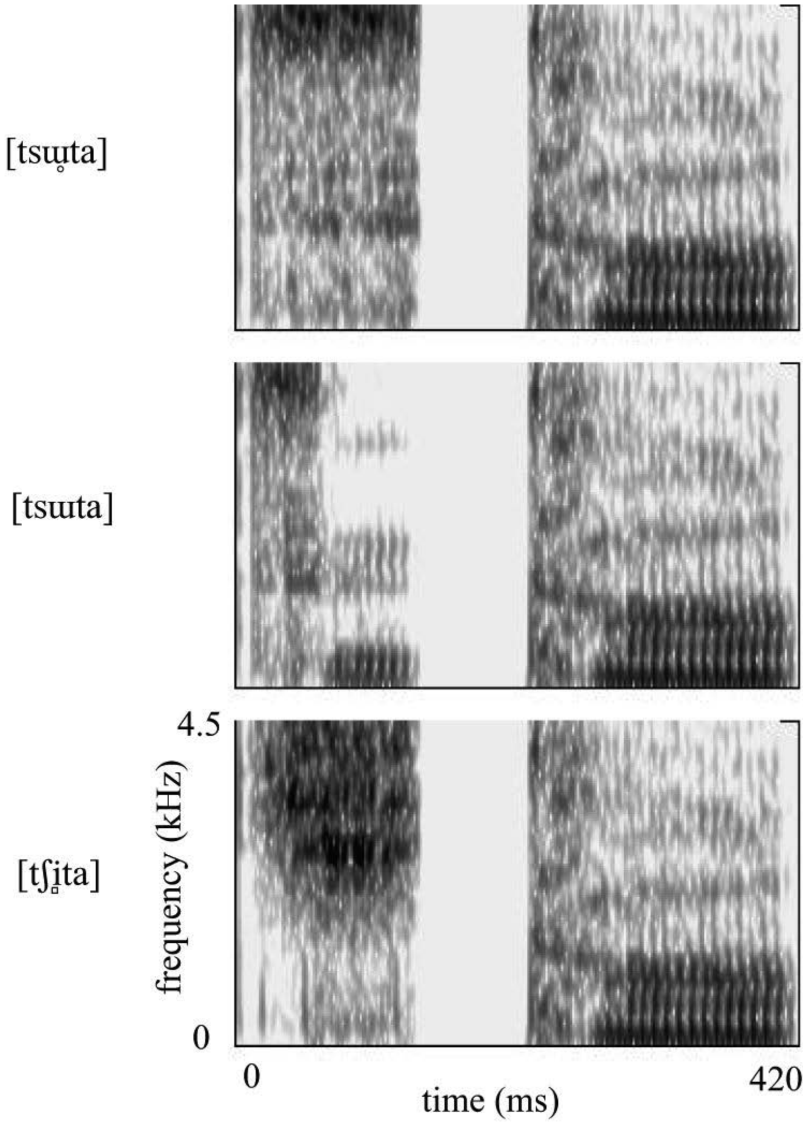

Fig. 1. Spectrograms of the stimuli.

[ʧi̥ta] vs. [tsɯ̥ta], were conducted with an interleave 5 min. rest in a counter-balanced order. As can be seen from the spectrograms in figure 1, the acoustical difference between [tsɯ̥ta] and [ʧi̥ ta] is the central frequency of frication in [tsɯ̥ta] is higher than that in [ʧi̥ta], and vowel [ɯ̥] in [tsɯ̥ta] and vowel [i̥] in [ʧi̥ta] are devoiced.

Subjects

The subjects were 12 native speakers of Japanese (3 males, 9 females) aged 20-46 years. Six subjects were speakers of the dialects where vowel devoicing occurs frequently (3 from Kanto, 2 from Kyushu, and 1 from Hokuriku: devoiced vowel subjects) and the others were from the dialects where vowel devoicing usually doesn't occur (3 from Hiroshima, 2 from Kinki, and 1 from Chubu: voiced vowel subjects). Nine subjects were university students and 3 were university staff. Four of these students graduated, so they could not take part in Experiment 2. All of the subjects were right-handed and had no hearing loss. The subjects were instructed to not concentrate on the stimuli but to listen passively. The stimuli were binaurally presented to the subjects using inserted earphones. Stimuli were presented at the most comfortable, subject-determined sound level. The average presentation level was a sensation level of 60 dB. All subjects gave their written informed consent before the experiments.

MEG recordings and analyses

The recordings were performed in a magnetically shielded room using a 204-channel whole head gradiometer (Neuromag Ltd., Finland). The MEG epochs, starting 100 ms before, and ending 600 ms after each stimulus onset, were averaged separately for the standard and the deviant stimuli, and filtered using a 1.2-26 Hz digital band-pass filter. The MMF was determined from subtracting waves from the deviant stimulus response minus the standard stimulus response. That is, in the phonemic condition, subtracting waves were calculated as follows: (1) MMF of [ts] deviant was calculated such that the magnetic wave of [ts] which served as a standard was subtracted from the magnetic wave of [ts] served as a deviant, and (2) MMF of [ʧ] deviant was calculated such that the magnetic wave of [ʧ] which served as a standard was subtracted from the magnetic wave of [ʧ] which served as a deviant. The MMF was determined between 160 and 310 ms (allophonic condition) or between 100 and 250 ms (phonemic condition). This was done because the second phonemes, [ɯ] and [ɯ̥], differ in the allophonic condition, while the first phonemes, [ts] and [ʧ], differ in the phonemic condition. Figure 2 shows an example of magnetoencephalographic wave forms. In this figure, a dashed line represents the deviant stimulus response, a dotted line represents the standard response, and a solid line represents the curve obtained by subtracting responses to standards from responses to deviants. The MMF response appears in 170 ms or thereabout in this figure.

