PERCEPTION OF TASHLHIYT CONSONANT QUANTITY CONTRASTS BY NATIVE VS. NONNATIVE LISTENERS FROM THREE LANGUAGES

Pierre Hallé[a,b], Philipp Buech[a], Yueh-chin Chang[c], Feng-fan Hsieh[c], Jiayin Gao[a], and Rachid Ridouane[a]

[a] LPP (CNRS/Sorbonne Nouvelle, Paris), [b] LMC2lab (Paris Cité), [c] Tsing Hua University (Hsinchu, Taiwan)
pierre.halle@sorbonne-nouvelle.fr, philipp.buech@sorbonne-nouvelle.fr, ycchang@mx.nthu.edu.tw,
ffhsieh@mx.nthu.edu.tw, jiayin.gao@cnrs.fr, rachid.ridouane@sorbonne-nouvelle.fr

ABSTRACT

Remarkably, Tashlhiyt has gemination contrasts for all consonants in all word positions. Perception of these contrasts by natives vs. nonnatives is little documented, with the exception of [Hallé, Ridouane, and Best, 2016], which found that natives performed near ceiling on all contrasts under scrutiny, whereas nonnative French listeners were much better on word-final than -initial contrasts. The authors tentatively proposed that the word-initial and word-final positions engage different perceptual processes. We re-examine these findings using a more comprehensive design including initial, medial, and final positions, using Tashlhiyt native listeners and nonnative listeners from three languages differing in terms of quantitative and qualitative incidence of gemination: Chinese, French, and Japanese.

Overall, easy perception of non-initial Tashlhiyt geminate-singleton contrasts by nonnative listeners is confirmed, supporting the earlier proposition that they engage universal perceptual processes. In contrast, word-initial contrasts would require language-specific learning.

Keywords: geminate consonants, within-word position, Tashlhiyt, cross-linguistic speech perception

1. INTRODUCTION

The perception of geminate vs. single consonants by listeners whose language has no geminates has been mainly studied in the framework of second language acquisition, most often using stimuli from a gemination language and nonnative learners vs. control native listeners. Japanese and Italian are the most popular gemination languages used in these studies. Worth to note, in these two languages, lexical geminates are restricted to the word-medial position. One common finding across studies is that no-gemination language listeners learning a gemination language have substantial difficulty, especially beginners, at perceiving word-medial geminates in identification or discrimination tasks (English learners: [1-6]; Korean learners: [7-9]; Chinese learners: [10-13]). [6] tested American listeners of four different learning levels of Japanese (including naïve listeners) on a same-different discrimination test and found error rates in the 40-30% range. Identification tasks seem to show better performances: in [8] (also see [9]), Korean learners of Japanese identify correctly Japanese singletons at the 64% level but geminates near ceiling level. Chinese listeners tend to show a reverse asymmetry [12]. More generally, perception of the Japanese geminate-singleton contrast differs according to the phonological and phonetic systems of the listener's native language.

A first cause of variation is the different extents to which languages allow gemination or not. Some languages, such as Lingala, have no gemination at all. Concatenated gemination occurs in Chinese dialects with checked syllables such as Cantonese (e.g., 拍佢, *pak koy* 'tap him'); it is restricted to /-n/#/n-/ in Mandarin. It is far more common in English or French, which allow many consonant codas. Korean "tense" stops are somewhat akin to geminates due to their long closure duration. Tashlhiyt is an extreme case of gemination incidence since it allows gemination in all utterance positions and for all consonants.

A second cause of variation is related to the possible secondary cues to gemination and how listeners exploit them. The primary cue –consonant duration– may be accompanied by prosodic cues such as flanking vowels' durations and/or pitch. [4] showed that consonant and preceding vowel durations interact in the perception of gemination in language-specific ways: a perceptual bias toward gemination is induced by preceding vowel *shortening* for Italian vs. preceding vowel *lengthening* for Japanese listeners, reflecting differences in how secondary cues to gemination are implemented in Italian versus Japanese.

