Perception of Italian and Japanese Consonant Length by Native Speakers of Australian English and Italian: A Pilot Study

Kimiko Tsukada[1], *Felicity Cox*[2], *John Hajek*[3], *Yukari Hirata*[4]

[1, 2]University Macquarie University, [3]The University of Melbourne, [4]Colgate University

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

Abstract

We examined the perception of Italian (IT) and Japanese (JP) consonant length contrasts (singleton vs geminate) in two groups of listeners: native speakers of IT and Australian English (OZ). Our preliminary results suggest that the IT listeners' experience with singleton/geminate contrasts was more beneficial than the OZ listeners' experience with vowel length contrasts in processing JP singleton and geminate consonants. Contrary to the previous literature, the OZ listeners identified stop length contrasts less accurately than fricative and affricate contrasts in both IT and JP. The IT listeners showed a manner effect only for JP with affricate length contrast being misperceived most.

Index Terms: cross-language speech perception, Italian (IT), Japanese (JP), singleton/geminate

1. Introduction

While all languages use durational variation (i.e. stretch or compress speech sounds) for various purposes, some languages use this effect contrastively, i.e. to differentiate the meaning of words. In this study, we focused on how Italian (IT) and Japanese (JP) consonants differing in length (i.e. singleton vs geminate) are perceived by listeners from different linguistic backgrounds.

Languages that employ consonant length contrasts vary widely with regard to the permissible location of the singleton/geminate contrasts in a word (initial, medial, final) [1, 2, 3] and the classes of sounds (e.g. stop, fricative, liquid, glide) [4, 5] that employ contrastive length. In general, singleton/geminate contrasts are more common in word-medial position and for obstruent sounds such as stops and fricatives [6].

In both IT [7, 8] and JP [9], consonant length is contrastive (e.g. /se**t**e/ *thirst* vs /se**tː**e/ *seven* for IT, /i**t**a/ *stayed, was/were* vs /i**tː**a/ *said, went* for JP). Singleton/geminate contrasts are an integral part of these languages where short and long consonants can precede and follow all vowel types. Australian English (OZ), another language under investigation in this study, does not use consonant length contrastively at all but does use length to contrast at least two pairs of vowels, e.g. /k**ɐ**t/ *cut* vs /k**ɐː**t/ *cart*, /ʃ**e**d/ *shed* vs /ʃ**eː**d/ *shared* [10, 11, 12, 13]. 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.

Given a wide range of typologically unrelated languages that use consonant length contrastively, a question arises as to whether length contrasts are equally perceptible across languages. Further, are all singleton/geminate contrasts perceived with equal accuracy regardless of phonetic features such as place or manner of articulation? The existing literature (e.g. [4, 5]) suggests a perceptibility hierarchy such that consonant length contrasts are least perceptible when the consonant in question is most similar in sonority to neighbouring vowels, e.g. glides and liquids [6]. The corollary is that length contrasts would be most perceptible for stops, which are the least sonorous sounds, by creating a salient contrast between the stop and the neighbouring vowels. There is in fact a relationship between sonority and the frequency with which singleton/geminate contrasts appear in world's languages with glide geminates reported to be less favoured than stop geminates [4, 5].

The ease with which singleton/geminate contrasts are perceived in any language by the same listeners is also likely to depend on their previous linguistic experience. Because IT, like JP, uses consonant length contrastively but OZ does not, it might be expected that their first language (L1) experience would give native IT listeners an advantage over OZ listeners in perceiving consonant length contrasts in JP as well as in IT. Alternatively, OZ listeners' experience with vowel length briefly mentioned above might transfer positively to unfamiliar languages and aid them in the length perception required for IT and JP *consonant* contrasts.

In this study, we examined the perception of IT and JP singleton/geminate contrasts in stops, fricatives and affricates. We sought to determine 1) if the IT and OZ listeners differ from each other in their consonant length perception of IT and JP and 2) if the IT and OZ listeners are equally influenced by the various manners of articulation represented in the singleton/geminate contrasts.

