

Differential cortical responses in second language learners to different vowel contrasts

Yasuyo Minagawa-Kawai,^{1,2,CA} Koichi Mori,¹ Yutaka Sato¹ and Toshizo Koizumi¹

¹Department of Sensory and Communicative Disorders, Research Institute, National Rehabilitation Center for Persons with Disabilities, 4-1 Namiki, Tokorozawa, Saitama 359-8555; ²Independent Administrative Institution, National Institute for Japanese Language, 3-9-14 Nishigaoka, Kita-ku, Tokyo 115-8620, Japan

^{CA}Corresponding Author: myasuyo@bea.hi-ho.ne.jp

Received 19 December 2003; accepted 12 January 2004

DOI: 10.1097/01.wnr.0000118983.36602.02

Near-infrared spectroscopy was used to measure hemodynamic responses in Korean subjects learning Japanese as a second-language to Japanese phonemic contrasts that are either phonologically distinctive or non-distinctive in their first-language. These results were compared with those of Japanese native listeners reported previously. Unlike the results observed in the Japanese subjects, the Korean subjects did not show category-specific neural responses to a durational contrast that is non-distinctive in

their first-language. The /a-e/ contrast elicited the larger response in the left auditory area, consistent with the results of the Japanese. These phoneme-dependent responses imply that the neuronal networks subserving the first and second-languages are either shared or exclusive according to the mutual relationship between the phonemic properties of the first and second-languages. *NeuroReport* 15:899-903 © 2004 Lippincott Williams & Wilkins.

Key words: Categorical perception; First language; Multilingualism; Near-infrared spectroscopy (NIRS); Phoneme perception; Second language; Suprasegmental; Vowel duration

INTRODUCTION

Since perception of phoneme contrasts is heavily shaped by linguistic experience with the first-language (L1) phonology [1], second-language (L2) learners as adults have difficulty in perceptually differentiating certain phonemic contrasts in L2 (e.g. English /r-l/ by Japanese) [2]. Neuronal correlates of phonemic perception have been examined by quite a few studies [3-7], most of which reveal cerebral representations of L1 phonemes in the left temporal area. Some of these studies [3,4] verified neurophysiological evidence of L1-specific phonemic perception, comparing neural responses to L1 phonemes and non-native ones which had not been learned at all. However, neural substrates for phoneme processing of L2 learners are only starting to be explored [8] and few studies have tested highly-proficient L2 learners.

Japanese has phonemic contrasts that differ only in duration and appear in any syllable position. In a few languages, such duration-based contrasts are phonologically distinctive, i.e. two segments are distinctive if they provide a minimal morphemic contrast in the same environment as in right and light for /r-l/ contrast in the English phonological system [9]. Accordingly, many learners of Japanese as adults find it difficult to discriminate Japanese long/short vowels and consonants [10,11]. The present study measured the cerebral hemodynamic responses to this form of durational contrast for Korean subjects in order to address the question whether non-native speakers with a highly-proficient language command of Japanese would process non-native phonemes similarly to

distinctive ones in their native language. For this purpose, cerebral responses to stimulus pairs of /a/ and /e/ contrast which is distinctive in Korean language were also measured as a control experiment.

MATERIALS AND METHODS

Subjects: Seven right-handed native speakers of Korean with normal hearing, aged between 27 and 40 years (mean (\pm s.d.) 32 ± 4.3 ; three males, four females) served as subjects. Written informed consent was obtained before the experiment. All the Korean subjects had stayed in Japan for more than three years during which time they used Japanese in their daily activities. They started learning Japanese at the age of ≥ 20 years. The average Japanese learning experience was 7.6 ± 4.1 years.