For exploring the magnitude of MMF in each hemisphere, we estimated the current source using a single dipole model [22]. When subjects perceive stimuli, many neurons in their cerebra are activated. Many activated neurons in a small area are approximated by an equivalent current dipole (ECD). ECDs are determined as the difference between measured magnetic

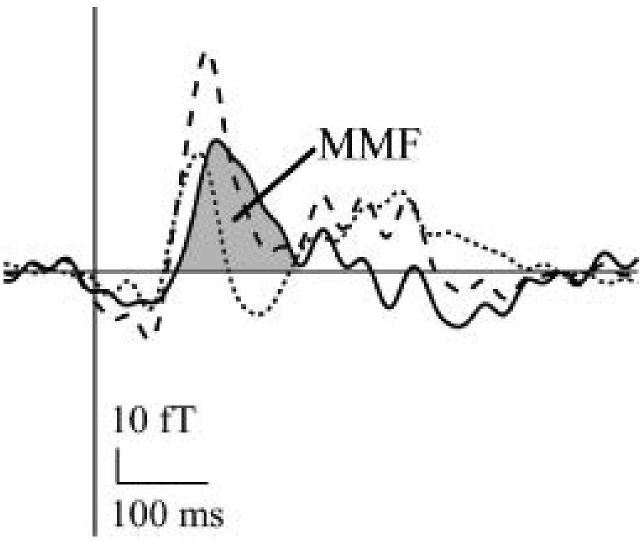

Fig. 2. An example of magnetoencephalographic wave forms of one channel in left hemisphere. Dashed line, deviant stimulus response; dotted line, standard stimulus response; solid line, subtracted standard response from deviant response.

distribution and magnetic distribution generated by an ECD becomes minimal. ECDs can best reproduce the magnetic fields measured in each channel. The locations and directions of these ECDs in a head are calculated by a computer. ECDs were determined for the MMF for each subject and condition in each hemisphere.

2.2 Results

Figures 3 and 4 represent the magnetic wave form of [ʧi̥ta] in the phonemic condition and [tsɯta] in the allophonic condition, respectively. MMF were elicited in the left and right hemispheres (LH and RH, respectively) in the phonemic condition, and also were elicited in the allophonic condition. In the phonemic condition MMF in LH was larger than that in RH, while in the allophonic condition MMF in RH was larger than that in LH.

Figures 5 and 6 represent the ECD moments in the phonemic condition and the allophonic condition, respectively. The ECD moment represents the strength of a dipole, i. e. the magnitude of brain response. In the phonemic condition, the ECD moments of deviant [tsɯ̥ta] and [ʧi̥ta] were almost the same strength in the left hemisphere and in the right hemisphere, respectively. In the allophonic condition, however, the ECD moments of the deviant [tsɯta] and [tsɯ̥ta] were not the same strength, as shown in figure 6. An ANOVA with two factors, Hemisphere and Deviant, was performed in both the phonemic (figure 5) and allophonic condition (figure 6). The ANOVA showed a significant main effect of Hemisphere in the phonemic condition ($F(1,44)=8.285$; $p=0.0061$), and significant main effects of both Hemispheres ($F(1,44)=8.759$; $p=0.0049$) and Deviant ($F(1,44)=39.422$; $p<0.0001$) in the allophonic condition, but Deviant ($F(1,44)=1.014$; n.s.) and interaction ($F(1,44)=1.503$; n.s.) in the phonemic condition, and interaction ($F(1,44)=0.997$; n.s.) in the allophonic condition. Namely, the phonemic condition showed left hemispheric dominance, whereas the allophonic condition, right hemispheric dominance. In general, phonemic perception is left hemispheric dominant [23], while prosodic perception [23], and/or the change in physical features of stimuli, e. g. pure tone or speech

