PERCEPTION OF ITALIAN AND JAPANESE SINGLETON/GEMINATE CONSONANTS BY LISTENERS FROM DIFFERENT LANGUAGE BACKGROUNDS

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

[1, 2]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 investigated if and how the use of one or multiple languages (bilingualism hereafter) affects the perception of intervocalic singleton/geminate consonants in Italian and Japanese. Two groups of non-native listeners (monolingual speakers of Australian English and bilingual speakers of Cantonese/English or Vietnamese/English) were examined. Two groups of native listeners (Italian and Japanese) residing in Australia acted as controls. Our aim was to test the hypothesis that the bilinguals process unfamiliar sounds more efficiently than the monolinguals due to their expanded phonetic inventories. Results showed that bilingualism did not result in superior performance overall. However, while the monolinguals identified consonant length in Italian slightly more accurately (albeit non-significantly) than in Japanese, the bilinguals showed the opposite pattern, i.e. greater accuracy with Japanese than with Italian. Generally, bilingual and monolingual non-native listeners misperceived geminates as singletons more often than they misperceived singletons as geminates in Japanese, but not in Italian.

Keywords: singleton/geminate, cross-language perception, Italian, Japanese, bilingualism.

1. INTRODUCTION

Would it help the processing of new information if one spoke more than one language? This study compared Australian English monolinguals and Cantonese/English or Vietnamese/English bilinguals as they perceived words containing short (singleton) or long (geminate) consonants inter-vocalically in Italian (IT) and Japanese (JP). In these languages, incorrect use of length can impair communication (e.g. eco *echo* vs ecco *here* in Italian, kako *past* vs kakko *parentheses* in Japanese). It is well-documented that non-native speakers have difficulty processing the Japanese special mora (unit of length) that forms the second part of geminates or long vowels [4, 5, 6].

Australian English (OZ), another language under investigation in this study, does not use consonant length contrastively within a word, but does use length to contrast at least two pairs of vowels, e.g. /kɐt/ *cut* vs /kɐːt/ *cart*, /ʃed/ *shed* vs /ʃeːd/ *shared* [2]. 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.

We sought to understand if a presumably expanded sound inventory resulting from proficiency in more than one language might be beneficial and help bilinguals to process unfamiliar sounds more efficiently than monolinguals. In this research, we use the term "bilingualism" to refer to the use of two or more languages. We decided to examine the perception of Cantonese/English and Vietnamese/English bilinguals (AS for Asian) for two main reasons. Firstly, their perception of IT and JP has been little studied before. Secondly, being largely monosyllabic tonal languages [1, 8, 9, 11], these languages do not exhibit intervocalic geminates and there is no phonemic contrast in consonant length. The OZ and AS listeners may still have enhanced sensitivity to durational variation compared to listeners of languages (e.g. Mandarin, Spanish) that do not use duration contrastively at all. This is because the former has experience with vowel length as mentioned above and the latter has access to multiple phonologies (and, like OZ listeners, shared experience of Australian English).

Our primary interest was to gain an insight into the processing of singleton/geminate consonants in IT and JP by listeners differing in their first languages (L1). We focused on the perception of naïve AS listeners with a view to determining which of the stimulus languages (IT or JP) had a more identifiable length category. The question of interest was whether or not bilingualism may give AS listeners an advantage over monolingual OZ listeners with less experience of other languages.

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. Two of the authors with expertise in IT and JP phonetics/phonology confirmed that all the IT and JP speakers clearly differentiated the singleton and geminate consonants by length. The speakers were recorded in a recording studio at a university in Sydney, Australia or at a research institute in 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 [10], 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 42 pairs of words and non-words contrasting in intervocalic consonant length (i.e. (C)V**C(C)**V where the medial C was /p t k b d g f v s dʒ n/ and V was /i e a o u/) spoken by the three IT speakers. The JP items included six pairs of non-words 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 words and non-words contrasted in intervocalic consonant length (i.e. (C)V**C(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 the number of items representing each manner and singleton/geminate type and some examples of the test words used. Only stops, fricatives and affricates that were presented in both IT and JP stimuli were considered for analysis in the present study with stops making up the majority of tokens.

Table 1: Examples of Italian test words used.