Procedures: Changes in the hemoglobin concentration and its oxygenation level in the bilateral temporal areas were recorded using a 24-channel near-infrared spectroscopy (NIRS) system (ETG-100, Hitachi Medical Co., Japan). The instrument emits continuous near-infrared lasers whose wavelengths were ~ 780 nm and 830 nm. The laser beams were modulated at different frequencies and detected with lock-in amplifiers [12]. The sites of recording reside in the optical paths in the brain between the nearest pairs of incident and detection probes separated by 3 cm on the scalp surface. Five incident and four detection probes arranged in a 3×3 square lattice were fitted on each lateral

side of the head. After the optical measurements, the positions of the optical probes were recorded with a 3D digitizer (Polhemus, Vermont, USA). The centers of recording sites were confirmed by superimposing the 3D coordinates onto T1-weighted MR brain images for each subject (Fig. 1). The channels nearest to the lateral end of the border between the transverse temporal gyrus and the planum temporale in a parasagittal projection were presumed to be in the auditory area. The recording channels whose centers were within the 1.5 cm radius of the above-mentioned border were selected. These channels should include the signals in the auditory cortex due to the spread of laser in the brain tissue [13]. Averaged x , y , z coordinates for the center of the auditory area were -61 ± 2.3 , -9.4 ± 3.1 and 5.2 ± 1.7 mm on the left side and 62.8 ± 2.0 , -8.8 ± 4.4 and 5.0 ± 1.7 mm on the right side (Fig. 1), where positive x is to the right, y is to the anterior, and z superior (coordinate system of Talairach and Tournoux).

Stimuli: The stimuli consisted of four instances of /mama/ or /mama:/ whose final vowel were 151 ms (A), 184 ms (B), 217 ms (C) and 250 ms (D). These pseudo words were synthesized to have a Japanese unaccented pattern with a steady formant structure in the final vowel, using a PARCOR procedure by *Rokubunken* [14]. Since a pilot behavioral experiment with Japanese listeners showed that the phonemic boundary (PB) of long and short vowels resided between stimuli B and C, three sessions including across- and within-category conditions were performed using the adjacent two stimuli among the stimuli A–D. In session AB (a within-category condition), stimulus A was repeated for 20 s as a baseline block and then stimuli A and B were presented in a pseudo-random order with equal probabilities for another 20 s as a target block with SOA of 1.25 s. These two blocks were alternated and repeated at least five times. Similar procedures were carried out for session BC (an across-category condition) and session CD (a within-category condition). As a control, Japanese

analysis-synthesized words /itta/ and /itte/ [15] were presented in a similar paradigm as above, with /itta/ used both in baseline and target blocks and /itte/ appearing only in target blocks 50% of the time. A pilot behavioral experiment of identification confirmed that Koreans were able to identify /itta/ and /itte/ with an accuracy of 99.1%.

The four sessions (AB, BC, CD and /a-e/) were presented in random order to each subject. The subjects were instructed to listen to the stimuli delivered through insert earphones (Ear Tone 3A) at a comfortable level (~ 70 dB SPL).

Behavioral test: After the NIRS recording, the subjects participated in an identification test of Japanese long/short categories. The stimuli used in this test were the stimuli A–D and three stimuli whose second vowels were 173 ms, 206 ms and 228 ms in duration. The subjects were instructed to listen to the stimuli repeated 16 times each in a random order played back from a PC through headphone, and to identify whether the second vowel was phonologically long or not. Inter-stimulus interval (ISI) was initially 2 s and adjusted as each subject needed during practice: this adjusted ISI was referred to as a requested time for judgment (RTJ). The practice session presented some of /mama/ or /mama:/ stimuli that native Japanese judged long or short at 100% of accuracy. More than 70% of correct responses were required until RTJ for each subject was determined. The PB between the long and short vowels was defined as the vowel length where the fitted identification curve crossed 50% of the responses long. The curve was fit by the least square method weighted by the Müller Urban process [16].

Data processing: The concentrations of oxygenated (oxy-), deoxygenated (deoxy-) hemoglobin (Hb) and total Hb were calculated from the absorbance changes of 780 and 830 nm laser beams sampled at 10 Hz. After discarding the blocks with artifacts, the Hb concentrations of the remaining blocks were averaged 5 or 6 times synchronously to the target blocks, and smoothed over a 5 s moving average window.

