

# The Effect of the Katakana Writing System on Second Language Pronunciation: A Suggestion for Katakana Sound Instruction in English Textbooks

Hideo IMAI

## Abstract

The ultimate aim of this study is to consider whether katakana-based pronunciation of wasei eigo debases Japanese EFL learners' pronunciation of English. Advanced and non-advanced Japanese EFL learners who do not have prior knowledge of phonetics participated in a series of auditory English speaking experiments that investigated whether voiced velar nasal /ŋ/ was correctly pronounced even when they visually perceived katakana. The significant finding of this study was that non-advanced EFL learners correctly produced the target sound when they visually perceived Japanese semantic words including the velar nasal that is often replaced with wasei eigo (i.e., 指輪 “yubiwa” which means *ring* is often replaced with the wasei eigo リング “ringu”). However, non-advanced learners' pronunciation of /ŋ/ worsened and an unnecessary /g/ was added when they visually perceived simple katakana-based wasei eigo (i.e., スピーキング “*supiikingu*” which means *speaking* is katakana-based wasei eigo that is widely used in place of the Japanese word 話す “*hanasu*”). The result implies that the pronunciation of /ŋ/ which does not have L1 phonological contrast worsens when non-advanced learner perceive katakana because they rely heavily on the katakana sound system when they study L2 pronunciation. Thus, it is possible to postulate that katakana-based pronunciation guides should be used with caution in second language teaching situations.

## 1. Introduction

Recently, the Ministry of Education, Culture, Sports, Science and Technology (MEXT, 2002, 2008, 2010) of Japan has begun to adopt communicative language teaching methods by focusing on the way the language is used rather than the grammatical structure of the language (Savignon & Wang, 2003) to further develop the English speaking ability of Japanese learners in the classroom. This new English teaching method made both Japanese students and teachers more aware of English pronunciation. In order to comply with these directives, some schools have started using textbooks, which include katakana-based pronunciation guides in order to give students a better understanding of the English sound systems.

However, the katakana-based pronunciation of *wasei eigo*, which is derived from English loan words, often lacks phonological contrast in Japanese (first language, L1), which can lead to great difficulty when trying to discriminate and produce English (second language, L2) phonemes as explained by Kohmoto (1969) in his book *New English Phonology* in which he states, “the key to discriminating the degree of difficulty in foreign language learning lies in the comparison of the foreign with the native language” (p.114). For example, Japanese people mispronounce /sɪŋər/ as /ʃɪŋgər/ due to the L1 katakana influence of シンガー. Furthermore, auditory perception errors caused by speakers’ mispronunciation or transfer of L1 can lead to an insurmountable breakdown in human communication (Uchida, 2008). This means that accurate English pronunciation instruction should be regarded as a fundamental element in the English language classroom, and the implementation of katakana-based pronunciation instruction in English textbooks must be reexamined because it could undermine the student’s ability to accurately learn and reproduce English sounds.

To determine the cause of a pronunciation error commonly made by Japanese learners of English, this study examines the difference in pronunciation between /ɪ/ and /g/ through a series of auditory speaking experiments. Using data extracted from these experiments, the researcher will show how katakana-based knowledge of the test words negatively affects the learners’ ability to correctly discriminate and produce English sounds.

English learners of different abilities participated in this experiment to reveal (1) the effects of the Japanese katakana sounds *グ* (/gu/) or *ガ* (/ga/) in the English pronunciations such as *ring* /rɪŋg/ and *singer* /sɪŋgər/ and (2) the correlation between the level of English proficiency and the fluency of pronunciation. The researcher hypothesized that due to the proliferation of wasei eigo in the Japanese language, participants who do not have high English proficiency would tend to mispronounce /ŋ/ as /g/ when they saw the katakana リング or シンガー (e.g., /rɪŋg/, /sɪŋgər/).

## 2. Basic Knowledge of Nasal Sounds

The study of phonetics categorizes three nasal sounds as separate phonemes: /m/, /n/, and /ŋ/. First, when we pronounce voiced bilabial nasal /m/, our two lips are completely closed to pronounce words such as *many*, *Rome*, and *mix*. Usually, bilabial closure occurs at the beginning of the word (Ladefoged & Johnson, 2011). The Japanese sound system has /m/ which make it relatively easy for Japanese people to discriminate between the sound contrast /m/ and /n/ because of the existence of phonological contrast. Second, voiced alveolar closure /n/ is produced when we pronounce *night*, *no* and *nobody*. When we pronounce /n/, the tip of the tongue touches the alveolar ridge, a spot between the hard palate and the teeth. Similar to the /m/ sound, Japanese speakers can discriminate between /m/ and /n/ because both sounds have L1 phonetic contrast. The third nasal is the voiced velar nasal /ŋ/. To produce this sound the back of the tongue touches near the soft palate in the very back of the mouth, and is found in such words as *playing*, *ring* and *singer*.

