



Cross-language Perception of Japanese Singleton and Geminate Consonants: Preliminary Data from Non-native Learners of Japanese and Native Speakers of Italian and Australian English

Kimiko Tsukada¹, Felicity Cox¹, John Hajek²

¹ Macquarie University, Australia

² The University of Melbourne, Australia

{kimiko.tsukada, felicity.cox}@mq.edu.au, j.hajek@unimelb.edu.au

Abstract

We examined the perception of Japanese consonant length contrasts (singleton vs geminate) in four groups of listeners: two groups of non-native learners of Japanese, three each at highly-advanced (NNJ1) and upper-intermediate (NNJ2) levels, native speakers of Italian (IT) (n=10) and Australian English (OZ) (n=8) with no knowledge of Japanese. Because Italian, like Japanese, uses consonant length contrastively but Australian English does not, we were interested in whether first language (L1) knowledge of consonant length might have an effect on the ability to perceive short and long Japanese consonants. The NNJ1 learners were more accurate in identifying Japanese singleton and geminate consonants than were the IT listeners who, in turn, were more accurate than the OZ listeners. The NNJ2 learners' results showed similarities and differences to the IT and OZ listeners. Our preliminary results suggest that L1 experience with consonant length may not necessarily guarantee accurate perception of consonant length in an unknown language. However, it may offer some advantage over lack of exposure to consonant length. In addition, the results for learner proficiency demonstrate that non-native learners need much time/practice before they clearly differentiate themselves from naïve listeners, reconfirming previous research that consonant length contrasts are difficult to acquire.

Index Terms: singleton/geminate, Japanese, Italian

1. Introduction

Speech sounds are stretched or compressed for various purposes. While it is common in language generally for some sounds to be physically shorter or longer than other sounds, some languages make use of this effect for contrastive purposes, i.e. to differentiate the meaning of words. "One of the most interesting languages in the way that it uses length is Japanese [1]", because it has length distinctions for both vowels (e.g. /i/ *stomach* vs /i:/ *good*) AND consonants (e.g. /ita/ *stayed, was/were* vs /ita:/ *said, went*). Vowel length is not contrastive in Italian, but consonant length is for a wide variety of phonemes (/p b t d k g m n r tʃ dʒ f v s l/) (e.g. /set/ *thirst* vs /set:/ *seven*) [2, 3, 4, 5]. Thus, both Italian and Japanese use consonant length (i.e. singleton vs geminate) for communicative purposes and native speakers of these languages possess two distinct consonant length categories in their L1s. Australian English, another language under investigation in this study, uses length to contrast at least two pairs of vowels, e.g. /kʌt/ *cut* vs /kʌ:t/ *cart*, /ʃed/ *shed* vs /ʃe:d/

shared [6, 7, 8]. The critical difference in these pairs is that the phonetic duration of the vowel in the second word is greater than that of the first.

The most salient and robust cue to the geminate vs singleton contrast is the duration of consonants in both Italian and Japanese [2, 9]. The durational ratio between short and long consonants in disyllabic Italian words is reported to be 1:2 or lower [2] and the ratio is 1:2 – 1:3 in Japanese [10].

Various studies report that vowels preceding geminate are longer in Japanese than vowels preceding singleton [9, 10], but the opposite is the case in Italian where the duration of the vowel is up to 37% shorter before geminate than before singleton consonants [11]. Thus, the complementarity of adjacent short-long sounds varies from one language to another, which means that this knowledge is not innate but must be learned. It is likely that such stark cross-linguistic phonetic differences have a measurable influence on the perception of these length contrasts.

Length contrasts, in general, may be difficult to acquire, because they are not as frequent cross-linguistically or as robust as other phonetic contrasts such as the voicing contrast (e.g. *tip/dip*) which is supported by as many as 16 co-varying acoustic cues [12]. Length contrasts appear to have many fewer such cues [9] and they are substantially influenced by speaking rate [13, 14].