Manner	Singleton	N	Geminate	N
stop	sete thirst	26	sette *seven*	26
fricative	rosa a rose	9	rossa red	9
affricate	agio ease	1	aggio premium	1

Table 2: Examples of Japanese test words used.

Manner	Singleton	N	Geminate	N
stop	saka a slope	108	sakka an author	108
fricative	sosen ancestor	14	sossen to take the initiative	14
affricate	ichi one	4	icchi agreement	4

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. Only tokens produced in isolation were used as stimuli in this study.

2.2. Listeners

Four groups of listeners participated. The first two groups consisted of five (1 male, 4 females) native listeners of JP (mean = 23.2 years) with no prior knowledge of IT and ten (7 males, 3 females) native listeners of IT (mean = 29.3 years) with no prior knowledge of JP. These listeners were included as controls. The other two groups consisted of eight female native listeners of OZ (mean = 27.3 years) and ten (2 males, 8 females) bilingual AS listeners (mean = 24.6 years) with no prior knowledge of IT or JP. The AS group consisted of six (2 males, 4 females) Cantonese and four female Vietnamese bilinguals. Both early and late bilinguals were included in the AS group based on their availability. Another three female Cantonese/English bilinguals were tested and, although their results did not substantially differ from the remaining Cantonese-speaking listeners, they were excluded from analysis because they had formally studied JP in the past. The listeners were recruited from the student/staff population at the university or from the community. The IT and JP listeners' mean length of residence in Sydney was less than 2 years. They were all tested in a quiet room on campus and received $20 (or equivalent in gift vouchers) for their participation.

2.3. Task

The listeners participated in a forced-choice identification task and listened to a total of 336 isolated tokens arranged in four blocks of 84. The first three 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 computer at a self-selected, comfortable level over the high-quality headphones (Sennheiser 200PX-II).

The JP group was given four response categories (1.「Ｘ Ｘ」 2.「Ｘ っ Ｘ」 3.「Ｘ Ｘ Ｘ」 4.「Ｘ っ Ｘ Ｘ」) with the presence of geminates encoded by the symbol (「っ」). The other groups were given two response categories (1. "Singola (Single) X", 2. "Doppia (Double) XX") according to the language most familiar to them (IT or English), which means they only needed to decide whether the medial consonant was short or long regardless of the word length. It may be preferable to use the same response categories across all listener groups in our future work. The English-speaking listeners (OZ, AS), for whom lexical items do not contrast in consonant length, 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. 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

Except for one Cantonese bilingual, the remaining nine AS listeners were more accurate in identifying the length category in the JP than the IT stimuli. Five out of the eight OZ listeners showed the opposite pattern.

Figure 1: Mean % correct identification by four groups of listeners.

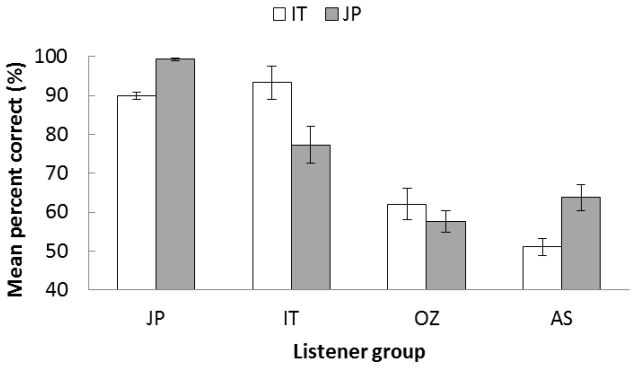


Figure 1 shows the mean percentage of correct identification by four groups of listeners. The JP and AS groups were more accurate in identifying the consonant length in JP than in IT whereas the IT and OZ showed the opposite pattern. It is not surprising that the IT and JP groups more accurately identified the singleton/geminate consonants in their L1s.

Two-way ANOVA with Group (JP, IT, OZ, AS) as a between-subjects factor and Stimulus Language (IT, JP) as a within-subjects factor showed a significant effect of Group [$F(3, 29) = 29.1$, $p < 0.001$] and the two-way interaction [$F(3, 29) = 18.8$, $p < 0.001$], but not the main effect of Stimulus Language. The simple effect of Group was significant for both IT [$F(3, 38) = 34.8$, $p < 0.001$] and JP [$F(3, 38) = 19.4$, $p < 0.001$] tokens. The OZ and AS groups, who did not differ from each other, were significantly less accurate than the IT and JP listeners in identifying both IT and JP tokens. The IT and JP groups did not differ from each other for the IT tokens, but the JP group was more accurate than the IT group in identifying the singleton/geminate consonants in the JP tokens.