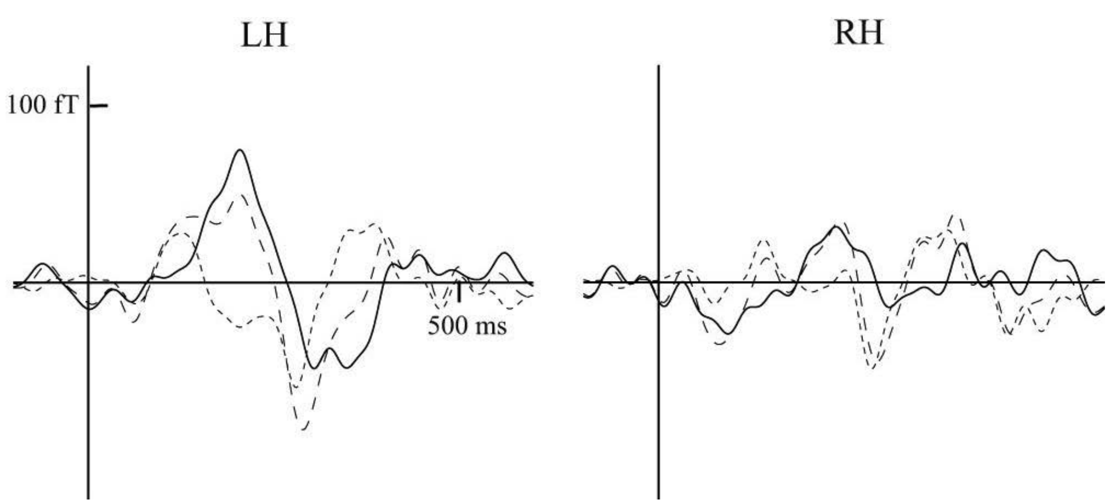

Fig. 3. Magnetoencephalographic wave forms of one subject in the phonetic condition. The dashed line represents [ʧi̥ta], which served as a deviant, the dotted line represents [ʧita], which served as a standard, and the solid line represents the subtracted line (deviant minus standard).

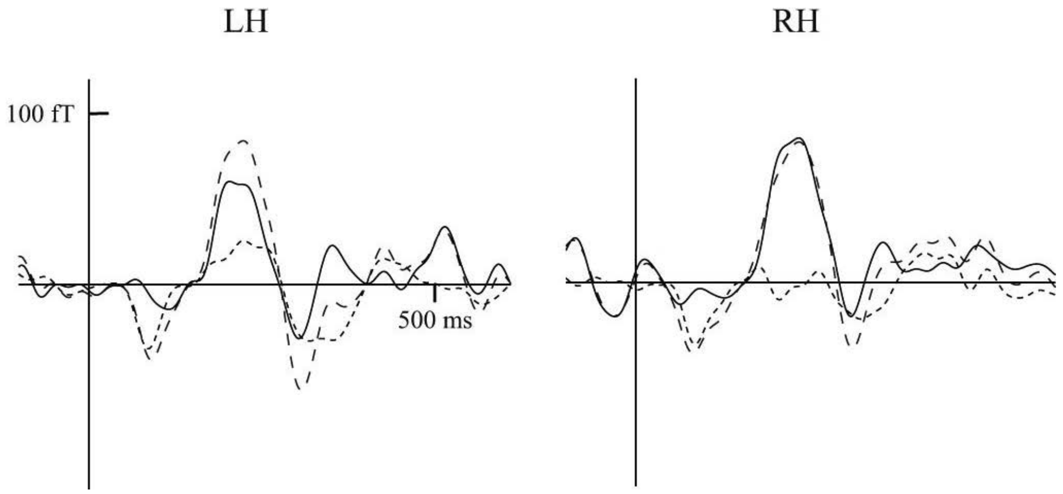

Fig. 4. Magnetoencephalographic wave forms of one subject (different from figure 3's subject) in the phonetic condition. The dashed line represents [tsɯta], which served as a deviant, the dotted line represents [tsɯta], which served as a standard, and the solid line represents the subtracted line (deviant minus standard).

sound duration, [24] is right hemispheric dominant. Namely, our result is consistent with these previous investigations.

In the phonemic condition, the difference between the ECD moment of the deviant [ʧi̥ta] and the deviant [tsɯ̥ta] was not significant. That is to say, there was no dependence of the ECD moment upon the deviant stimuli in the phonemic condition. On the other hand, in the allophonic condition, the ECD moment of the voiced [tsɯta] was larger than that of the devoiced [tsɯ̥ta] in both hemispheres. In the allophonic condition, in contrast to the phonemic condition, there was a dependence of the strength of the ECD moment upon the deviant stimuli. In this experiment the most important result is that the strength of the ECD moment depended upon deviants. In previous mismatch studies as reported in the Introduction, this mismatch dependence upon deviants was not mentioned.

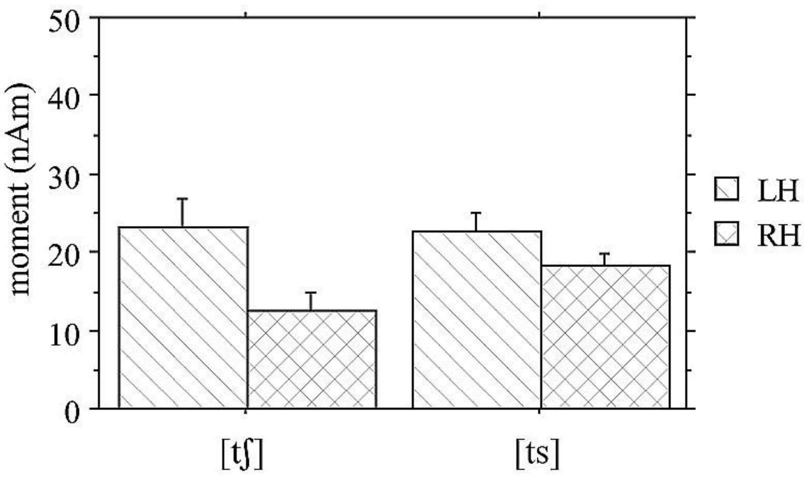

Fig. 5. Dependence of the ECD moment upon the deviant stimuli in the phonemic condition. [ʧ], [ʧi̥ta] was presented as deviant; [ts], [tsɯ̥ta] was presented as deviant. LH, left hemisphere; RH, right hemisphere. Error bars indicate the standard error. There was no deviant dependence in this condition.