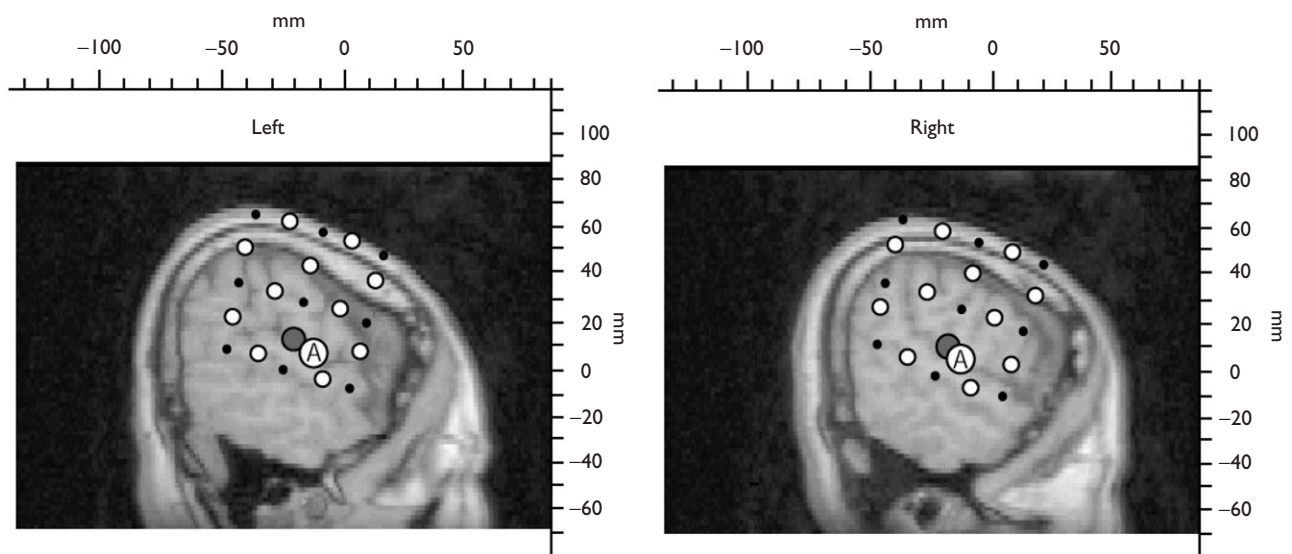


Fig. 1. Locations of the NIRS probes and the channels. NIRS probe positions (black dots) and the centers of the measurement channels (white circles) are superimposed onto the parasagittal MR brain images. The lateral posterior borders of the Heschl gyri are labeled A. The channels with the maximal responses are shown with gray circles. Positive is towards the anterior and the vertex in the scales.

The response peaks of the averaged target blocks were measured, and the z-scores of these peaks were calculated against the 10s baseline period just before the target block. A response with a peak z score > 1.96 ($p < 0.05$, uncorrected for multiple comparison) was considered a significant response.

RESULTS

NIRS results: The largest hemodynamic changes to the target stimuli were frequently found in the auditory area. In contrast, the responses observed in parietal and frontal area were small and varied in their temporal and amplitude pattern in all the subjects although limited parts (inferior part) of the frontal and parietal lobes were sampled (Fig. 1). One of the auditory channels that demonstrated the maximal peak responses on each side was selected for statistical analysis. Averaged coordinates for the channels chosen were $y = -13 \pm 2.8$ and $z = 9.2 \pm 2.8$ mm on the left side and $y = -13 \pm 4.1$ and $z = 11.7 \pm 2.7$ mm on the right side (only y and z coordinates are measured on parasagittal projection because the precise recording depths/or x, are not known). The hemodynamic responses peaked 6–14s after the target stimulus onset.

