ACOUSTIC PROPERTIES OF PALATALIZED CONSONANTS IN JAPANESE

Yu Tanaka

Doshisha University
yutanak@mail.doshisha.ac.jp

ABSTRACT

Palatalization affects the acoustics of segments in various ways. The present study investigates the spectral and durational properties of Japanese palatalized stops and fricatives in the data of 90 female and 111 male speakers taken from a large speech corpus. The results show, as expected, that palatalized consonants tend to have higher F2 and higher CoG (except /sʲ/) of their release or frication noise and mostly (all but labial /pʲ/) have longer VOT or total duration than their plain counterparts. The findings suggest that phonemically palatalized consonants in Japanese, though variously realized in phonetic terms, exhibit cross-linguistically common acoustic characteristics of palatalization.

Keywords: Palatalization, F2, CoG, VOT, Japanese

1. INTRODUCTION

The phonology of Japanese systematically contrasts plain and palatalized consonants [1, 2]; many consonants have a plain form /C/ and a palatalized form /Cʲ/. As is common cross-linguistically [3, 4, 5], the actual phonetic realizations of the phonemically palatalized consonants vary. Some are produced with secondary palatalization, as in /pʲ/ [pʲ] and /kʲ/ [kʲ], while others are true palatals or alveolo-palatals, as in /hʲ/ [ç] and /sʲ/ [ɕ].

There is a large body of literature describing the phonetic correlates of palatalization in other languages (see e.g. [6, 7, 8, 9, 10] and references therein). Common acoustic effects of palatalization on stops and fricatives are as follows. When palatalized, they tend to have release or frication noise that is louder, longer, and spectrally distinct, with higher second formant frequency (F2) and higher center of gravity (CoG) values (except for sibilants) than their plain counterparts.

Relatively little research has examined the acoustic properties of Japanese palatalized stops and fricatives. One experimental study [11] analyzed /bʲa, gʲa, mʲa, nʲa, ɾʲa/ and /ba, ga, ma, na, ɾa/, showing that the syllable duration of the former group is longer; however, it did not examine the duration of voiceless stops or fricatives. Other studies [12, 13] conducted acoustic analyses of word-medial /t, k, tʲ, kʲ/ in recordings taken from a speech corpus, reporting that /kʲ/ tends to have shorter closure and longer release noise than /k/, but no such compensatory effect is found between /t/ and /tʲ/. Note, however, that /tʲ/ [tɕ] involves affrication, and direct comparison with plain /t/ [t] or with palatalized dorsal /kʲ/ [kʲ] is not easy. The study also excluded the labial stop /pʲ/ and fricatives such as /sʲ/ and /hʲ/ from consideration.

The current study provides further analyses of palatalized consonants in Japanese. Specifically, I measure the spectral and durational features of palatalized stops and fricatives /pʲ, kʲ, sʲ, hʲ/, and compare them to those of plain /p, k, s, h/.

2. METHOD

Following [12, 13], I used the *Corpus of Spontaneous Japanese* (CSJ) [14], which contains speech data of 90 female and 111 male speakers from various age groups with detailed annotations. Fig. 1 is an image of annotated speech containing *ky* /kʲ/ in the word *keNkyuH* /keɴkʲuː/ 'research'. The palatalized stop is segmented into closure (<cl>) and release (labeled *ky*) in the segment ("seg") tier.

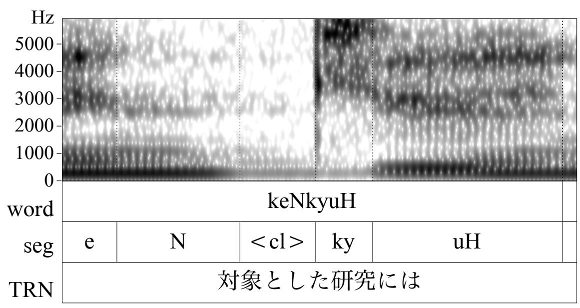

Figure 1: Annotated speech in CSJ

To examine the acoustic properties of palatalized stops, I measured the duration and midrange F2 and CoG of the release of /pʲ/ and /kʲ/, and those of /p/ and /k/ for comparison. (The midrange was defined