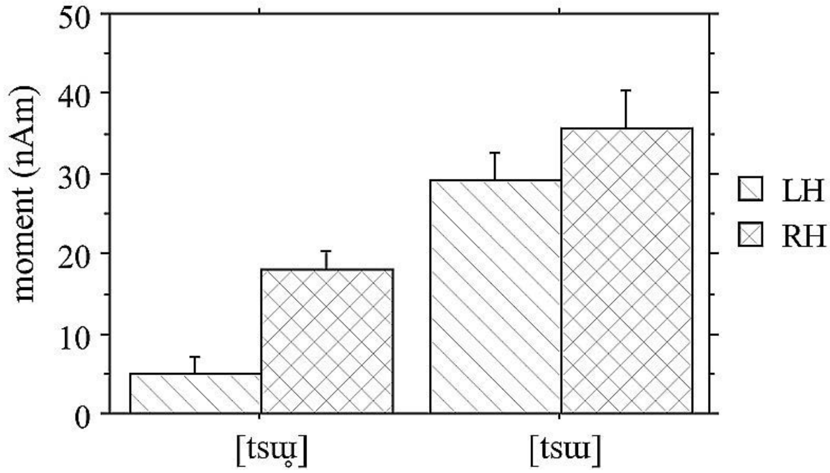

Fig. 6. Dependence of the ECD moment upon the deviant stimuli in the allophonic condition. [tsɯ̥], [tsɯ̥ta] was presented as deviant; [tsɯ], [tsɯta] was presented as deviant.

As mentioned in the Introduction, devoiced vowel dialects and non-devoiced vowel dialects exist in Japanese. Noting this, we studied effects from listeners' dialects upon deviant dependence. Figures 7 and 8 are the ECD moments obtained from the 6 subjects who spoke in devoiced vowel dialects and the 6 subjects who spoke in non-devoiced vowel dialects, respectively. In the phonemic condition (figure 7), the ECD moments of the deviant [ʧi̥ta]

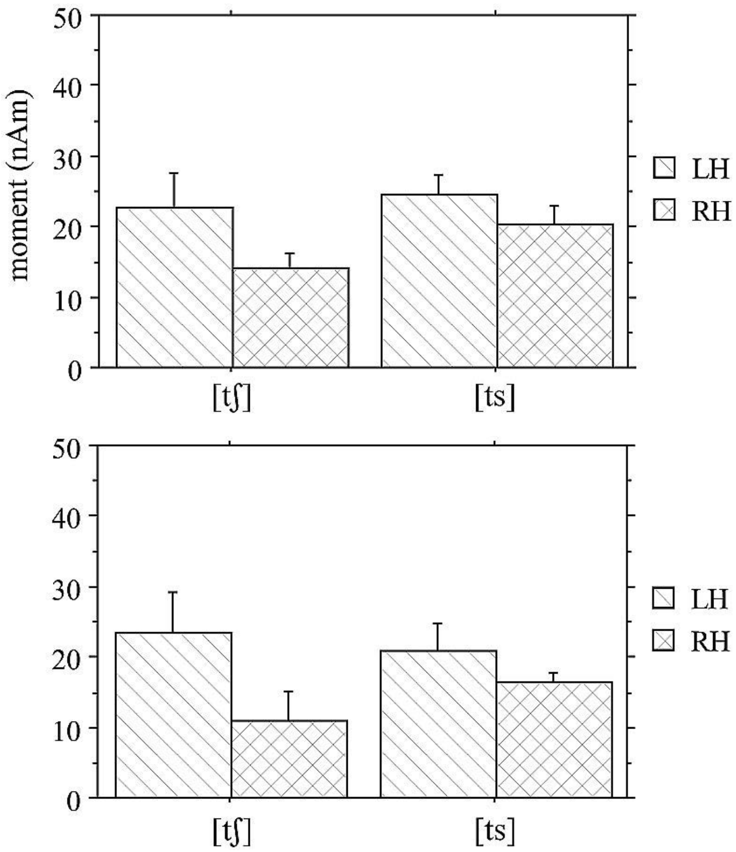

Fig. 7. Dependence of the ECD moment upon the deviant stimuli in the phonemic condition. Top, devoiced vowel dialect speakers; Bottom, non-devoiced vowel dialect speakers.