Length contrast perception is a topic that has been attracting much interest among scholars in second language (L2) perception and learning [15, 16, 17]. In teaching L2 Japanese, length contrasts are known to be difficult for non-native learners from diverse L1 backgrounds [10, 18] including Australian English [19, 20, 21]. However, accurate processing of length contrasts is related to the knowledge of grammar (e.g. /kat/ *past tense of to buy* if the pitch pattern is Low-High or *to win* if it is High-Low) and basic vocabulary (e.g. /gako/ *school*, /kite/ *stamp*) and is an essential part of Japanese language acquisition. In particular, geminate /t/ frequently occurs in Japanese, as it is used to mark the past tense as exemplified above.

This study presents some pilot data that examined the identification of Japanese singleton and geminate consonants by non-native listeners differing in their previous linguistic experience. Two groups of non-native learners of Japanese (NNJ1 and NNJ2, three each) differed in their proficiency of Japanese. Native speakers of Italian (IT) and Australian English (OZ) with no prior knowledge of Japanese were also tested using the same stimuli (with response categories adjusted) to examine if and to what extent their L1 experience might transfer to the perceptual processing of an unknown

language. The question of interest was whether IT listeners' L1 knowledge of consonant length might give them an advantage over NNJ learners and OZ speakers in perceiving singleton and geminate consonants in Japanese. The knowledge gained from this research will be beneficial in advancing our current understanding of L1 transfer effects in cross-language speech perception. It could also form a basis for future developments in improving L2 Japanese pedagogy.

2. Methods

2.1. Stimuli preparation

2.1.1. Speakers

Seven (4 males (M1, M2, M3, M4), 3 females (F1, F2, F3)) native speakers of Japanese aged 20-40 each participated in a sixty-minute recording session in a recording studio at the National Institute of Japanese Language and Linguistics (NINJAL), Tokyo, Japan or at Macquarie University, Sydney, Australia and received \$20 (or equivalent in Japanese yen) for their participation. All spoke standard Japanese having been born or having spent most of their life in the Kanto region.

2.1.2. Speech materials

Seventy-nine real words and 40 non-words were recorded by M1 and 196 real words (including the 79 words recorded by M1) were recorded by the other six speakers (M2, M3, M4, F1, F2, F3). Of these materials, a total of 252 items were selected and presented to the participants. They consisted of six pairs (12 items) of non-words (/kak(t)a/, /kek(t)e/, /kok(t)o/, /tat(t)a/, /tet(t)e/, /tot(t)o/) spoken by M1 and two replications of randomly selected 60 pairs (2 x 120 items) of real words spoken by the remaining six (M2, M3, M4, F1, F2, F3) speakers. Items produced by each of the six speakers were distributed as evenly as practicable. Table 1 shows some of the test words used.

Table 1. Example of test words used.

	Type 1 X X	Type 2 X っ X	Type 3 X X X	Type 4 X っ X X
# mora	2	3	3	4
# syll	2	2	2/3	2/3
/t/	歌 Uta song	売った Utta sold	離党 Ritoo leave a (political) party	立冬 Ritto the beginning of winter
/k/	部下 Buka subordinate	物価 Bukka price	世間 Seken world	石鹸 Sekken soap

Words and non-words belonging to Types 1 and 2 and Types 3 and 4 contained the same number of syllables but differed in the presence or absence of geminate consonants (marked by っ in Table 1) with an extra mora (unit of length marked by X in Table 1) for items belonging to Types 2 and 4. (Note that both X and っ count as one mora). These materials were presented on the computer screen in random order and produced once in isolation and once in a short carrier sentence (/sokowa _____ to jomimasu/ "You read it as _____ there"). The pace of presentation was controlled by the experimenter

(the first author). The real words were intended to include all or most of the vowels in combination with /t k s/ and included one pair with medial /ʃ/ (/toʃin/ city centre - /toʃin/ a rush) and two pairs with medial /tʃ/ (/ittʃi/ one - /ittʃi/ agreement, /ʃittʃi / seven - /ʃittʃi / a marsh). The speech materials were digitally recorded at 44.1 kHz using Edirol R-09 HR and the target words were segmented and stored in separate files. Tokens produced in isolation were used as stimuli in this study.