The simple effect of Stimulus Language was significant for the JP [$F(1, 29) = 5.2$, $p < 0.05$], IT [$F(1, 29) = 29.8$, $p < 0.001$] and AS [$F(1, 29) = 19.6$, $p < 0.001$] groups, but not the OZ group. As seen in Figure 1, the JP and AS groups were more accurate in identifying the JP than the IT length category while the IT group unsurprisingly showed the opposite pattern. The next two sections provide an analysis of the direction of misperception.

3.2. IT stimuli

Figure 2: Mean % incorrect identification (%) for the IT stimuli by four groups of listeners.

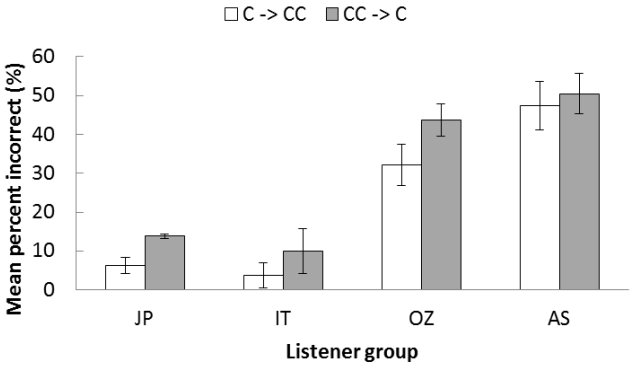


Figure 2 shows the mean percentage of singletons misperceived as geminates (C -> CC) and geminates misperceived as singletons (CC -> C) for the IT stimuli. All four groups of listeners tended to misperceive geminates as singletons more often than the other way around.

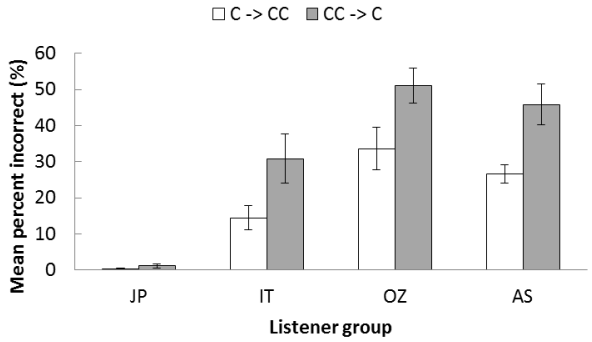
Two-way ANOVA with Group (JP, IT, OZ, AS) as a between-subjects factor and Direction of Misperception (C -> CC, CC -> C) as a within-subjects factor only showed a significant effect of Group [$F(3, 29) = 36.5$, $p < 0.001$]. Although the

listeners misperceived IT geminates as singletons more often than they misperceived singletons as geminates, this difference did not reach significance. As is clearly seen in Figure 2, the IT and JP groups made fewer incorrect responses than did the OZ and AS groups.

3.3. JP stimuli

Figure 3 shows the mean percentage of singletons misperceived as geminates (C -> CC) and geminates misperceived as singletons (CC -> C) for the JP stimuli. Three non-native groups (IT, OZ, AS) misperceived JP geminates as singletons more often than the other way around.

Figure 3: Mean % incorrect identification (%) for the JP stimuli by four groups of listeners.



Two-way ANOVA with Group (JP, IT, OZ, AS) as a between-subjects factor and Direction of Misperception (C -> CC, CC -> C) as a within-subjects factor showed significant main effects of Group [$F(3, 29) = 17.8$, $p < 0.001$] and Direction [$F(1, 29) = 15.5$, $p < 0.001$], but not the two-way interaction. This indicates that listeners generally misperceived JP geminates as singletons more often than the other way around. Averaged across the two directions of misperception, the JP group made significantly fewer incorrect responses than the IT group who, in turn, made fewer errors than the AS and OZ groups. The two groups of naïve listeners did not differ from each other in their direction of misperception. These results indicate that although the IT listeners were not as accurate as the JP listeners, their experience with a consonant length contrast was still beneficial in differentiating singleton from geminate consonants in JP.