Although velar nasal /ŋ/ exists in Japanese, the strict pronunciation of the nasal quality of the sounds is not as strictly adhered to as in the past and a plosive /g/ can often be heard in its place (Takayama, 2015; Vance, 1987). For example, Japanese learners often mispronounce *sing* /sɪŋ/ as [sɪŋg] or *cigarette* /sɪgəˈrɛt/ as [sɪŋərˈɛt] (Matsuzaka, 1986). Thus, different from the other two nasals, phonological contrast of the velar nasal /ŋ/ would be less obvious, and could result in an inability to accurately discriminate between both /ŋ/-/m/ and /ŋ/-/n/.

The qualities of these three nasal sounds can be determined by using sound analysis software such as *Praat* (developed by Paul Boersma and David Weenink). All of our spoken sounds, including nasals, contain a certain amount of sound spectrum, and the spectral peaks of the sound spectrum are called formant. Analyzing each of the three nasal formants, we can clearly figure out numerical differences among each formant based on the frequency of the speaker's pronunciation. For example, the first formant (F1), which gives an indication of up-down movement of the tongue, does not record a remarkable difference in the size of the mouth opening when native English speakers produce nasal sounds. When measured by auditory analysis software, all of these three nasal consonants are usually very low F1 centered at about 250Hz. (Ladefoged & Johnson, 2011). On the other hand, each nasal has different second formants (F2) that serve as a barometer for forward and backward tongue movement in our mouths. Recasens (1988) noted that F2 is between 1000-1500Hz for /m/, 1500-2000Hz for /n/ and around 2000Hz for /ŋ/. Ergo, when we pronounce the velar nasal /ŋ/, a low F1 and high F2 can be observed simultaneously.

### **3. Velar Nasal Perception and Pronunciation Errors for Japanese English Learners**

L2 pronunciation errors directly relate to perceptual problems caused by the learners' natural tendency to translate L2 sound phonemes into familiar L1 phonemes already established. This results in the learners wrongly mimicking sounds at the beginning of L2 acquisition (O'Connor, 1980). Therefore, it is important to probe into both listening and pronunciation experiments conducted previously.

Nozawa and Sang (2012) examined three different types of English learners (Japanese, English and Korean) to clarify whether they can discriminate nasal sounds of /m/, /n/, /ŋ/ in a listening experiment. They had 14 native speakers of each language hear recorded words which all ended in final syllable nasals. They predicted that Japanese speakers' identification accuracy of syllable-final non-native nasals would be lower than that of English and Korean speakers because recently, Japanese people rarely discriminate velar nasal due to the lack of phonological contrast in L1 (Takayama, 2015; Vance, 1987).

As Nozawa and Sang expected, the result showed exactly what they had hypothesized. Although English and Korean listeners chose the correct answer approximately 90% of the time for all three nasals, Japanese listeners chose the right answer only about 60% of the time. The results also proved the error patterns of Japanese listeners. They estimated that Japanese listeners made more errors when discriminating /n/ and /ŋ/, and /ŋ/ was often mistakenly heard as /n/. In addition, Aoyama (2003) also proved that syllable-final velar nasal /ŋ/ is more difficult for Japanese learners than Korean learners. Therefore, it is possible to state that Japanese learners of English have particular difficulty in discriminating velar nasal due to the deficiency of phonological contrast system in L1.

The occurrence of nasal error that transfers /ŋ/ as /g/ also appears in the speaking experiment. Misawa and Mari (1995) recorded 45 Japanese students' voices when they pronounced English words including velar nasal /ŋ/ (e.g., *studying*, *long*). The result was that approximately half of the students incorrectly combined both plosive and nasal sounds. Misawa and Mari concluded that the appropriate use of voiced velar plosive /g/ and voiced velar nasal /ŋ/ have been inconsistently acquired for Japanese learners of English.