2.2. Listeners

Four groups of listeners participated. Two groups (NNJ1, NNJ2) consisted of non-native learners of Japanese, three each at highly-advanced and upper intermediate levels of proficiency. These NNJ learners were all undergraduate students at Macquarie University. The three NNJ1 listeners had completed the most advanced language unit (approximately 400-600 hours of instruction) at tertiary level (equivalent to N1/N2 level according to the Japanese Language Proficiency Test (JLPT)). The other three NNJ2 listeners had just enrolled in the second most advanced language unit (approximately 300-450 hours of instruction, JLPT N2 level). Thus, the NNJ learners formed two distinct groups in terms of their Japanese language learning experience.

Some of the NNJ learners were L2 English speakers and spoke languages other than English and Japanese or had family members (mostly parents) who did. Their other languages included Cantonese, Mandarin, Maltese, Tagalog or Visayan. This type of multilingual background of NNJ learners is typical of the Japanese language classroom in a multi-cultural city such as Sydney. Although it would have been desirable to control for NNJ learners' L1 more tightly, it was not possible, as they participated on the basis of their availability, and no screening was carried out. This should be better controlled in future studies.

The third and fourth groups consisted of ten native speakers of Italian (IT) and eight native speakers of Australian English (OZ) with no prior knowledge of Japanese. The IT and OZ listeners were recruited from the student/staff population at Macquarie University or from the community. The IT listeners' mean length of residence in Sydney was less than 2 years. They were all tested in a quiet room at Macquarie University and received \$20 (or equivalent in gift voucher) for their participation.

2.3. Task

The listeners participated in a forced-choice identification task and listened to a total of 252 tokens arranged in 3 blocks of 84. They were tested individually in a session lasting between 30 and 40 minutes. The presentation of the stimuli and the collection of perception data were controlled by the UAB (University of Alabama at Birmingham) software [22]. The listeners could take a break after each block if they wished. They heard the stimuli at a self-selected, comfortable level over the high-quality headphones (Sennheiser 200PX-II) on a notebook computer and were asked to decide which of the four (for NNJ listeners) or two (for IT and OZ listeners) types the "word" belongs to.

While the two NNJ groups were given four response categories (「X X」 「X っ X」 「X X X」 「X っ X X」) explicitly indicating the presence or absence of geminate

consonants, the IT and OZ groups were given two response categories (“Singola (Single)”, “Doppia (Double)”) according to their L1 (Italian or Australian English), which means they only needed to decide whether the medial consonant was short or long regardless of the word length. The OZ listeners, for whom consonant length is not phonemic, were given *car tool* vs *cart tool* and *some others* vs *some mothers* as examples of “short” and “long” /t/ and /m/. The listeners were allowed, but not encouraged, to replay the stimulus tokens multiple times and were asked to guess if uncertain.

3. Results

3.1. Comparison between two groups of learners with more (NNJ1) and less (NNJ2) experience with Japanese

The overall results show that the NNJ1 listeners were more accurate than the NNJ2 listeners in identifying the words and non-words according to the presence or absence of the geminate consonants (95% vs 72%). Figure 1 shows the mean percentage of correct identification as a function of word type.

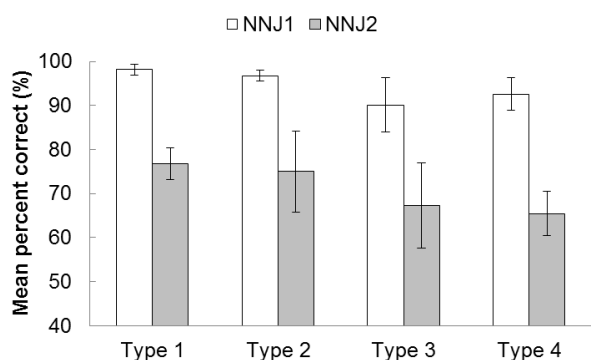


Figure 1: Mean percent correct identification (%) by NNJ1 and NNJ2 listeners as a function of word type.