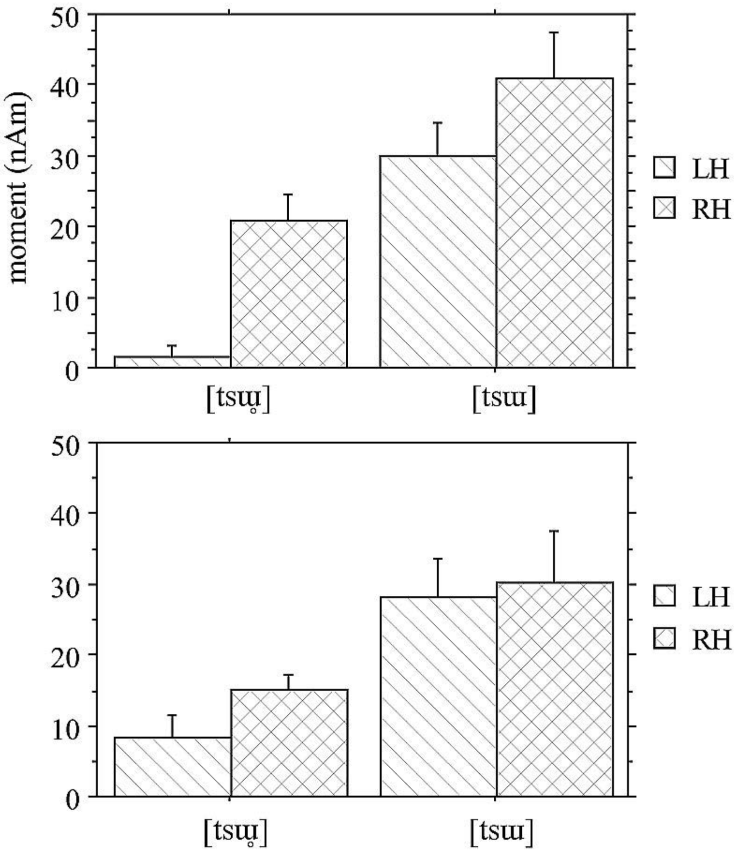

Fig. 8. Dependence of the ECD moment upon the deviant stimuli in the allophonic condition. Top, devoiced vowel dialect speakers; Bottom, non-devoiced vowel dialect speakers.

and the deviant [tsɯ̥ta] in both hemispheres were nearly equal in the devoiced vowel and non-devoiced vowel subjects. In the allophonic condition (figure 8), similar to the phonemic condition, the ECD moments of the deviant [tsɯta] and deviant [tsɯ̥ta] in both hemispheres were nearly equal in the devoiced vowel and non-devoiced vowel subjects. Thus, no differences in the ECD moments between the devoiced vowel subjects and non-devoiced vowel subjects were seen. Namely, the strength of the ECD moment did not depend upon deviants in both devoiced vowel dialects and non-devoiced vowel dialects. This is the most important result in our experiments.

As also mentioned in the Introduction, as the acoustical difference between the standard stimulus and deviant stimulus becomes larger, the magnitude of the mismatch response increases. In this experiment, the acoustical differences between the standard and deviant stimuli were equal irrespective of the tasks (e.g., the standard voiced [tsɯta] vs. the deviant devoiced [tsɯ̥ta], the standard devoiced [tsɯ̥ta] vs. the deviant voiced [tsɯta]) in both the allophonic and phonemic conditions. Despite the acoustic equality, the magnitude of the mismatch response in the allophonic condition changed depending on deviants (asymmetric), but in the phonemic condition, it was almost constant independent of deviants (symmetric). Therefore, we conducted a second experiment to examine the effect of the voiced and devoiced vowel portions in the stimuli on the mismatch response. We performed this experiment using the voiced and devoiced vowel stimuli cut out from the original stimuli.

3 EXPERIMENT 2

3.1. Method

Stimuli

The stimuli were the voiced [ɯ] portion of [tsɯta] and the devoiced [ɯ̥] portion of [tsɯ̥ta]. Both were 75 ms segments extracted from the stimuli used in experiment 1. Avoiding artifacts, stimuli were carefully extracted and 20 ms tapered at both ends. Figure 9 shows the wave forms and the spectrograms of these stimuli.

As in experiment 1, an oddball paradigm was used, in which a standard stimulus was presented at a probability of 85% and a deviant stimulus was presented at a probability of 15%. Deviants never occurred in immediate succession. The SOA was 1 s. Two sessions with respective standard-deviant pairs of [ɯ] vs. [ɯ̥] and [ɯ̥] vs. [ɯ] were conducted with an interleave 5 min. rest in a counter-balanced order.

Subjects

The subjects were 8 native speakers of Japanese (3 males, 5 females) aged 20-46 years. All of the subjects had participated in experiment 1. The subjects were instructed to not concentrate on the stimuli but to listen passively. The stimuli were binaurally presented to the subjects with