Figure 2a shows the averaged total Hb response peaks to the target stimuli for the three conditions of long/short

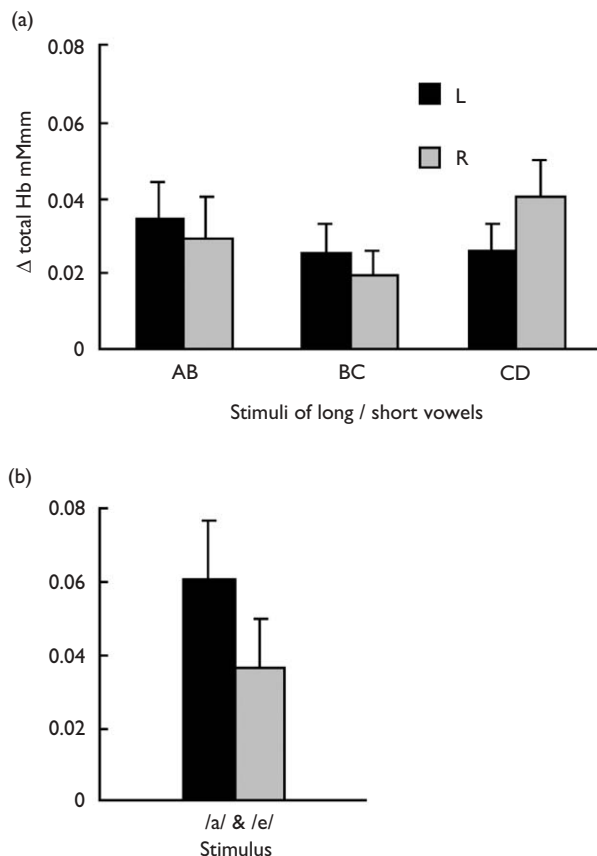


Fig. 2. Averaged values of the total Hb changes in the left auditory area (L) and the right auditory area (R). The stimulus conditions were (A) AB, BC and CD consisted of long/short vowels and (B) /a/ and /e/ stimuli. Error bars indicate s.e.

vowels. Two-way analysis of variance with condition (AB, BC and CD) and side (left and right) as factors revealed no significant main effects or interaction ($F(2,36) = 1.38$, $p = 0.27$; $F(2,36) = 1.19$, $p = 0.32$). There were at most two or three out of seven subjects showing significant peak hemodynamic responses for any of the three conditions (Fig. 3a), with no consistent hemispheric differences. Under the condition of distinctive phonemic contrast with /itta/ - /itte/ (/a-e/) stimuli, the averaged total Hb responses to the target stimuli were larger (Fig. 2b) than those for the long/short vowels. The peak hemodynamic responses were significant for all subjects, with five subjects in the left and the two remaining in the right or both sides of the auditory area (Fig. 3b).

To compare the response magnitude in the two different across-category conditions, two-way repeated measures ANOVA was performed with phonemic condition (/a-e/ vs BC in long/short vowels) and side (left and right) as factors. A significant main effect was observed for the phonemic condition ($F(1,24) = 5.08$, $p < 0.05$) with *post-hoc* Fisher's PLSD test, indicating that the /a-e/ phonemic contrast elicited the larger responses than the long/short vowel contrast ($t = 2.16$, $p = 0.02$). The leftward laterality was only significant for the /a-e/ contrast, and the laterality index calculated from $(L-R)/(L+R)$, where L and R are