In this study, we focused on the first cause of variation: the extent of gemination incidence in the listener's native language. Tashlhiyt fits this purpose quite well since its geminates are produced with primary durational cues, with little or no secondary cues to gemination, whether prosodic or phonetic-acoustic, as shown in [14-15]. That is, the differences that may be found across listeners will not likely be attributable to differences in the language-specific ways they exploit secondary cues to gemination in perception. [16] examined the discrimination of some Tashlhiyt geminate-singleton contrasts by French vs. Tashlhiyt listeners. The contrasts were

/b/-/bb/, /d/-/dd/, /g/-/gg/, /f/-/ff/, and /s/-/ss/ in word-initial position, and /t/-/tt/ in word-final position. Whereas Tashlhiyt listeners performed near ceiling level on all contrasts, French listeners did rather poorly on word-initial contrasts (in the 65-75% correct discrimination range) but did quite well on word-final /t/-/tt/ contrasts (~90% correct). From these rather scarce data, the authors nevertheless proposed that the good French performance on word-final /t/-/tt/ reflects a universal sensitivity to the "beat" given by successive salient phonetic events (here, between the vowel preceding /t, tt/ and the release of /t, tt/). In contrast, the perception of word-initial contrasts requires sensitivity to "sound substance duration," which would need language-specific learning. Anecdotal evidence supporting the authors' claim was drawn from a single word-medial contrast used in the training phase (*jutid-juttid*). To begin with, we believed that more data and a more complete design were needed before drawing reliable conclusions about the perceptual mechanisms at work when listening to geminate compared to single consonants. In particular, the design should cover all the three initial, medial, and final positions, different types of consonants, and use a more varied sample of languages with regards to gemination incidence. Moreover, since [16] was restricted to word-initial and -final positions, no comparison could be made with the data reported in the literature, which indeed massively deal with the medial position.

In the present study, we used a rather exhaustive design including the three within-word positions and four different target consonants /d, t, s, n/ covering major manners of articulation and voice settings. We recruited participant listeners from four different languages: Tashlhiyt as the control language, Mandarin (limited incidence of false gemination), French (higher incidence of false gemination arising from word contact or vowel deletion), and Japanese (gemination restricted to word-medial obstruents and /n/).

Several remarks are in order. First, word-initial /t/-/tt/ is very difficult to perceive in the absence of acoustic cues to duration, even for Tashlhiyt listeners [17]. We nevertheless included this contrast for sake of completeness. Second, regarding Japanese, the native "yamato" lexical stratum bans geminate voiced obstruents. However, established loanwords such as /eggu/ 'egg' or /beddo/ 'bed' are quite common [18]. Also, Japanese /nn/ may be special for gemination as it most often occurs as a concatenated geminate (in Sino-Japanese compounds or /-n/#/n-/ contacts). In Mandarin, /nn/ can only be a concatenated geminate, as in 山鳥 *shān niǎo* 'mountain bird'. Note that we recruited Mandarin-speaking subjects in Taiwan: they all had Mandarin as their main language but some of them also had a good command of the Hakka and/or Minnan dialects spoken in Taiwan, in which /p, t, k/ false geminates do occur, as in Cantonese. Since [12] reported an advantage of mainland China Cantonese- over Mandarin-speaking listeners in the perception of Japanese geminates, we collected detailed language background data on Taiwanese participants for a possible effect of their familiarity with Hakka or Minnan.

Keeping these points in mind, several predictions can be made. First, given the incidence and phonological status of gemination in the four languages under scrutiny, we expect the best performance for the Tashlhiyt control group, followed by the Japanese, French and Chinese groups in that order, with possible differences within the Chinese group. Weak differences across groups for the non-initial positions plus poor nonnative performance in initial position would support [16]. In more detail, Japanese listeners might have extra trouble with /d/-/dd/ and possibly /n/-/nn/, whereas Mandarin-dominant Taiwan listeners might find the latter contrast easier.