Two-way ANOVA with Group (NNJ1, NNJ2) as a between-subjects factor and Word Type (T1, T2, T3, T4) as a within-subjects factor only showed a significant effect of Group [$F(1, 12) = 20.6, p = 0.01$]. The two-way interaction was not significant, indicating that the NNJ1 listeners were more accurate than the NNJ2 listeners in identifying all four types of words.

3.2. Comparison between four groups of listeners with (NNJ1, NNJ2) and without (IT, OZ) experience with Japanese

Figure 2 shows the mean percentage of correct identification for the four groups of listeners.

As NNJ groups and IT/OZ groups were given a different number of response categories (4 vs 2), the NNJ listeners' results were collapsed into with and without geminates to make them directly comparable to those of the IT and OZ listeners. As the group size differed widely ($n=3$ for NNJ1, NNJ2, $n=8$ for OZ, $n=10$ for IT), we only provide a descriptive analysis of the results for now.

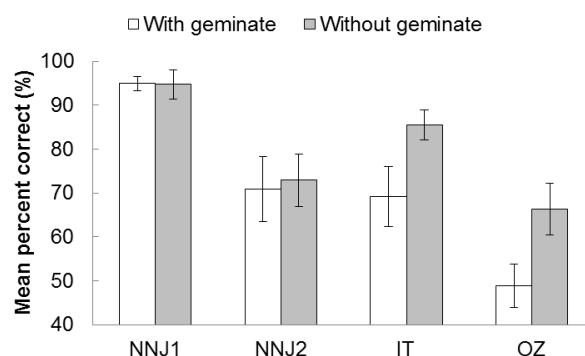


Figure 2: Mean percent correct identification (%) for words with and without geminate consonants by NNJ1, NNJ2, IT and OZ listeners.

The NNJ1 learners were more accurate in identifying Japanese singleton and geminate consonants than were the IT listeners who were, in turn, more accurate than the OZ listeners. These three groups differed from one another to a greater extent in their identification of tokens with (95% vs 69% vs 49%) than without geminates (95% vs 86% vs 66%). The NNJ2 learners were comparable to the IT listeners in identifying tokens with geminates (71% vs 69%) but were less accurate in identifying tokens without geminates (73% vs 86%). The NNJ2 learners and the OZ listeners differed less in their identification of tokens without geminates (73% vs 66%) than those with geminates (71% vs 49%).

As Figure 2 shows, the two NNJ groups identified tokens with and without geminate consonants with comparable accuracy (95% vs 95% for NNJ1, 71% vs 73% for NNJ2). The IT and OZ listeners, on the other hand, were more accurate in identifying tokens without geminates (86% for IT and 66% for OZ, respectively) than those with geminates (69% for IT and 49% for OZ, respectively).

The IT and OZ listeners' pattern of results suggests that their L1 experience with long sounds (consonants for IT and vowels for OZ) did not automatically transfer to the perceptual processing of Japanese geminates and that, by extension, phonetic realizations of geminates in Italian and Japanese are not equivalent. Acoustic analyses of the stimuli are in progress to shed light on this.

4. Discussion

We examined the identification accuracy of Japanese singleton and geminate consonants in words and non-words by four groups of listeners differing in their experience with length contrasts. There were three main findings. Firstly and perhaps not surprisingly, the highly-advanced learners were clearly more accurate than upper-intermediate learners in identifying Japanese words and non-words contrasting in the presence or absence of geminate consonants. Neither groups of learners showed any bias in their responses to the two types of stimuli, i.e. singleton vs geminate. This can be taken as a positive (or neutralizing) effect of Japanese language learning, because none of the NNJ learners had L1s that used consonant length contrastively in the way Japanese does. Secondly, the highly-advanced, but not upper-intermediate, learners were more accurate in identifying Japanese geminate consonants than were the naïve IT and OZ listeners. Thirdly, the IT listeners were more accurate than the OZ listeners in their perception of

Japanese consonant length (difference of 20% for both singleton (86% vs 66%) and geminate (69% vs 49%), respectively), possibly a benefit of L1 experience with consonant length contrasts.