4. GENERAL DISCUSSION

We have observed that two groups of listeners from monolingual (OZ) and bilingual (AS) backgrounds differed in the perception of consonant length contrasts in IT and JP, languages unknown to them. While there was no clear bilingual advantage and both groups were less accurate than the other two groups of listeners of quantity languages, IT and JP, bilingual listeners were more accurate in JP than in IT.

Another finding was that geminates were misperceived as singletons in JP but not in a statistically significant way in IT. In IT, vowels preceding geminates are known to be up to 37% *shorter* than vowels preceding singletons [3], but in JP, vowels preceding geminates are *longer* than vowels preceding singletons [4, 7]. When non-native listeners responded to the JP stimuli, they may have paid greater attention to vowels (high in sonority) than to consonants (low in sonority) and used vowel duration as a cue to the length category of the following consonant. The IT, AS and OZ listeners misperceived geminates as singletons, possibly because, unlike in IT, vowels preceding geminates were not shortened in JP.

Being of Asian background, the AS listeners may be generally more familiar with and exposed to the JP culture and language than IT even though they never studied either language. In other words, the AS listeners may find JP to be less dissimilar (or foreign) than IT to their own L1s. Conversely, although we might expect the OZ listeners to be more familiar with IT than JP due to linguistic-typological similarity, their performance did not favour either one or the other. Needless to say, this requires verification with a larger number of AS listeners. It would be useful to examine how JP sounds are processed by bilinguals from non-Asian backgrounds whose languages do not have an intervocalic consonant length contrast. It would also be necessary at some point to test listeners from a wider range of Asian language backgrounds such as Korean, Mandarin and Thai to confirm that our findings are not limited to Cantonese and Vietnamese listeners.

As we observed in this study, the IT and JP listeners were more accurate than naïve listeners (both monolingual and bilingual) in identifying singleton and geminate consonants in an unfamiliar consonant quantity language. It would be valuable to verify if this generalizes to native listeners of other consonant quantity languages such as Bengali or Punjabi. Answering these questions would lead to a better understanding of the role of previous linguistic experience such as L1 and bilingualism in spoken language processing across different languages.

5. ACKNOWLEDGEMENTS

This research was supported by Macquarie University internal research grants. We thank anonymous reviewers for their time.

6. REFERENCES

[1] Bauer, R. S., Benedict, P. K. 1997. *Modern Cantonese Phonology*. Berlin/New York: Mouton de Gruyter.

[2] Cox, F., Palethorpe, S. 2007. Illustrations of the IPA: Australian English. *J. Int. Phonetic Assoc.* 37, 341-350.

[3] Hajek, J., Stevens, M., Webster, G. 2007. Vowel duration, compression and lengthening in stressed syllables in Italian. *Proc. 16th ICPhS* Saarbrücken, 1057-1060.

[4] Han, M. S. 1992. The timing control of geminate and single stop consonants in Japanese: A challenge for non-native speakers. *Phonetica* 49, 102-127.

[5] Hardison, D. M., Motohashi Saigo, M. 2010. Development of perception of second language Japanese geminates: Role of duration, sonority, and segmentation strategy. *Applied Psycholinguistics* 31, 81-99.

[6] Hirata, Y. 2004. Effects of speaking rate on the vowel length distinction in Japanese. *J. Phonetics* 32, 565-589.

[7] Hirata, Y. 2007. Durational variability and invariance in Japanese stop quantity distinction: Roles of adjacent vowels. *J. Phonetic Society of Japan* 11, 9-22.

[8] Kirby, J. P. Vietnamese (Hanoi Vietnamese). 2011. *J. Int. Phonetic Assoc.* 41, 381-392.

[9] Pajak, B., Levy, R. 2014. The role of abstraction in non-native speech. *J. Phonetics* 46, 147-160.

[10] Vance, T. J. 2008. *The Sounds of Japanese.* Cambridge University Press.

[11] Zee, E. 1991. Cantonese. *J. Int. Phonetic Assoc.* 21, 46-48.