2. DISCRIMINATION EXPERIMENT

We tested the nonnative language groups (Chinese, French, and Japanese) and the control Tashlhiyt group on their discrimination of Tashlhiyt geminate-singleton contrasts, using the AXB paradigm.

2.1. Methods

2.1.1. Participants

Fifty-four Chinese-speaking students at Tsing Hua University in Hsinchu, Taiwan (aged 18-47 y, mean 23.1, SD 5.0; 21 males), 20 Japanese-speaking students at Sophia University in Tokyo (20-28 y, mean 22.4, SD 2.4; 5 males), 21 French-speaking students at Université Sorbonne Nouvelle in Paris (20-39 y, mean 25.1, SD 5.5; 5 males), and 26 Tashlhiyt-speaking students at Ibnou Zohr University in Agadir (18-53 y, mean 33.2, SD 12.4; 16 males) participated in the experiment voluntarily. Nonnative participants never had exposure to Tashlhiyt. French participants were naïve to any true-gemination language. Based on the collected metadata, Taiwan participants were either clearly Mandarin-dominant speakers (n=22), or with good (n=14) or moderate (n=16) practice of Minnan and/or Hakka. Japanese participants had not been exposed to other gemination language than Japanese. Participants reported no hearing deficit or any kind of language impairment.

2.1.2. Stimuli and design

Twelve geminate-singleton contrasts (4 consonants x 3 positions) of Tashlhiyt were used (/d/: *diɤ-ddiɤ,*

tidi-tiddi, fad-fadd; /t/: *tili-ttili, juti-jutti, jufat-jufatt*; /s/: *sir-ssir, tisi-tissi, ifis-ifiss*; /n/: *niʁ-nniʁ, inas-innas, imun-imunn*). This made a total of 24 words. Four repetitions of each item produced in isolation by a native speaker were retained as experimental stimuli. Tashlhiyt /d/ is entirely voiced (i.e., [d]) and differs from /dd/ by a longer voiced closure duration (geminate-to-singleton duration ratio ~2.5 in the recorded materials). The mean geminate-to-singleton duration ratios were 2.01, 2.25, and 2.53 for initial, medial, and final position, respectively. Closure could not be measured acoustically for word-initial /t/ and /tt/ but electropalatographic articulatory data show much longer closures for geminates than single consonants in absolute initial position [14].

As can be seen, the duration ratios increased from the initial to the final position. They were the highest for /n/ and /d/ and the lowest for /s/. Apart from consonant duration, gemination was not accompanied by *f0* or intensity cues in our stimuli (except the smaller intensity of prevoicing with /dd/ than /d/, as found in [16]). Yet, gemination was *accompanied* by consistent, however weak secondary duration cues in flanking vowels. The vowel preceding word-final consonants, was shorter before geminates than singletons (158 < 214 ms, p<.01); this trend was much weaker in word-medial position (94 < 108 ms, n.s.). A weak trend in the opposite direction was found in the following vowel (194 vs. 188 ms, n.s.).

Each contrast was presented 16 times (four times for each of the AAB, ABB, BBA, and BAA patterns), hence 192 trials divided in two blocks of 96 trials with the same distribution of the 12 contrasts. Each stimulus appeared equiprobably in each within-AXB position. The test phase was preceded by a training phase of eight trials on six contrasts –not used in the test trials– bearing on manner (*dar-daʁ),* voicing (*dar-tar, giji-kiji,* each used twice), or gemination (*fit-ffit, kiji-kkiji*, and *tafa-ttafa*).

2.1.3. Procedure

Nonnative participants were tested individually in a quiet room and received the speech stimuli through professional quality headphones. On each AXB trial, participants were presented with the stimuli and had to press one of the keys labeled “1” or “3” to indicate whether X was a better category match to the first (A) or third (B) stimulus. The inter-stimulus (offset to onset), inter-trial, and inter-block intervals were set to 1, 4, and 9 s, respectively. Response times (RTs) were measured from the onset of the B stimulus, with a 3 s time-out. The experiment was run by a custom-made Python program [19], using the Pygame extension for timing control and RT measurement. For the Tashlhiyt-native subjects, the experiment was run using an on-line PsyToolkit [20-21] script, which replicated the settings of the off-line Python program. All participants received feedback on accuracy, time-out, and RT in the training phase, and only on possible time-out in the test phase, in order to foster speed and attention.