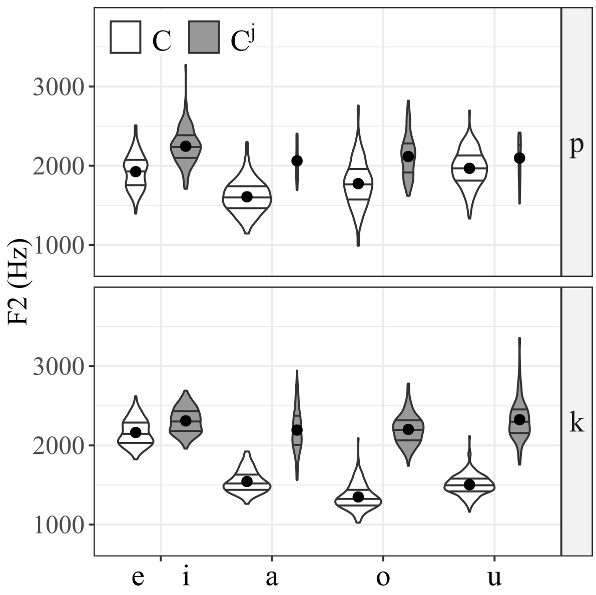

Figure 2: F2 of stop release noise

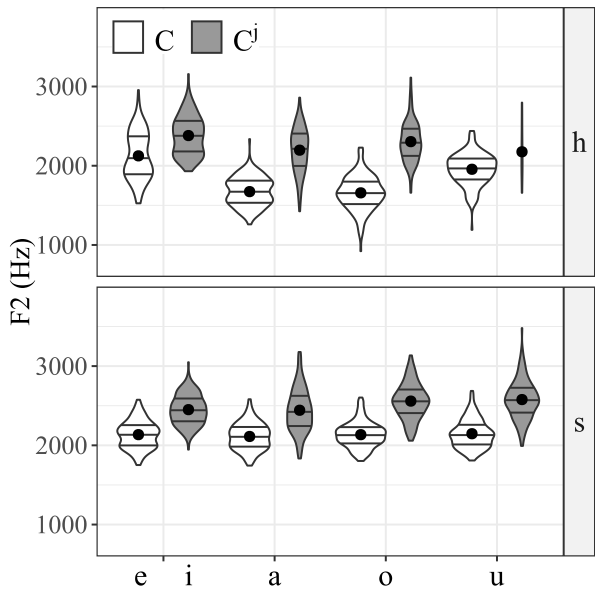

Figure 3: F2 of frication noise

as a 20ms window, or a 5ms window if the release was shorter than 20ms, centered at the midpoint.) For fricatives, I measured the duration and midrange F2 and CoG of /s^j/ and /h^j/, and those of /s/ and /h/. (I excluded cases where the following vowel was devoiced.) The measurement procedure was automated using a modified version of the *formant-logger* script [15] on *Praat* [16]. In total, 157,823 tokens of the target segments were analyzed.

3. RESULTS

3.1. F2 and CoG of release and frication noise

I first present the results of F2 of stop release noise. Fig. 2 shows a violin plot of the speakers' mean F2 of plain /C/ and palatalized /C^j/ by consonant type and vowel context. Lines (—) and dots (•) represent quartiles and grand means, respectively. There are differences in sample size, as shown by width, since some speakers' data do not contain any instances of the relevant sequences. The /e/ and /i/ contexts are given for reference; in principle, /e/ only occurs with plain /C/ (i.e. [pe], [ke]), and /i/ only with palatalized /C^j/ (i.e. [p^{j}i], [k^{j}i]) on the surface level.

It can be seen that /C^j/ has relatively high F2 in general. A mixed-effects model was constructed on *R* [17, 18, 19] to predict F2 values with the presence of palatalization (/C/, /C^j/), the consonant type (/p/, /k/), and their interaction as its predictors. The random structure included by-speaker random intercepts and slopes, as well as random intercepts for the vowel context (/a/, /o/, /u/). The data in the /e/ and /i/ contexts with no real contrast in palatalization were excluded, and the remaining 117,120 tokens were used as fitting data.

The model shows that /C^j/, or palatalization, significantly raises F2 in /p^j/ ($\beta = 525.77$, $t = 20.47$, $p < 0.001$) with the baseline intercept being /p/; the same is true for /k^j/ ($\beta = 830.26$, $t = 77.12$, $p < 0.001$) if the baseline is set to /k/. The main effect of the consonant type is also significant; /p/ has higher F2 than /k/ ($\beta = 219.44$, $t = 17.93$, $p < 0.001$). (However, see §4 for discussion of possible measurement errors.) In addition, the interaction term of the two predictors is significant, indicating that the F2-raising effect of /C^j/ is greater on /k^j/ than /p^j/ ($\beta = 304.49$, $t = 11.43$, $p < 0.001$).