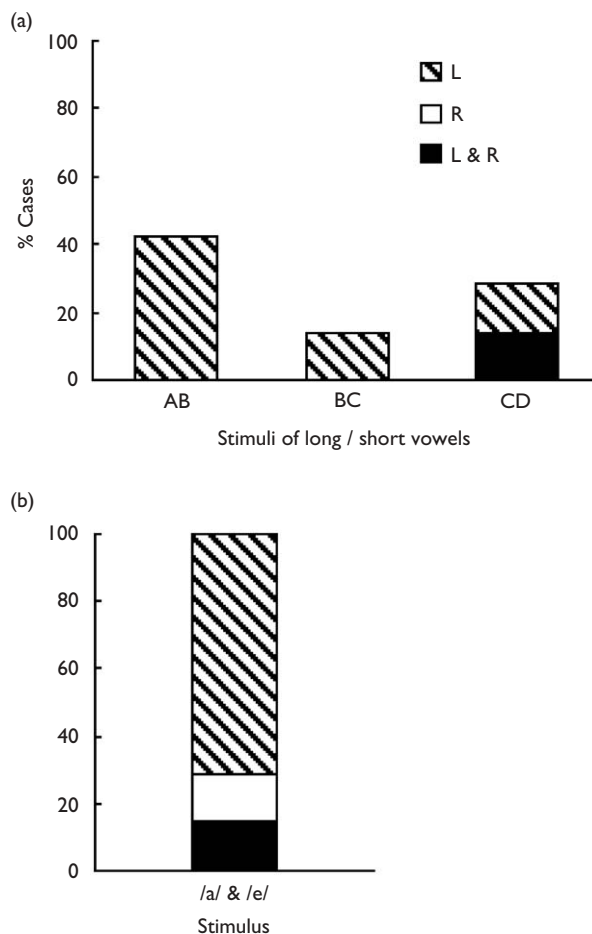


Fig. 3. Percentage of subjects who showed significant responses according to the stimulus conditions in the left (L) hemisphere, the right (R) hemisphere and both hemispheres (L + R). The stimulus conditions were (A) AB, BC and CD consisted of long/short vowels and (B) /a/ and /e/ stimuli.

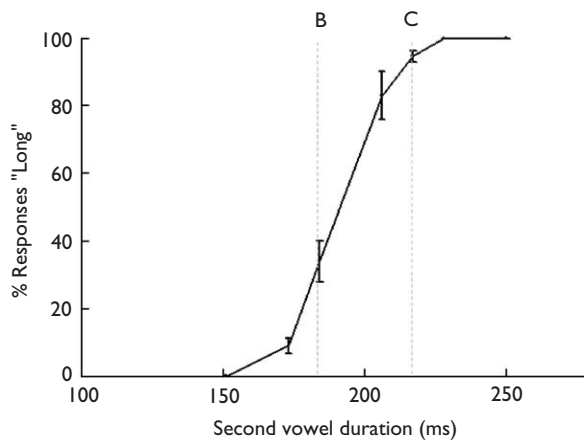


Fig. 4. Averaged identification rate of /mama-/mama:/ continuum for the Korean subjects. Durations of stimuli B and C are indicated above the graph. The PB is estimated between stimuli B and C. Error bars indicate s.e.

maximal total Hb responses on the left and right side respectively was 0.32 (zero-test, $p = 0.026$) for the /a-e/ and 0.084 (zero-test, $p > 0.05$) for the long/short vowel contrast.

Behavioral results: The behavioral test revealed that the averaged PB value was 191 ± 8 ms ($n = 7$; Fig. 4), which fell between the stimuli B and C. Here the slope of identification curve was steepest which is suggestive of categorical perception. The averaged RTJ was 3.4 ± 0.37 s ($n = 7$).

DISCUSSION

Language-specific phoneme representations were revealed by the mismatch negativity (MMN) which reflected a phonological difference between frequent and infrequent stimuli [3,4]. As hemodynamic measurements with PET and NIRS can capture such change detection responses [17], the present study was also able to measure neural responses in L2 learners. The native Korean speakers who had become fluent in Japanese after puberty, showed differential cortical activities in response to the two kinds of Japanese phonemic contrasts. For the durational contrast, neither phoneme-specific nor left-dominant responses were observed, although behavioral tests showed categorical perception indistinguishable from that of native Japanese speakers [18]. In contrast, the across-category stimuli of /a/ and /e/ contrast elicited a larger response in the left auditory area than in the right, which is suggestive of linguistic perception. Furthermore, a pilot behavioral test showed that Korean subjects almost perfectly identified /a/ and /e/ contrast.