2.2. Results

The data of two Taiwanese and two Tashlhiyt participants were discarded due to high miss-rates, leaving 52 Taiwanese, and 24 Tashlhiyt subjects for accuracy and RT data analyses. We first looked whether Taiwanese subjects should be partitioned into subgroups, based on their metadata: 22 subjects were clearly Mandarin-dominant (native mandarin-speaking parents, little or no practice of Minnan or Hakka), 14 had substantial practice of Minnan or Hakka (both parents native speakers of one of these dialects, high self-rating in hearing/speaking fluency), and 16 were in-between. The 14 Minnan- or Hakka-fluent subjects did not perform better than the 22 Mandarin-dominant ones, as [12] would suggest. On the contrary, their average proportion of correct discrimination was numerically lower (.75 vs. .77), and this held for most position × consonant conditions. We therefore treated the Chinese-speaking group as a single, unified group.

We ran analyses of variance on the unfiltered d-prime and correct-response RT data with the following fixed factors: Group (four language groups) as a between-subject factor, Position (three positions), Consonant (d, t, s, n), and Block (two blocks) as within-subject factors. All the contrasts were quite well perceived except word-initial /t/-/tt/, which was perceived at chance level by all four groups. We therefore present the data separately for word-initial /t/ and /d, s, n/ in Figs. 1 and 2. Block was marginal for d-prime, p=.055 (numerical trend for increased performance in block 2) but n.s. for RTs. The other factors were all significant, ps<.0001. More detail is given in the next sections. We begin with d-prime.

The four groups did not differ for word-initial /t/, F<1 (~55% correct); for word-initial /d, s, n/ only, the pattern was Tashlhiyt > French > Japanese ≈ Chinese, p<.0001, p<.0001, and F<1, respectively; for word-medial all four /d, t, s, n/, it was Tashlhiyt ≈ French > Japanese > Chinese, F<1, p<.05, p<.01, respectively; word-finally, it was Tashlhiyt ≈ French >> Japanese ≈ Chinese, F<1, p<.0001, F<1, respectively. For /d, s, n/ only, Tashlhiyt subjects performed near ceiling regardless of Position, which was not significant, p>.05. For all four /d, t, s, n/, the three nonnative groups’ performance differed with Position, following the pattern initial << medial ≥ final, p<.0001, p=.16, respectively. The same pattern

held for Tashlhiyt subjects on /t/ alone. Overall, conflating the medial and final positions, all four groups performed very well, with Tashlhiyt and French subjects near ceiling (~96%), somewhat better than Japanese and Chinese subjects (~90%). In contrast, the three nonnative groups performed quite poorly on word-initial contrasts, as shown in Fig. 1.

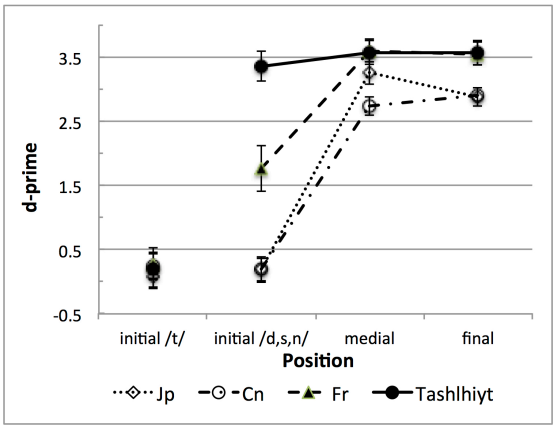

Figure 1: d-primes according to Position and Group, separately for /t/ and /d, s, n/ word-initially.