While the NNJ listeners were introduced to the concept of contrastive length through their Japanese language learning, the IT listeners have two distinct consonant length categories in their L1. Having Australian English as an L1 would not equip the OZ listeners with consonant length categories which would be necessary to efficiently perceive Japanese singleton and geminate consonants. However, Australian English does use length contrastively for some vowel pairs, e.g. /kɛt/ *cut* vs /kɛt/ *cart*, /ʃɛd/ *shed* vs /ʃɛd/ *shared* [6, 7, 8] as mentioned in the Introduction. The OZ listeners in this study did not appear to transfer their L1 experience with vowel length to the processing of consonant length in Japanese. Their ability to identify the length category was poorer by 17% (49% vs 66%) when the target consonant was a geminate than when it was a singleton. Whether L1 experience with vowel length transfers positively to consonant length perception or not needs to be verified with listeners whose L1 does not use any length contrast for either consonants or vowels (e.g. American English, Mandarin or Spanish) on the one hand and then on the other hand with listeners whose L1 uses contrastive length across the entire vowel system (e.g. Thai), and not just partially (e.g. Australia English).

Some important questions which need further investigation relate to asymmetry in the ease of perception or learning. For example, is it easier to perceive and learn vowel or consonant length? For vowel length perception, are short vowels, which are much more frequent in Japanese [23], or long vowels, which are expected to be perceptually more salient and sonorant, easier to perceive? It has been reported that “phonetic extra-long voiceless obstruents are frequent” in Japanese [24]. As noted above, geminate /t/ may be particularly prevalent and noticeable in Japanese as a highly productive past tense marker (*ita stayed, was/were* vs *itta said or went, shita did* vs *shitta knew/found out, kita came or put on/wore* (depending on the pitch pattern) vs *kitta cut*, etc). Another question then arises as to if and how frequency of occurrence affects ease of perceptual learning.

Our future work includes 1) testing a larger number of Japanese learners at different levels of proficiency to better understand when and how non-native learners acquire the skills to accurately perceive length contrasts in Japanese and 2) testing native Italian speakers’ perception of vowel length contrasts in Japanese [25] and Thai, another language that contrasts short and long vowels, to determine if their L1 experience with consonant length might vary in the level of transfer to the perception of vowel length in different unknown languages. The findings will be useful in understanding the role of L1 experience in cross-language speech perception.

5. Conclusions

We presented preliminary data which showed that native speakers of Italian and Australian English were less accurate than the highly-advanced, but not upper-intermediate, learners in identifying singleton and geminate consonants in Japanese, a language unfamiliar to them. This suggests that L1 experience with length, whether consonant or vowel, may not necessarily guarantee accurate perception of consonant length in an

unknown language. This tentative conclusion is consistent with the results of a recent study on the discrimination of short vs long Japanese vowels [25], which found that familiarity with length contrasts in L1 may not guarantee efficient processing of length contrasts in unfamiliar languages. Furthermore, the extent of success may be related to how the contrasts in question are phonetically implemented in listeners’ L1. Italian and Japanese are known to differ from each other in this respect as briefly mentioned in the Introduction and this may have an impact on the perception of Japanese consonant length by native Italian listeners. In addition, our results for the two groups of learners at different levels of proficiency demonstrate that non-native learners need much time and practice before they differentiate themselves from naïve listeners, reconfirming that Japanese consonant length contrasts are difficult to acquire.

6. Acknowledgements

This research was supported by Macquarie University internal research grants. We thank Prof. Haruo Kubozono and Dr. Shinji Ogawa for their help with the recording sessions at NINJAL and our speakers and listeners who willingly participated in this research. We are grateful to the three reviewers for their valuable input on an earlier version of this paper.