The observed differences in the cortical responses to the two types of contrasts could be accounted for by the distinctiveness of the respective contrasts in the subjects' native language. The /a/ and /e/ contrast is distinctive in the Korean phonology and shares a relatively similar phonetic space to that of the Japanese contrast [19]. Significant cortical responses to this contrast were observed in all subjects, although the experimental stimuli were apparently in a L2 context. These neural responses were consistent with the Japanese results that also demonstrated a leftward laterality [17]. On the other hand, the long/short durational contrast, which current phonological system in

Korean lacks, i.e. non-distinctive contrast, did not evoke the phoneme-specific responses as the /a/ vs /e/ contrast did for the same subjects. This was in clear contrast to the native Japanese results where the across-category, but not within-category, durational contrast elicited a larger response in the left hemisphere [18], indicative of linguistic processing. A question remains why the Koreans showed different neural response patterns to the long/short vowels even though their identification results were similar to those of the native listeners. The differences in response time between L1 and L2 speakers in the behavioral test may shed light on this question: RTJ was longer for the Korean group (RTJ = 3.4 s) than for the Japanese group (RTJ = 2.2 s, reanalysis of the data in [18]). The RTJ difference may reflect a different strategy from that of the L1 group for identifying phonemes which may result from the lack of phonemic representations in L2 Korean learners as revealed from the current NIRS results. These suggest that late L2 learners, who seem to have acquired L2 phonemes behaviorally, as shown by their identification scores, may still employ different neural process in dealing with L2 phonemes. However, this assumption should be further tested by detailed behavioral data. It has been a controversial issue whether individual bilinguals employ different neural substrates for L1 and L2 use. Some researchers have suggested that different networks are employed for L1 and L2 processing [20–22], others have reported the same neuronal substrates subserving L1 and L2 [23,24]. Although the discrepancy may be partly due to the differences in their task and/or L2 proficiency of the subjects [25], the present study proposes the following regarding the neuronal networks for L2 phonemic processing. Whether the same or different cerebral networks are recruited for L1 and L2 phonemes depends on the distinctiveness of the contrast and the position of the target phoneme in the L1 and L2 phonetic spaces. When a target phonemic contrast of L2 is also distinctive in L1 and fills the same or similar positions as in the L1 phonetic space, the same cerebral substrates for the L1 phonemes may be used in processing L2 phonemes. This was most likely the case with the present control stimuli (/a/ and /e/). However, when a contrast is not shared by the L1 and L2 phonetic space, cerebral networks that are not necessarily specific to phonemic discrimination could be employed, as was observed for Koreans listening to the long/short phonemic contrast.

CONCLUSION

The current work provides crucial evidence in understanding the neural processes of L1 and L2 phonemes. The hemodynamic responses of the Koreans to the two types of L2 phonemic contrasts were different according to the phonological properties of the contrasts in L1, i.e., the distinctiveness and phonetic status in L1. These findings imply that the neuronal networks subserving L1 and L2 are either shared or different according to the mutual relationship between the phonemic properties in L1 and L2.

REFERENCES

1. Jusczyk PW. *The Discovery of Spoken Language*. Cambridge: The MIT Press; 1997.
2. Miyawaki K, Strange W, Verbrugge R, Liberman AM, Jenkins JJ and Fujimura O. An effect of linguistic experience: the discrimination of [r]