2. Methods

2.1. Stimuli preparation

2.1.1. Speakers

Three (2 males, 1 female) native speakers of IT and seven (4 males, 3 females) native speakers of JP in their 20-60s participated in the recording sessions lasting between 45 and 60 minutes. One of the authors with expertise in IT phonetics/phonology confirmed that the three IT speakers clearly differentiated the singleton and geminate consonants by length. All three IT speakers were highly experienced university teachers of IT with a clear understanding of pronunciation norms. According to self-report, all native JP speakers spoke standard JP having been born or having spent most of their life in the Kanto region. The speakers were recorded in a recording studio at Macquarie University, Sydney, Australia and at the National Institute of Japanese Language and Linguistics (NINJAL), Tokyo, Japan. They received $20 (or equivalent in Japanese yen) for their participation.

2.1.2. Speech materials

A total of 84 IT and 252 JP items were presented to the participants. Although minimal pairs contrasting in singleton/geminate consonants appear to be more prevalent in JP than in IT, the number of the IT items should be increased in future research to ensure a balance of items between the two stimulus languages. The IT items included (C)VC(C)V words and non-words (e.g. /ta**t(ː)**a/, /ku**k(ː)**u/, /si**s(ː)**i/) where the medial C was /p t k b d g f v s dʒ n/ and V was /i e a o u/. The JP items included six pairs of non-words (/ka**k(ː)**a/, /ke**k(ː)**e/, /ko**k(ː)**o/, /ta**t(ː)**a/, /te**t(ː)**e/, /to**t(ː)**o/) spoken by one of the seven speakers and two replications of 60 pairs of real words spoken by the remaining six (3 males, 3 females) speakers. The real words were (C)VC(C)V(V/n) where the medial C was /p t k s tʃ/ and V was /i e a o u/. Voiced geminates are limited in JP. Tables 1 and 2 show some of the test words used. Only stops, fricatives and affricates that were presented in both IT and JP stimuli are considered for analysis in the present study with stops making up the majority of tokens. These materials were presented on a computer screen in random order and produced once in isolation and once in a short carrier sentence (/diko ___ di nuovo/ "I say __ again" for IT, /sokowa _____ to jomimasu/ "You read it as ______ there" for JP). The pace of presentation was controlled by the experimenter (the first author). The recorded speech materials were digitized at 44.1 kHz and the target words were subsequently segmented and stored in separate files. Tokens produced in isolation were used as stimuli in this study.

Table 1. *Examples of IT test words used.*

Manner	Singleton	N	Geminate	N
stop	Sete *thirst* Eco *echo*	26	Sette *seven* Ecco *here*	26
fricative	Rosa *a rose* Beve *he/she drinks*	9	Rossa *red* Bevve *he/she drank*	9
affricate	Agio *ease*	1	Aggio *premium*	1

Table 2. *Examples of JP test words used.*

Manner	Singleton	N	Geminate	N
stop	Kotoo *an isolated island* Sokoo *behaviour*	108	Kottoo *an antique* Sokkoo *a swift attack*	108
fricative	Kasee *Mars* Toshin *city centre*	14	Kassee *active* Tosshin *a rush*	14
affricate	Ichi *one* Shichi *seven*	4	Icchi *agreement* Shicchi *a marsh*	4

2.2. Listeners

Two groups of listeners participated. The first group consisted of ten (7 males, 3 females) native IT speakers (*mean* = 29.3 years) with no prior knowledge of JP and the second group consisted of eight (8 females) native OZ speakers (*mean* = 28.0 years) with no prior knowledge of IT or JP. The 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 336 tokens arranged in 4 blocks of 84. The first 3 blocks contained JP tokens and the last block contained IT tokens. The first three items in each block were for practice and were not analyzed. No feedback was given for practice or test items. The listeners were tested individually or in a small group of two to three in a session lasting between 30 and 40 minutes. The experimental session was self-paced and the listeners could take a break after each block if they wished. They listened to the stimuli on a notebook computer at a self-selected, comfortable level over the high-quality headphones (Sennheiser 200PX-II).

The listeners were given two response categories ("Singola (Single)", "Doppia (Double)") according to their L1 (IT or OZ). The OZ listeners, for whom consonant length is not contrastive, were given *car **t**ool* vs *car**t t**ool* and *so**me** others* vs *so**me m**others* as examples of "short" and "long" /t/ and /m/. Clarification of the instructions was provided if necessary. The listeners only needed to decide whether the medial consonant was short or long regardless of the word length. They were allowed, but not encouraged, to replay the stimulus tokens multiple times and were asked to guess if uncertain.