7. References

- [1] Ladefoged, P., *A Course in Phonetics* [3rd Ed]. Fort Worth: Harcourt Brace College Publishers, 1993.
- [2] Esposito, A. and Di Benedetto, M. G., “Acoustical and perceptual study of gemination in Italian stops”, *Journal of the Acoustical Society of America*, 106: 2051-2062, 1999.
- [3] Rogers, D., and d’Arcangeli, L., “Italian”, *Journal of the International Phonetic Association*, 34, 117-121, 2004.
- [4] Payne, E. M., “Phonetic variation in Italian consonant gemination”, *Journal of the International Phonetic Association*, 35: 153-181, 2005.
- [5] Pickett, E. R., Blumstein, S. E., and Burton, M. W., “Effects of speaking rate on the singleton/geminate consonant contrast in Italian”, *Phonetica*, 56: 136-157, 1999.
- [6] Cox, F., “Australian English pronunciation into the 21st century”, *Prospect*, 21: 3-21, 2006a.
- [7] Cox, F., “The acoustic characteristics of /hVd/ vowels in the speech of some Australian teenagers”, *Australian Journal of Linguistics*, 26: 147-179: 2006b.
- [8] Cox, F. and Palethorpe, S., “Illustrations of the IPA: Australian English”, *Journal of the International Phonetic Association*, 37: 341-350, 2007.
- [9] Idemaru, K., and Guion, S., “Acoustic covariance of length contrast in Japanese stops”, *Journal of the International Phonetic Association*, 38: 167-186, 2008.
- [10] Han, M. S., “The timing control of geminate and single stop consonants in Japanese: A challenge for non-native speakers”, *Phonetica*, 49: 102-127, 1992.
- [11] Hajek, J., Stevens, M. and Webster, G., “Vowel duration, compression and lengthening in stressed syllables in Italian”, *The Proceedings of the 16th ICPHS*, 1057-1060, 2007.
- [12] Lisker, L., “‘Voicing’ in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees”, *Language and Speech*, 29: 3-11, 1986.

- [13] Hirata, Y., "Effects of speaking rate on the vowel length distinction in Japanese", *Journal of Phonetics*, 32: 565-589, 2004.
- [14] Hirata, Y. and Tsukada, K., "Effects of speaking rate and vowel length on formant frequency displacement in Japanese", *Phonetica*, 66: 129-149, 2009.
- [15] Altmann, H., Berger, I. and Braun, B. "Asymmetries in the perception of non-native consonantal and vocalic length contrasts", *Second Language Research*, 28: 387-413, 2012.
- [16] Bohn, O.-S., "Cross-language speech perception in adults", in W. Strange [Ed], *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 279-304, Timonium, MD: York Press, 1995.
- [17] Cebrian, J., "Experience and the use of duration in the categorization of L2 vowels", *Journal of Phonetics*, 34: 372-387, 2006.
- [18] Sonu, M., Kato, H., Tajima, K., Akahane-Yamada, R. and Sagisaka, Y., "Non-native perception and learning of the phonemic length contrast in spoken Japanese: Training Korean listeners using words with geminate and singleton phonemes", *Journal of East Asian Linguistics*, 22: 373-398, 2013.
- [19] Kato, A. and Cox, F., "Duration feature learning for an intervocalic voiceless velar stop consonant", *The Proceedings of the 17th ICPhS*, 1038-1041, 2011.
- [20] Toda, T., "Issues regarding geminate consonants in Japanese language education", *Journal of the Phonetic Society of Japan*, 11: 35-46, 2007.
- [21] Tsurutani, C., "Acquisition of geminate consonants by English-speaking learners of Japanese", *Review of Asian and Pacific Studies*, 32: 25-41, 2007.
- [22] Smith, S., *User Manual for UAB Software*. University of Alabama at Birmingham (UAB), Department of Rehabilitation Sciences, 1997.
- [23] Aoyama K., *A Psycholinguistic Perspective on Finnish and Japanese Prosody: Perception, Production and Child Acquisition of Consonantal Quantity Distinctions*. Boston: Kluwer Academic Publishers, 2001.
- [24] Vance, T. J., *The Sounds of Japanese*. New York: Cambridge University Press, 2008.
- [25] Tsukada, K., Hirata, Y., and Roengpitya, R., "Cross-language perception of Japanese vowel length contrasts: Comparison of listeners from different first language (L1) backgrounds", *Journal of Speech, Language, and Hearing Research*, 57: 805-814, 2014.