## **4. The Study**

Previously, listening and pronunciation studies have shared the idea that Japanese learners of English do not correctly perceive and pronounce the velar nasal /ŋ/ and sometimes change the original sound of /ŋ/ into /g/ because of the lack of phonological contrast in L1 (Aoyama 2003; Misawa & Mari 1995; Nozawa & Sang 2012). In addition to this, the present study primarily explores whether the Japanese katakana system or wasei eigo, which widely dominates Japanese society, debases the L2 sounds system.

### **4.1 Research Questions**

To examine the katakana effect towards Japanese learners of English as a foreign language (EFL), the following two research questions were posed:

**Research Question (1):** Do Japanese katakana sounds such as *グ* (/gu/) or *ガ* (/ga/) debase English velar nasal pronunciation?

**Research Question (2):** Can Japanese EFL learners with high English proficiency accurately pronounce English words even when they see katakana-based *wasei eigo*?

## 4.2 Hypothesis

According to previous research, the lack of phonological contrast in /ŋ/ L2 discrimination, in both listening and speaking, lead Japanese speakers to pronounce /g/ for /ŋ/. Therefore, the likelihood of low-proficiency Japanese learners of English to correctly pronounce the English /ŋ/ and /g/ has decreased due to the increase of katakana and *wasei eigo* in the Japanese language. Eventually, a distortion of F1 or F2 will probably be acquired by non-advanced English learners when compared to advanced learners.

## 4.3 Participants and Procedure

13 participants (7 male/6 female) who did not have prior knowledge of phonology and speech disorders were chosen from undergraduate university students. They were categorized into two groups: advanced English learners and non-advanced English learners depending on their TOEIC score, which can be an indicator of their English proficiency. Advanced learners were categorized as having a TOEIC score of between 525 and 900, while non-advanced learners had scores of lower than 500. Each participant was randomly asked to demonstrate the following two speaking experiments (see Figure 1) in a soundproofed room:

### 4.3.1 Experiment 1: Katakana Word Task

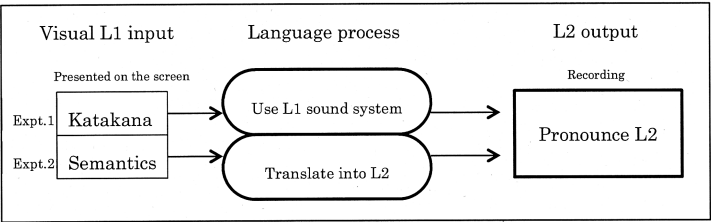
Six katakana words, which are commonly known as *wasei eigo* were selected as visual stimuli and presented on the computer screen. Participants were required to speak as if they were native speakers and pronounce English sounds for each katakana word presented on the screen. For example, if the katakana word ミーティング which represents *meeting* in English appeared on the screen, they were required to say /mí:tɪŋ/

the same as a native speaker’s English pronunciation. To avoid understanding the aim of this experiment, another six katakana words were added to the list as distractors. Twelve katakana words were presented in total (see Appendix. 1).

### 4.3.2 Experiment 2: Japanese Semantic Task

Six words in Japanese, which are often replaced with wasei eigo, include /ŋ/ phoneme were selected as visual stimuli and presented on the computer screen. Participants were asked to produce the English translation of each word. In this case, because each stimulus was different from the original English pronunciation, they instinctively changed the input semantic signals into the appropriate English translation. This language process is quite similar to real English conversation in that it requires a code-switching process in their brain without any katakana information obstacle. For instance, if the participant sees 会議 “*kaigi*” which means meeting in English, they first translate this semantic word into English, and then produce native-like English pronunciation. An additional six unrelated distractor words were included to get rid of the participants’ expectation of the target of the experiment (see Appendix. 2).

Figure.1 Experiment detail based on expected language process



*Note:* Two separate language processes can be predicted in the current two experiments. The Japanese semantic task may need a natural language translation process when compared to the katakana word task, which uses the Japanese participants’ prior knowledge of the L1 sound system.

## 5. Data Acquisition

The auditory recording software *Speech Recorder* (developed by the Institute of the Phonetics and Speech Processing of Ludwig-Maximilian University in Munich,

Germany) allows researchers to show a link between the stimulus displayed on the screen and the recording of the participants' vocalization of the stimulus (Draxler, 2011). The data extracted from *Speech Recorder* were then analyzed on the auditory analysis software *Praat*. In the current experiment, F1 and F2 were carefully analyzed to observe whether /ŋ/ was correctly pronounced, or not. Theoretically, the accurate formant would be around 250Hz for F1 and 2000Hz for F2 (Ladefoged & Johnson, 2011; Recasens, 1988).