3. Results

3.1. Overall

The overall results show that, averaged across the IT and JP tokens, the IT listeners correctly categorized the word-medial singleton/geminate tokens 83% of the time whereas the OZ listeners did so 64% of the time.

Three-way ANOVA with Group (IT, OZ) as a between-subjects factor and Stimulus language (IT, JP) and Manner of articulation (stop, fricative, affricate) as within-subjects factors showed significant effects of Group [$F(1, 16) = 16.8$, $p < 0.001$], Language [$F(1, 16) = 22.9$, $p < 0.001$] and Manner [$F(2, 32) = 3.7$, $p < 0.05$]. The two-way interactions involving the Group factor were significant [Group x Language: $F(1, 16) = 5.3$, $p < 0.05$, Group x Manner: $F(2, 32) = 4.1$, $p < 0.05$]. These interactions are plotted in Figures 2 and 3, respectively. The three-way interaction and the Language x Manner interaction were not significant.

Figure 1 shows the mean percentage of correct identification of all tokens by the IT and OZ listeners as a function of stimulus language. The IT listeners were more accurate than the OZ listeners in identifying the word-medial singleton or geminate in both IT and JP (94% vs 67% in IT and 74% vs 60% in JP) with the between-group difference twice as large for the IT (27%) as the JP (14%) consonants. This suggests that L1 experience with singleton/geminate contrasts aided the IT listeners in identifying the length category in JP and that the OZ listeners' experience with vowel length did not boost their perception to the level of the IT listeners. While the IT listeners were clearly more accurate in perceiving the length category in familiar IT than in unfamiliar JP, the simple effect of Stimulus Language was not significant for the OZ listeners who were inexperienced in both IT and JP.

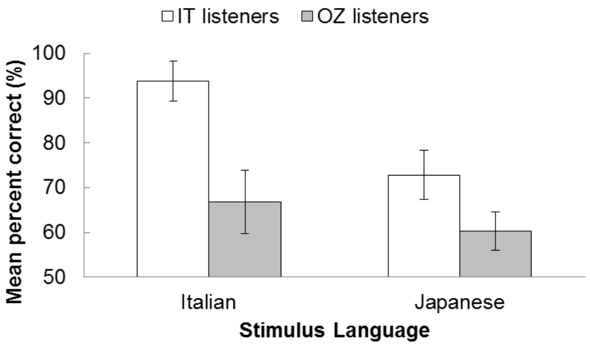

Figure 1: *Mean percent correct identification (%) of singleton and geminate consonants in IT and JP by IT and OZ listeners. The error bars indicate ± one standard error of the mean.*

Figure 2 shows the mean percentage of correct identification by the IT and OZ listeners as a function of manner of articulation averaged across stimuli in IT and JP. The IT listeners were less affected by the manner (stop: 85%, fricative: 85%, affricate: 83%) than the OZ listeners who showed lower length categorization accuracy for stop (57%) than for fricative (70%) and affricate (64%). Figure 2 clearly illustrates that the IT listeners were more accurate than the OZ listeners in identifying the length category for all three classes of consonants. While it would be desirable to examine the effect of preceding and following vowels, the number of items presented was not balanced to permit this type of analysis.

We now turn to how the IT and OZ listeners perceived singleton and geminate consonants in IT and JP separately.

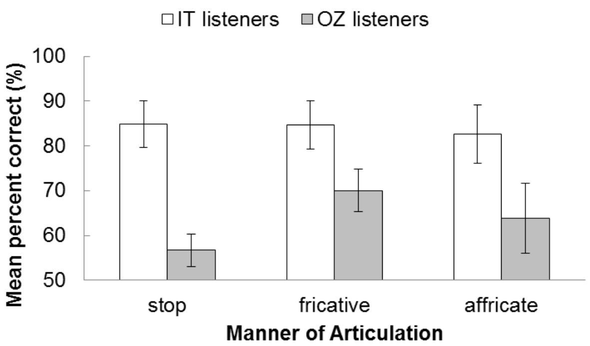

Figure 2: *Mean percent correct identification (%) of singleton and geminate consonants by IT and OZ listeners of all tokens as a function of manner of articulation. The error bars indicate ± one standard error of the mean.*

3.2. The perception of IT singleton/geminate

Figure 3 shows the mean percent correct identification of IT singleton/geminate consonants by the IT and OZ listeners as a function of manner of articulation. As expected, the IT listeners were more accurate than the OZ listeners in identifying the length category of their L1 consonants.