- and [l] by native speakers of Japanese and English. *Percept Psychophys* **18**, 331–340 (1975).
3. Dehaene-Lambertz G. Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport* **8**, 919–924 (1997).
 4. Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huottilainen M, Iivonen A *et al.* Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* **385**, 432–434 (1997).
 5. Koyama S, Gunji A, Yabe H, Oiwa S, Akahane-Yamada R, Kakigi R *et al.* Hemispheric lateralization in an analysis of speech sounds: left hemisphere dominance replicated in Japanese subjects. *Brain Res Cogn Brain Res* **10**, 119–124 (2000).
 6. Rivera-Gaxiola M, Csibra G, Johnson MH and Karmiloff-Smith A. Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behav Brain Res* **111**, 13–23 (2000).
 7. Simos PG, Diehl RL, Breier JL, Molis MR, Zouridakis G and Papanicolaou AC. MEG correlates of categorical perception of a voice onset time continuum in humans. *Brain Res Cogn Brain Res* **7**, 215–219 (1998).
 8. Imaizumi S, Ito H, Tamekawa Y, Deguchi T and Mori K. Plasticity of non-native phonetic perception and production: a training study. *Proc ICSLP* **5**, 1887–1890 (1998).
 9. Chomsky N and Halle M. *The Sound Pattern of English*. Cambridge: The MIT Press; 1968.
 10. Minagawa-Kawai Y, Maekawa K and Kiritani S. Effects of pitch accent and syllable position in identifying Japanese long and short vowels: comparison of English and Korean speakers. *J Phonet Soc Jpn* **6**, 88–97 (2002).
 11. Toda T. Interlanguage phonology: Acquisition of timing control in Japanese. *Austr Rev Appl Ling* **17**, 51–76 (1994).
 12. Watanabe E, Maki A, Kawaguchi F, Takashiro K, Yamashita Y, Koizumi H *et al.* Non-invasive assessment of language dominance with near-infrared spectroscopic mapping. *Neurosci Lett* **256**, 49–52 (1998).
 13. Yamashita Y, Maki A and Koizumi H. Near-infrared topographic measurement system: imaging of absorbers localized in a scattering medium. *Rev Sci Instrum* **67**, 730–732 (1996).
 14. Imagawa H and Kiritani S. High speed speech analysis system using a personal computer with DSP and its application to pronunciation training. *Ann Bull Res Inst Logoped Phoniatr Univ Tokyo* **23**, 175–164 (1989).
 15. Imaizumi S, Mori K, Kiritani S, Hosoi H and Tonoike M. Task-dependent laterality for cue decoding during spoken language processing. *Neuroreport* **9**, 899–903 (1998).
 16. Müller GE. *Die Gesichtspunkte und die Tatsachen der Psychophysischen Methodik*. Wiesbaden: Bergmann; 1904.
 17. Furuya I and Mori K. Cerebral lateralization in spoken language processing measured by multi-channel near-infrared spectroscopy (NIRS). *Brain Nerve* **55**, 226–231 (2003).
 18. Minagawa-Kawai Y, Mori K, Furuya I, Hayashi R and Sato Y. Assessing cerebral representations of short and long vowel categories by NIRS. *Neuroreport* **13**, 581–584 (2002).
 19. Umeda H. Some problems on the contrastive study of Korean and Japanese. *J Jpn Lang Teaching* **48**, 31–42 (1982).
 20. Albert ML and Obler LK. *The Bilingual Brain: The Neuropsychological and Neurolinguistic Aspects of Bilingualism*. New York: Academic Press; 1978.
 21. Dehaene S, Dupoux E, Mehler J, Cohen L, Paulesu E, Perani D *et al.* Anatomical variability in the cortical representation of first and second language. *Neuroreport* **8**, 3809–3815 (1997).
 22. Kim KH, Relkin NR, Lee KM and Hirsch J. Distinct cortical areas associated with native and second languages. *Nature* **388**, 171–174 (1997).
 23. Berthier ML, Starkstein SE, Lylyk P and Leiguarda R. Differential recovery of languages in a bilingual patient: a case study using selective amygdala test. *Brain Lang* **38**, 449–453 (1990).
 24. Nakada T, Fujii Y and Kwee IL. Brain strategies for reading in the second language are determined by the first language. *Neurosci Res* **40**, 351–358 (2001).
 25. Perani D, Paulesu E, Galles NS, Dupoux E, Dehaene S, Bettinardi V *et al.* The bilingual brain: proficiency and age of acquisition of the second language. *Brain* **121**, 1841–1852 (1998).

Acknowledgements: The authors thank D. Erickson for assistance with English in the earlier version of the manuscript. This work was supported by Japan Society for Promotion of Science (No. 8484), Research grants from the Ministry of Health, Labor and Welfare of Japan (K.M.) and Japan Aging Research Foundation (Y.M.).