## 6. Results

### 6.1 Experiment 1

The result of the current experiment that analyzed nasal phonemes showed the negative effect of katakana, especially for non-advanced English learners. Even though katakana words ending in ゴ or ガ which lead Japanese English learners subconsciously to produce the incorrect pronunciation by adding /g/ after the velar nasal /ŋ/. In contrast, the advanced learners' pronunciations were correct. The acquired data of low F1 ( $M=644\text{Hz}$ ) and high L2 ( $M=2000\text{Hz}$ ) means that the back of the tongue clearly hit the velar spot and accurate /ŋ/ was produced (Recasens, 1988; Ladefoged and Johnson 2011). In contrast, non-advanced learners had a tendency to add unnecessary /g/ when they perceived katakana stimuli in experiment 1, making F1 ( $M=1711\text{Hz}$ ) nearly the same value as F2 ( $M=2080\text{Hz}$ ). More precisely, a huge inclination differences was not visible between F1 and F2 in experiment 1 meaning that non-advanced learners were unconsciously influenced by katakana (see Figure 2).

### 6.2 Experiment 2

Interestingly, this katakana effect was less visible when non-advanced learners perceived semantic visual stimuli, resulting a high inclination between F1 ( $M=981\text{Hz}$ ) and F2 ( $M=2000\text{Hz}$ ) in experiment 2. When they perceived Japanese semantic words and translated them into L2, their F1 declined compared to experiment 1. This pronunciation improvement was presumably because they did not see katakana on the screen that eventually made their pronunciation incorrect due to the current dominance



of katakana-based wasei eigo. In contrast, the advanced learners' F1 ( $M=500\text{Hz}$ ) were not so significantly changed in either experiment 1 or experiment 2. Also, a high F2 ( $M=2000\text{Hz}$ ) was recorded from the advanced learners in experiment 2 meaning that they produced the correct English sounds when they visually perceived Japanese semantics (see Figure 3).

Figure. 2 Katakana word task (Expt. 1)

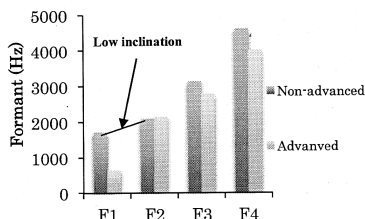
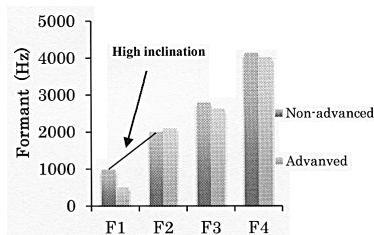


Figure. 3 Japanese Semantic task (Expt. 2)

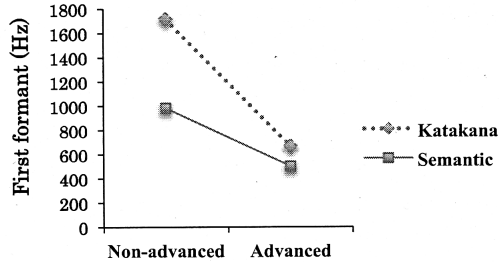


*Note:* Formant inclinations between F1 and F2 / $\eta$ / were obvious in the semantic task (Expt.2) for non-advanced English learners because they were able to pronounce the words without the katakana letter effect that unconsciously added an extra /g/ after the velar nasal / $\eta$ /.

### 6.3 Experiment 1 and Experiment 2

To provide a clearer definition of the katakana effect towards English pronunciation, only F1 was extracted from all formant data and reanalyzed to observe the differences between katakana and semantic visual perceptions (see Figure 4). Notably, F1 acquired from advanced learners were stabilized in both katakana and semantic visual stimuli. However, acquired F1 from non-advanced learners decreased when semantic stimuli were presented, meaning their pronunciation became much closer to native pronunciation. Regarding these outcomes, although both stimuli were presented in Japanese, katakana-based pronunciation was worsened for L2 pronunciation when compared to the semantic stimuli.

Figure. 4 Extracted F1 data from both experiments



*Note:* The F1 results were averaged and a slope line was created to visualize whether English pronunciation of /ŋ/ was affected by Katakana. This figure directly shows that katakana input creates a large inclination between advanced and non-advanced Japanese learners of English when comparing the semantic inputs. It means that non-advanced learners' pronunciation was greatly affected by katakana input.