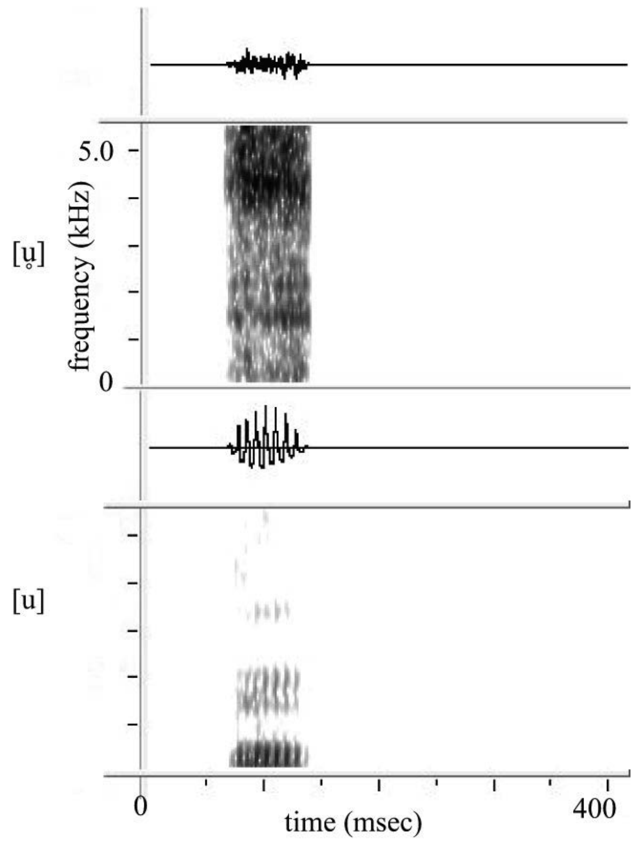

Fig. 9. Wave forms and spectrograms of stimuli.

inserted earphones.

MEG recordings and analyses

The recordings were made by the same methods of experiment 1. The MMF was determined from subtracting waves from the deviant stimulus response minus the standard stimulus response between 160 and 310 ms. ECD moments were determined for the MMF for each subject and condition.

3.2. Results

Figure 10 shows the ECD moment in this experiment. The ECD moment was larger in the right hemisphere than in the left hemisphere in both the voiced and devoiced conditions. Yet in contrast to the results of experiment 1, the ECD moments of the deviant voiced stimulus and the deviant devoiced stimulus did not differ from each other in either hemisphere. An ANOVA with two factors, Hemisphere and Deviant, revealed a significant main effect of Hemisphere ($F(1,28)=4.241$; $p=0.0489$) but no significant effect of Deviant ($F(1,28)=0.006$; n.s.) and interaction ($F(1,28)=0.190$; n.s.). Consequently, the asymmetry of the ECD moment in deviants did not arise from the acoustical difference between the voiced and devoiced portions. Incidentally, these results reveal that the fundamental frequency has no effect on the asymmetry.

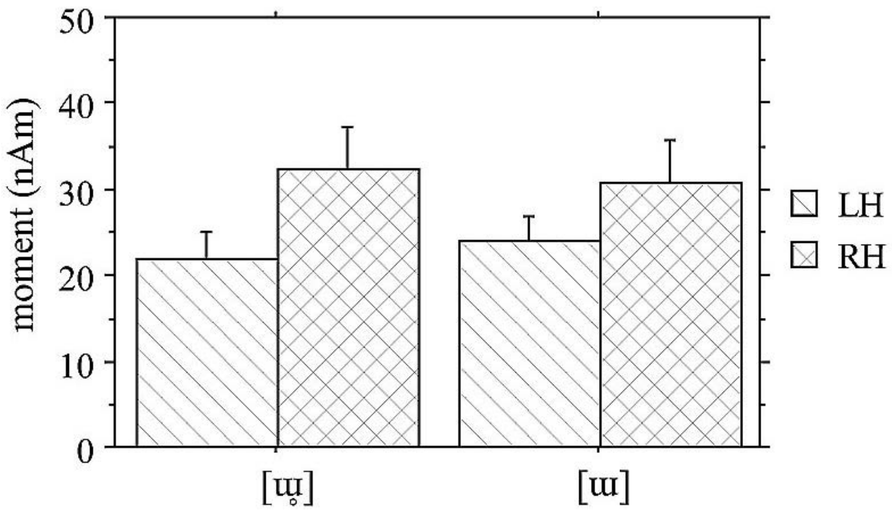

Fig. 10. ECD moment in experiment 2. [ɯ̥] was presented as deviant; [ɯ̥], [ɯ] was presented as deviant [ɯ].

4 DISCUSSION

The results of our present study suggest that the ECD moment elicited by the voiced vowel deviant stimulus is significantly larger than that elicited by the devoiced vowel deviant stimulus in the allophonic condition in both hemispheres. However, the deviants elicited no such effects in the phonemic condition.

In general, as mentioned earlier, as the acoustic difference between a standard stimulus and a deviant stimulus becomes larger, the strength of the ECD moment becomes larger [e.g., 20]. Yet in our experiments, the acoustic difference between the standard stimulus and deviant stimulus was equal, irrespective of the deviants. Accordingly, we expected the ECD moment of the voiced vowel deviant to be almost equal to that of the devoiced vowel deviant. In fact, the ECD moments were almost equal in the phonemic condition but in the allophonic condition, the ECD moment was asymmetric with regard to which stimulus became the deviant.

Experiment 2 was performed using stimuli, constructed only by the voiced and devoiced portions of [tsɯta] and [tsɯ̥ta]. The results did not show asymmetric ECD moment. That is, the asymmetry of the mismatch originated not from the acoustic difference (i.e., the sound pressure, fundamental frequency) between the standard and the deviant stimulus per se, but presumably from the phonetic environment in the words. Namely, when [ɯ] and [ɯ̥] are embedded in the word [ts_ta], the asymmetry of the mismatch occurs.