Two-way ANOVA with Group (IT, OZ) as a between-subjects factor and Manner (stop, fricative, affricate) as a within-subjects factor only showed a significant main effect of Group [$F(1, 16) = 24.6$, $p < 0.001$]. The Manner and the two-way interaction were not significant, indicating that the IT listeners were more accurate than the OZ listeners in their length identification for all three classes of sounds (92% vs 57% for stop, 94% vs 75% for fricative, 95% vs 69% for affricate). For the OZ group, the simple main effect of Manner approached significance [$F(2, 32) = 3.2$, $p = 0.059$]. Figure 3 shows that the OZ listeners' length categorization accuracy was less accurate when the consonant was stop (57%) compared to when it was fricative (75%) or affricate (69%). This is puzzling given that singleton/geminate contrasts are expected to be most easily perceptible for stops. It is necessary to include more fricative and affricate tokens to verify if the OZ listeners have additional perceptual difficulties when the target consonant is stop.

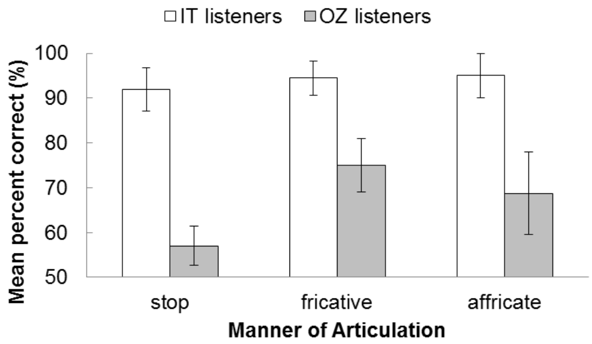

Figure 3: *Mean percent correct identification (%) of IT singleton and geminate consonants by IT and OZ listeners as a function of manner of articulation. The error bars indicate ± one standard error of the mean.*

3.3. The perception of JP singleton/geminate

Figure 4 shows the mean percent correct identification of the JP singleton/geminate consonants by the IT and OZ listeners as a function of manner of articulation. Again, the IT listeners were consistently more accurate than the OZ in identifying the length category of JP consonants.

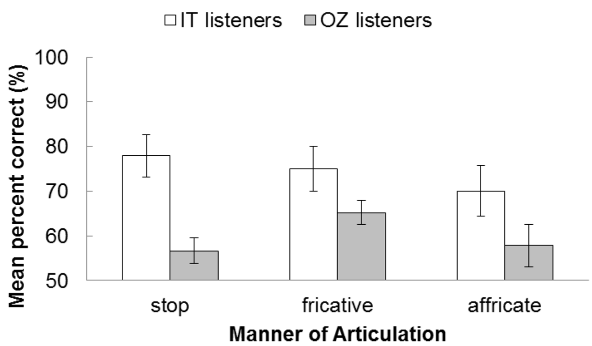

Figure 4: *Mean percent correct identification (%) of JP singleton and geminate consonants by IT and OZ listeners as a function of manner of articulation. The error bars indicate ± one standard error of the mean.*

Two-way ANOVA with Group (IT, OZ) as a between-subjects factor and Manner (stop, fricative, affricate) as a within-subjects factor showed a significant effect of Group [$F(1, 16) = 5.7$, $p < 0.05$] and the two-way interaction [$F(2, 32) = 3.4$, $p < 0.05$], but the main effect of Manner narrowly missed reaching significance [$F(2, 32) = 3.3$, $p = 0.051$]. The simple effect of Group was significant only for stop [$F(1, 22) = 10.4$, $p < 0.01$]. The simple effect of Manner reached significance for the OZ group [$F(2, 32) = 3.5$, $p < 0.05$] and approached significance for the IT group [$F(2, 32) = 3.1$, $p = 0.059$]. The OZ listeners were less accurate in identifying the length category when the consonant was stop (57%) than when it was fricative (65%).

4. Discussion

We examined the identification accuracy of IT and JP singleton and geminate consonants in words and non-words by two groups of listeners differing in their experience with length contrasts.