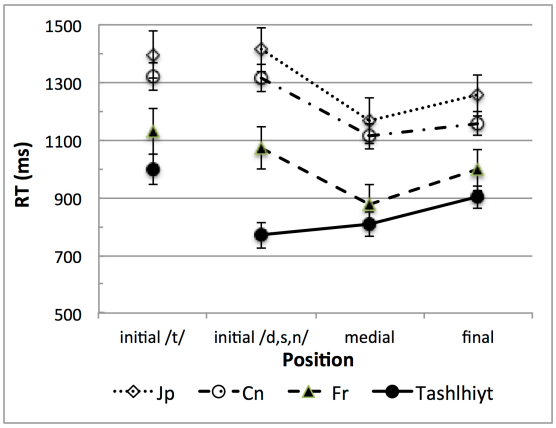

Figure 2: RTs according to Position and Group, separately for /t/ and /d, s, n/ word-initially.

The RT data paralleled well this overall picture. Tashlhiyt subjects were the fastest, $p<.0001$. In non-initial position, the order in RTs was Japanese ≥ Chinese >> French ≥ Tashlhiyt, $F<1$, $p<.0001$, $p=.23$, respectively. Word-initially, it was Japanese ≥ Chinese >> French > Tashlhiyt, $p=.23$, $p<.0001$, $p<.001$, respectively. Nonnative groups responded faster for non-initial than initial position (all four consonants) by ~180 ms, $p<.0001$, with the fastest RTs in medial position, $F(1,90)=99.3$, $p<.0001$. In contrast, Tashlhiyt subjects were faster on non-final (799 ms), than final (918 ms) positions, $p<.0001$ for /d, s, n/, and the slowest for word-initial /t/ (1000 ms vs. 809-902 ms non-initially), $p<.0005$.

In Japanese, /dd/ is nonnative and /nn/ is a weaker case of gemination, hence the prediction of a lesser performance on /d/-/dd/ or /n/-/nn/. This prediction is borne out only word-initially but clearly not in non-initial position (see Table 1). As for Taiwanese, they performed the worst on /n/-/nn/ although /nn/ is the sole concatenated geminate in Mandarin. Unpredictably, /s/-/ss/ was generally the best-perceived contrast word-initially but the worst perceived in non-initial positions.

consonant	initial	medial	final	means
d	**-0.24**	**3.84**	**3.40**	**2.33**
t	0.07	3.04	3.30	2.14
s	**1.11**	**2.82**	**1.89**	**1.94**
n	**-0.31**	**3.32**	**2.95**	**1.99**
means	0.16	3.26	2.88	2.10

Table 1: Japanese d' data: Consonant x Position.

3. DISCUSSION

Both the accuracy and RT data pointed to a consistent pattern: Nonnative listeners performed well on Tashlhiyt word-medial or -final gemination contrasts, but much less well on word-initial contrasts. (The case of word-initial /t/-/tt/ is special: it was perceived at chance level in all four groups of listeners and induced the longest RTs.) Therefore it seems that, in non-initial position, singleton-geminate contrasts are much easier than can be expected from the incidence of gemination in the listeners' languages: Japanese listeners did worse than French and not better than Mandarin-speaking Taiwanese. French listeners did as well as Tashlhiyt listeners. The data thus fit well with [16]: sensitivity to the "beat" given by successive salient events (e.g., the vowels surrounding a consonant) is likely universal; in contrast, sensitivity to word-initial consonant duration, where no beat can help, likely requires language-specific learning.

An alternative account may be provided by the secondary durational cues that accompany gemination in Tashlhiyt: vowels are shorter before geminate than singleton consonants, which should help perceiving gemination contrasts in non-initial position. Yet, this shortening is more marked for word-final than -medial geminates (26 vs. 12%), which runs counter to the perceptual advantage for medial over final gemination we observed. The "beat" account provides a simple explanation for this differential: the beat given by two vowels flanking a medial consonant must be more salient than that given by a preceding vowel and a following consonant-release because the latter phonetic event is conceivably perceptually weaker than a vowel is.

At any rate, further investigation of the potential contribution of vowel duration as a secondary cue to gemination is warranted as a follow-up study.