## 7. Discussion

As previous phonological studies have pointed out, non-native contrast, or the lack of phonological contrast in sounds such as /g/ and /ŋ/ for Japanese people are difficult to discriminate and pronounce (Aoyama 2003; Misawa & Mari 1995; Nozawa & Sang 2012). This phenomenon is known for the markedness differential hypothesis (MDH) (Eckman, 1977) in the field of second language acquisition (SLA). The basic idea of MDH came from the contrastive analysis approach that compares two different language structures. For instance, if a markedness relationship exists between two languages and language A has marked structure whereas language B does not, then the speaker of B learning A should have more difficulty understanding the structure of language A (Carlisle, 1988). Put simply, the sounds which are less commonly used in daily life (i.e., marked), are usually considered to be difficult to acquire as L2 phonemes. On the other hand, the sounds which are widely used because of their regularity and productivity (i.e., unmarked) are relatively easy to acquire as L2 phonemes. Hence, if the daily used unmarked structure of katakana letters such as *ヶ* was used as a guide for L2 pronunciation, it is difficult for Japanese learners of English to produce accurate velar nasal sounds, because /ŋ/ is a marked sound for native Japanese people and it has non-native contrast in Japanese language.

When applied to this study, the MDH theory shows that although non-advanced learners knew how to pronounce the marked sound of L2 /ŋ/ accurately (see the result of Expt.2), they mispronounce the English word when they perceive a marked structure of katakana-based *wasei eigo* (see the result of Expt.1). This result shows a critical notion of the priority order of the language sound process. For non-advanced learners, whether or not they have prior knowledge of the correct pronunciation of the marked sound of /ŋ/, they mispronounced it when they perceived katakana, implying that their priority of pronunciation rule is katakana.

When thinking about a child playing on a seesaw as a metaphorical example, this priority phonological rule becomes clear. Non-advanced learners rely heavily on the unmarked structures. When unmarked structure was used to describe marked structure (e.g., to describe /ŋ/ by using katakana as a hint), the unmarked structure goes up and marked structure goes down. From this perspective, it is possible to find an additional implication based on the MDH theory that if a markedness relationship exists in two language and both marked phoneme and unmarked phoneme can be used at the same time to understand the target language sound system, non-advanced learners are unconsciously apt to use unmarked structures to understand the target language phonemes rather than using marked structures.

This finding may give a pedagogical implication when teaching English sound structure. Generally, L2 learning sometime uses the L1 structure as a support tool for giving hints to students when trying to understand the phonological rules of L2. This method of teaching is presumably following the innate human language process because the results of the present study showed that language learners rely on their unmarked L1 structure when they learn target unmarked structure language. However, using katakana-based sound instruction to teach the L2 sound system is sometimes problematic when considering L2 phonemes which do not have any phonological contrast in L1. As discussed in this study, if the sound does not have phonological contrast and has a marked structure, L2 learners' erroneous pronunciations occur due to the natural tendency of using unmarked structures in L1 (i.e., katakana). This may lead to a break down in the construction of L2 phonological rules especially at the beginning of language learning. Put simply, when non-advanced English learners see Katakana

instruction in their textbook, their pronunciation is nearly the same as the katakana sounds. Thus, L2 instructors should be cautious when using katakana-based sound instruction especially when teaching non-native contrast phonemes, because it may warp the students' L2 sound rules.

## 8. Conclusion

The current speaking experiment discovered the effect of the katakana writing system on second language pronunciation and proved that the Japanese learners of English who do not have high English proficiency tend to mispronounce English velar nasal sounds when they visually perceive katakana letters. In addition, the present study showed that non-advanced learners rely excessively on unmarked structure (i.e., katakana-based sound instructions) when they learn the second language that includes marked structures (i.e., non-native contrast sound such as /ŋ/). This can create a negative effect for language learners because L1 reliance can worsen L2 phonological rules especially when L1 does not have phonological contrast, or L2 has a marked structure.

Although this experiment only focuses on the velar nasal /ŋ/ by using scientific evidence, many difficult sound pairs still remain such as /b/-/v/ and /s/-/ʃ/. Thus, more research is needed to examine whether or not the katakana writing system distorts English pronunciation. Still, the present study gives us fragmental proof that when teaching second language pronunciation, we must be very careful when using katakana instruction, especially when L1 sounds do not have phonological contrast such as /ŋ/ and urges us to reexamine the true effect of katakana instruction in English textbooks.