Näätänen et al. [25] found in the Finnish language that the prototype deviant elicited a larger mismatch response than the non-prototype deviant, independent of the acoustic distance between the standard and the deviant stimulus. In their experiments, they measured mismatch responses by using EEG and MEG for Finns and Estonians. Finnish and Estonian have very similar vowel systems. Vowels /e/, /ö/ and /o/ are common to both languages. The only exception is Estonian /õ/, which has roughly intermediate acoustic characteristics between /ö/ and /o/. Estonian /õ/ is the non-prototype of Finnish /ö/. They ran an oddball paradigm, in which the standard stimulus was the vowel /e/ and the deviants were /ö/, /õ/ and /o/, and they measured mismatch responses. The mismatch responses in Estonians increased /ö/, /õ/, /o/, in the order as acoustic distance increased, while in Finns, the mismatch response in /õ/ was smaller than that in /ö/, in spite of the greater acoustic distance.

In Japanese standard pronunciation, vowels /i/ and /u/ are devoiced between voiceless consonants, but not voiced. The devoicing rate of vowel /i/ between /tʃ/ and /t/ is 84.98 % and that of vowel /ɯ/ between /ts/ and /t/ is 91.59 % [5]. Namely, if the devoiced vowel between voiceless consonants is prototypical, devoiced vowel deviants elicited larger mismatch responses than voiced vowel deviants. However, in fact, voiced vowel deviants elicited larger mismatch responses than devoiced vowel deviants. Is the Japanese prototype the voiced vowel even between voiceless consonants, irrespective of the standard pronunciation? Or, are there cases where the magnitude of the mismatch response of the prototype (standard pronunciation, i.e., devoiced vowel) is smaller than that of the non-prototype? In our study, the mismatch responses in the allophonic condition showed no differences between the hemispheres for the listeners with devoiced vowel dialects and non-devoiced dialects (figure 7). That is, in both dialects, the mismatch response of the non-devoiced vowel deviant was larger than that of the devoiced vowel deviant. There were no differences in the brain responses between the dialects.

According to the study of Imaizumi, Fuwa, and Hosoi [26], Japanese children living in both voiced vowel areas and devoiced vowel areas first acquired the voiced vowels between voiceless consonants by the age of 4 years. At the age of 5 years, the children living in an area with standard pronunciation acquired devoiced vowels between voiceless consonants. On this basis, Imaizumi et al. concluded that the voiced vowels, but not the devoiced, are acquired as vowel prototypes, even in areas with standard pronunciation. We can thus conceive that the prototype deviant (utterance with the voiced vowel between voiceless consonants) may elicit in both hemispheres significantly larger mismatch responses than the non-prototype deviant (utterance with the devoiced vowel between voiceless consonants). Thus, Japanese speakers acquire voiced vowel pronunciation even between voiceless consonants as the prototype in early childhood, and retain it as prototype even after growing up and pronouncing the devoiced vowel between voiceless consonants. Once the prototypes of phonemes in a native language are acquired in childhood, they remain robust in the brain for many years to come, even after allophones are learned and used in daily life. Consequently, we can conclude, in the light of these results on brain activation, that the Japanese prototype is the voiced vowel, not the devoiced vowel, irrespective of the dialects spoken.

This study with native listeners of Japanese showed that mismatch responses were elicited in the discrimination between voiced and devoiced vowels. These mismatch responses were different from those in the phonemic discrimination. In the phonemic discrimination, the magnitude of the mismatch response was almost invariable for the exchange between a deviant stimulus and a standard stimulus. In the allophonic discrimination, the magnitude of the mismatch response varied according to the nature of the stimulus. Namely, the magnitude of the mismatch response is larger in the prototype stimulus than in the non- prototype one.

5 CONCLUSIONS

In this study we analyzed MMF using MEG to test how the brain responses to allophonic contrast arise in native speakers of Japanese. Our results suggest that mismatch responses were elicited in the discrimination between allophones. The voiced deviant appeared to elicit significantly larger MMF than the devoiced deviant in the allophonic discrimination, while the MMF had almost the same amplitude irrespective of the deviant types in the phonemic condition. This suggests that voiced vowels between voiceless consonants are processed as the vowel prototype in Japanese. From a brain science perspective, the voiced vowel (e. g. [ɯ] in [tsɯta]) between voiceless consonants may be the prototype in Japanese, even though the devoiced vowel (e. g. [ɯ̥] in [tsɯ̥ta]) between voiceless consonants is the one that occurs in standard Japanese.

ACKNOWLEDGEMENTS

We thank Prof. Donna Erickson (Kanazawa Medical University) for her helpful advice on this article. This study was partially supported by a Grant-in-Aid for Scientific Research (No. 18520327, 20520394) from the Japan Society for the Promotion of Science.

REFERENCES

[1] M. S. Han, "Unvoicing of vowels in Japanese," *The Study of Sounds*, 10. 81-100 (1962).

[2] F. Kuriyagawa and M. Sawashima, "Word accent, devoicing and duration of vowels in Japanese," *Annual Bulletin RILP*, Faculty of Medicine, University of Tokyo, 23. 85-108 (1989).

[3] K. Maekawa, "Probability of vowel devoicing in common Japanese," (in Japanese). *The world of language*, 1. 69-81 (1983).

[4] M. Fujimoto, "Effects of consonant type and syllable position within a word on vowel devoicing in Japanese," *Proc. Int. Conf. Speech Prosody.*, 625-628 (2004).