4. ACKNOWLEDGEMENTS

This study was supported by the French ANR (ANR- 10-LABX-0083, Labex-EFL).

5. REFERENCES

[1] Han, M. S. 1992. The timing control of geminate and single stop consonants in Japanese: A challenge for non-native speakers. *Phonetica* 49: 102–27.

[2] Harada, T. 2006. The acquisition of single and geminate stops by English-speaking children in a Japanese immersion program. *Studies in Second Language Acquisition* 28: 601–32.

[3] Toda, T. 2007. Issues regarding geminate consonants in Japanese language education. *Journal of the Phonetic Society of Japan* 11: 35–46.

[4] Kingston, J., Kawahara, S., Chambless, D., Mash, D., Brenner-Alsop, E. 2009. Contextual effects on the perception of duration. *J. Phon.* 37, 297–320.

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

[6] Hayes, R. 2001. The perception of novel phoneme contrasts in a second language: A developmental study of native speakers of English learning Japanese singleton and geminate consonant contrasts. *Coyote Papers* 12, 28–41.

[7] Sonu, M., Kato, H., Tajima, K., Akahane-Yamada, R., Sagisaka, Y. 2013. 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: 373–98.

[8] Tsukada, K., Cox, F., Hajek, J., Hirata, Y. 2018. Non-native Japanese learners' perception of consonant length in Japanese and Italian. *Second Language Research* 34, 179–200.

[9] Tsukada, K., Rong, Y., Kim, J. Y., Han, J. I., Hajek, J. 2021. Cross-Linguistic Perception of the Japanese Singleton/Geminate Contrast: Korean, Mandarin and Mongolian Compared. *Proc. Interspeech*, 3910-3914.

[10] Hung, H. Y. 2012. Perception of Japanese geminate stops among Taiwanese learners of Japanese. *Journal of the Phonetic Society of Japan* 16. 15–27.

[11] Lu, S., Chen, L., Yamada, J. 2016. Taiwanese learners' durational patterns for Japanese obstruent geminates/singletons and adjacent vowels. *Perceptual and Motor Skills* 122. 67–75.

[12] Ren, H. 2022. Perception of Japanese singleton and geminate contrasts: A case of Chinese learners with different dialectal backgrounds. *Front. Psychol.* 13, 1070107.

[13] Uchida, T. 1993. Characteristics of auditory cognition of long vowels and double consonants for Chinese students in learning Japanese language. *Japanese Journal of Educational Psychology* 4. 414–23.

[14] Ridouane, R. 2007. Gemination in Tashlhiyt Berber: an acoustic and articulatory study. *J. Int. Phon. Assoc*. 37, 119–142.

[15] Ridouane, R. 2010. Geminate at the junction of phonetics and phonology. In: Fougeron, C., Kühnert, B., D'Imperio, M., Vallée, N. (eds), *Papers in Laboratory Phonology X*. Mouton de Gruyter, 61–90.

[16] Hallé, P., Ridouane, R., Best, C. 2016. Differential difficulties in perception of Tashlhiyt Berber consonant quantity contrasts by native Tashlhiyt listeners vs. Berber-naïve French listeners. *Front. Psychol*. 7, 209.

[17] Ridouane, R., Hallé, P. 2017. Word-initial geminates: from production to perception. In Haruo Kubozono (ed.), *Phonetics and Phonology of Geminate Consonants,* 66-84. Oxford Studies in Phonology and Phonetics - Oxford University Press.

[18] Kawahara, S. 2015. The phonetics of sokuon, or geminate obstruents. In: Kubozono, H. (ed.) *Handbook of Japanese phonetics and phonology*. De Gruyter Mouton, 43–78.

[19] Yu, S. 2018. Psypsy program. Github repository: https://github.com/shi4yu2/psypsy.

[20] Stoet, G. 2010. PsyToolkit: a software package for programming psychological experiments using Linux. *Behavior Research Methods* 42, 1096–1104.

[21] Stoet, G. 2017. PsyToolkit: a novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology* 44, 24–31.