## References

- Aoyama, K. (2003). Perception of syllable-initial and syllable-final nasals in English by Korean and Japanese speakers. *Second Language Research*, 19(3), 251-265.
- Carlisle, R. S. (1988). The Effect of Markedness on Epenthesis in Spanish/English Interlanguage Phonology. *Issues and Developments in English and Applied Linguistics (IDEAL)*, 3, 15-23.

- Draxler, C. (2011). Speech recorder quick start and user manual. *Institute of Phonetics and Speech Processing, University of Munich, Tech. Rep.* Retrieved from [www.speechrecorder.org]
- Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language learning*, 27(2), 315-330.
- Kohmoto, S. (1969). *New English phonology: A contrastive study of English and Japanese pronunciation*. Tokyo: Nan'un-do.
- Ladefoged, P., & Johnson, K. (2011). A course in phonetics Wadsworth. *Cengage Learning, Scarborough*.
- Matsusaka, H. (1986). *Eigo Onseigaku Nyumon*. [An Introduction to English Phonetics.]. Tokyo: Kenkyu-sha.
- MEXT. (2002). Developing a strategic plan to cultivate 'Japanese with English abilities'. Retrieved from [http://www.mext.go.jp/english/]
- MEXT. (2008). *Chuga kekou gakushu shidou yoryo kaisetsu gaikokugo ben* [The guide to course of study: Foreign language]. Tokyo: Kairyudo
- MEXT. (2010). *Koutou gakkou gakushu shidou yoryo kaisetsu gaikokugo ben eigo ben* [The guide to course of study: Foreign language and English]. Tokyo: Kairyudo.
- Misawa, M., & Mari, P. (1995). The Role Of Japanese Pronunciation in English Phonetics Class. *Bulletin of Shinshu Honan Women's Junior College*, 12: 21-31.
- Nozawa, T., & Cheon, S.Y. (2012). The identification of Nasals in a Coda Position by Native Speakers of American English, Korean and Japanese. *Phonetics Society of Japan 16(2)*, 5-14
- O'Connor, J. D. (1980). *Better English Pronunciation*. Cambridge University Press.
- Recasens, D. (1988). Phonology, and speech production and perception: the case of nasal stops. In W.U. Dressler, H.C. Luschützky, O.E. Pfeiffer, & J.R. Rennison (Eds.). *Phonologica 1988: Proceedings of the 6th international phonology meeting*. Cambridge: Cambridge University Press, 229-238.
- Savignon, S. J., & Wang, C. (2003). Communicative language teaching in EFL contexts: Learner attitudes and perceptions. *IRAL*, 41(3), 223-250.
- Takayama, T. (2015). Historical phonology. In H. Kubozono (Ed.), *Handbook of Japanese Phonetics and phonology Vol. 2* (pp. 621-650). Walter de Gruyter GmbH & Co KG.
- Uchida, H. (2008). A benefit of pronunciation training: How orally practicing unfamiliar phonemes can positively affect aural identification of these phonemes. *Tottri University of Environmental Studies*, 6, 39-48.
- Vance, T. J. (1987). *An Introduction to Japanese Phonology*, Albany, NY: State University of New York Press.

Appendix. 1 Visual stimuli used in experiment 1 (Katakana word task)

	<b>stimuli</b>	<b>correct</b>	<b>incorrect</b>
<b>1</b>	ミーティング	/mí:tɪŋ/	/mí:tɪŋg/
<b>2</b>	クリーニング	/klí:nɪŋ/	/klí:nɪŋg/
<b>3</b>	スピーキング	/'spi:kɪŋ/	/'spi:kɪŋg/
<b>4</b>	トレーニング	/tréinɪŋ/	/tréinɪŋg/
<b>5</b>	ハンガー	/hæŋə(r)/	/hæŋgə(r)/
<b>6</b>	クライミング	/kláimin/	/kláiminŋ/

Appendix. 2 Visual stimuli used in experiment 2 (Japanese semantic task)

	<b>stimuli</b>	<b>correct</b>	<b>incorrect</b>
<b>1</b>	長い	/lɔ':ŋ/	/lɔ':ŋg/
<b>2</b>	王様	/kíŋ/	/kíŋg/
<b>3</b>	指輪	/ríŋ/	/ríŋg/
<b>4</b>	歌手	/síŋə(r)/	/síŋgə(r)/
<b>5</b>	喫煙	/smóukɪŋ/	/smóukɪŋg/
<b>6</b>	春	/spríŋ/	/spríŋg/

*Note:* Six target stimuli were automatically presented on a computer screen for each trial. To avoid participants' awareness of the aim of the experiment, six other stimuli were presented as distractors.