There were three main findings. Firstly and perhaps not surprisingly, the IT listeners were clearly more accurate than the OZ listeners in identifying the length category in both familiar IT and in unfamiliar JP. The finding that the IT listeners outperformed the OZ listeners in their perception of the JP stimuli suggests that there was a positive transfer of L1 experience with singleton/geminate contrasts in their cross-language perception. However, the IT listeners' perception of consonant length was much less efficient in JP than in IT and, as a result, their advantage over the OZ listeners was much more limited in JP.

Secondly, while the IT listeners were much more accurate in identifying the consonant length in IT than in JP, the same was not the case with the OZ listeners who would be regarded as unbiased judges of consonant length. It appears that the OZ listeners' experience with vowel length contrasts in their L1 was not particularly useful in processing singleton/geminate contrasts in IT and JP.

Thirdly, contrary to the expectations based on the existing literature, the effect of manner of articulation was fairly limited. Notably, however, the OZ listeners' length perception was least accurate for stops refuting the hypothesis that consonant length would be most perceptible for stops. The reason for this specific effect remains unclear and requires further investigation.

This study was limited in the range of consonants presented to the listeners. For example, IT but not JP has singleton/geminate contrasts for liquids. It would be interesting to examine how the gap in the sound inventory might impact on listeners' cross-language perception of length contrasts. Thus, our future work includes 1) testing the perception of native JP listeners as they identify familiar (e.g. nasals) and unfamiliar (e.g. liquids) singleton/geminate contrasts in IT as well as JP and 2) testing listeners whose L1 only uses vowel but not consonant length (e.g. Thai) or no length contrast of any kind (e.g. Spanish). The findings will be useful in advancing our current understanding of the role of L1 experience in cross-language speech perception.

5. Conclusions

We presented preliminary data which showed that the IT listeners were more accurate than the OZ listeners in identifying singleton and geminate consonants in JP as well as in IT. However, their processing advantage was much more limited when faced with JP stimuli than IT stimuli. On the basis of our current data, the OZ listeners' experience with vowel length contrasts in their L1 did not appear to transfer to successful perception of consonant length in unknown languages. We, therefore, tentatively conclude that L1 experience with length may be useful under certain conditions but may not necessarily guarantee accurate perception of length in an unknown language.

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] Abramson, A., "The perception of word-initial consonant length: Pattani Malay", Journal of the International Phonetic Association, 16: 8-16, 1986.

[2] Ridouane, R., "Gemination in Tashlhiyt Berber: An acoustic and articulatory study", Journal of the International Phonetic Association, 37: 119-142, 2007.

[3] Kraehenmann, A. and Lahiri, A., "Duration differences in the articulation and acoustics of Swiss German word-initial geminate and singleton stops", Journal of the Acoustical Society of America, 123: 4446-4455, 2008.

[4] Aoyama, K. and Reid, L. A., "Cross-linguistic tendencies and durational contrasts in geminate consonants: An examination of Guinaang Bontok geminates", Journal of the International Phonetic Association, 36: 145-157, 2006.

[5] Maddieson, I. "Glides and gemination", Lingua, 118: 1926-1936, 2008.

[6] Hansen, B. B. and Myers, S. "The consonant length contrast in Persian: Production and perception", http://uts.cc.utexas.edu/~smyers/Hansen&Myers.pdf, accessed on 2 Jun, 2014.

[7] 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.

[8] Rogers, D. and d'Arcangeli, L., "Italian", Journal of the International Phonetic Association, 34, 117-121, 2004.

[9] Vance, T. J., The sounds of Japanese. New York: Cambridge University Press, 2008.

[10] Bernard, J.R.L., "Length and identification of Australian English vowels", AUMLA: Journal of the Australasian Universities Language and Literature Association, 27: 37-58, 1967.

[11] Cox, F., "Australian English pronunciation into the 21st century", Prospect, 21: 3-21, 2006a.

[12] Cox, F., "The acoustic characteristics of /hVd/ vowels in the speech of some Australian teenagers", Australian Journal of Linguistics, 26: 147-179: 2006b.

[13] Cox, F. and Palethorpe, S., "Illustrations of the IPA: Australian English", Journal of the International Phonetic Association, 37: 341-350, 2007.