[5] K. Maekawa and H. Kikuchi, "Corpus-based analysis of vowel devoicing in spontaneous Japanese: an interim report," in *Voicing in Japanese*, van de Weijer, Nanjo, and Nishihara, Eds. (Mouton de Gruyter, Berlin, 2005) pp. 205-228.

[6] K. Maekawa, "Effects of speaking rate on the voicing variation in Japanese," *IEICE Tech. Rep.*, SP-89-148. 47-53 (1990).

[7] S. Imaizumi, A. Hayashi, and T. Deguchi, "Listener adaptive characteristics of vowel devoicing in Japanese dialogue" *J. Acous. Soc. Am.*, 98. 768-778 (1995).

[8] M. Sugito, "The effect of speech rate on devoiced accented vowels in Osaka Japanese," in *Voicing in Japanese*, van de Weijer, Nanjo, and Nishihara, Eds. (Mouton de Gruyter, Berlin, 2005) pp. 247-260.

[9] M. Fujimoto and S. Kiritani, "Comparison of vowel devoicing for speakers of Tokyo- and Kinki dialects," *J. Phon. Soc. Jpn.*, 7. 58-69 (2003).

[10] H. Hirose, "The activity of the adductor laryngeal muscles in respect to vowel devoicing in Japanese," *Phonetica*, 23. 156-170 (1971).

[11] H. Yoshioka, "Laryngeal adjustments in the production of the fricative consonants and devoiced vowels in Japanese," *Phonetica*, 38. 236-251 (1981).

[12] M. Fujimoto, E. Murano, S. Niimi, and S. Kiritani, "Difference in glottal opening pattern between Tokyo and Osaka dialect speakers: Factors contributing to vowel devoicing," *Folia Phoniatr. Logop.*, 54. 133-143 (2002).

[13] A. Cutler, T. Otake, and J. M. McQueen, "Vowel devoicing and perception of spoken Japanese words," *J. Acoust. Soc. Am.*, 125. 1693-1703 (2009).

[14] R. Näätänen, A. W. Gaillard, and S. Mäntysalo, "Early selective-attention effect on evoked potential reinterpreted," *Acta Psychol. (Amst)*, 42. 313-329 (1978).

[15] R. Hari, M. Hämäläinen, R. Ilmoniemi, E. Kaukoranta, K. Reinikainen, J. Salminen, K. Alho, R. Näätänen, and M. Sams, "Responses of the primary auditory cortex to pitch changes in a sequence of tone pips: neuromagnetic recordings in man," *Neurosci. Lett.*, 50. 127-132 (1984).

[16] M. Sams, P. Paavilainen, K. Alho, and R. Näätänen, "Auditory frequency discrimination and event-related potentials," *Electroencephalogr. Clin. Neurophysiol.*, 62. 437-448 (1985).

[17] M. Sams, E. Kaukoranta, M. Hämäläinen, and R. Näätänen, "Cortical activity elicited

by changes in auditory stimuli: different sources for the magnetic N100m and mismatch responses," *Psychophysiology*, 28. 21-29 (1991).

[18] O. Lounasmaa, R. Hari, SL. Joutsiniemi, and M. Hämäläinen, "Multi-SQUID recordings of human cerebral magnetic fields may give information about memory processes," *Europhysics Lett*, 9. 609-612 (1989).

[19] R. Näätänen, P. Paavilainen, K. Alho, K. Reinikainen, and M. Sams, "Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain?" *Neurosci. Lett.*, 98. 217-221 (1989).

[20] E. Kaukoranta, M. Sams, R. Hari, M. Hämäläinen, and R. Näätänen, "Reactions of human auditory cortex to a change in tone duration," *Hear. Res*. 41. 15-21 (1989).

[21] R. Näätänen, P. Paavilainen, and K. Reinikainen, "Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man?" *Neurosci. Lett.*, 107. 347-352 (1989).

[22] S. J. Williamson and L. Kaufman, "Biomagnetism," *J. Magnetism and Magnetic Materials*, 22. 129-201 (1981).

[23] S. Imaizumi, K. Mori, S. Kiritani, H. Hosoi, and M. Tonoike, "Task-dependent laterality for cue decoding during spoken language processing," *NeuroReport*, 9. 899-903 (1998).

[24] K. Kasai, H. Yamada, S. Kamio, K. Nakagome, A. Iwanami, M. Fukuda, K. Itoh, I. Koshida, M. Yumoto, K. Iramina, N. Kato, and S. Ueno, "Brain lateralization for mismatch response to across- and within-category change of vowels," *NeuroReport*, 12. 2467-2471 (2001).

[25] R. Näätänen, A. Lehtokoski, M. Lennes, M. Cheour, M. Huotilainen, A. Iivonen, M. Vainio, P. Alku, R. J. Ilmoniemi, A. Luuk, J. Allik, J. Sinkkonen, and K. Alho, "Language-specific phoneme representations revealed by electric and magnetic brain responses," *Nature*, 385. 432-434 (1997).

[26] S. Imaizumi, K. Fuwa, and H. Hosoi, "Development of adaptive phonetic gestures in children: evidence from vowel devoicing in two different dialects of Japanese," *J. Acoust. Soc. Am.*, 106. 1033-1044 (1999).