Fig. 3 plots the F2 values of fricatives, contrasting /C/ and /C^j/. Here too, /C^j/ shows consistently higher F2 in every vowel context. Another mixed-effects model was applied in the same manner as above, except that the consonant type was fricatives (/h/, /s/). It predicts that palatalization raises F2 in both /s^j/ ($\beta = 403.73$, $t = 28.19$, $p < 0.001$) and /h^j/ ($\beta = 587.48$, $t = 38.70$, $p < 0.001$). /h/ generally has lower F2 than /s/ ($\beta = -402.33$, $t = -41.05$, $p < 0.001$), but palatalization has a greater effect on /h^j/ than /s^j/, as is indicated by a significant interaction ($\beta = 183.75$, $t = 8.55$, $p < 0.001$).

Fig. 4 gives the results for the CoG of stop release noise. Generally speaking, /C^j/ shows higher CoG than /C/. A mixed-effects model indicates that the effect of /C^j/ is significant in both /k^j/ ($\beta = 1819.92$, $t = 43.96$, $p < 0.001$) and /p^j/ ($\beta = 625.76$, $t = 9.69$, $p < 0.001$). Also, CoG is higher in dorsal /k/ than

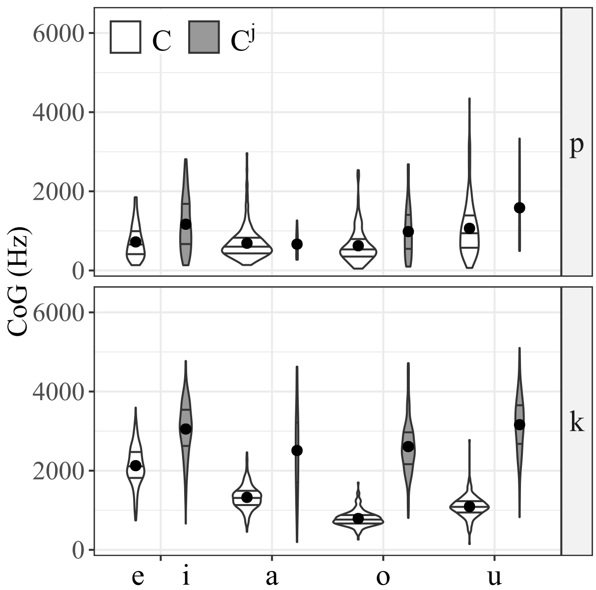

Figure 4: CoG of stop release noise

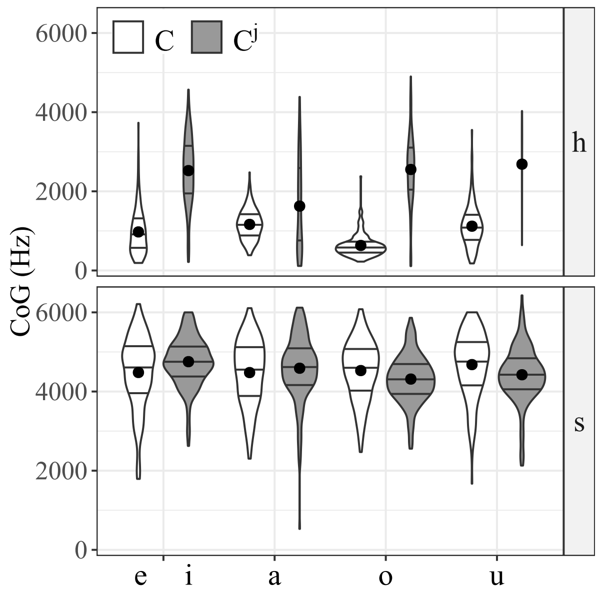

Figure 5: CoG of frication noise

labial /p/ (β = 446.19, t = 16.78, $p < 0.001$), and a significant interaction further shows that the effect of palatalization is greater on the dorsal (β = 1194.16, t = 16.08, $p < 0.001$).

Fig. 5 shows the results for fricatives. As can be seen, /s/ has higher CoG than /h/ (β = 3602.75, t = 69.18, $p < 0.001$). The effect of palatalization is significant, with CoG higher in /h^j/ than in /h/ (β = 1221.57, t = 17.63, $p < 0.001$). However, a significant interaction between palatalization and the consonant type suggests that no such relationship holds for /s^j/ and /s/ (β = –1355.83, t = –15.87, $p < 0.001$); in fact, when directly compared, the CoG of /s^j/ is lower than that of /s/ (β = –134.265, t = –2.90, p = 0.004). That is, palatalization rather lowers CoG in the case of sibilants.

In sum, palatalization is associated with high F2 and CoG of stop release and frication noise, except that sibilant /s^j/ has relatively low CoG.

3.2. VOT and total duration

Turning to the durational properties of palatalized consonants, I present in Fig. 6 the duration of voiceless stop release noise, namely voice onset time (VOT). Notice that in the /k/ conditions, /C^j/ tends to have longer VOT than /C/. As for the /p/ conditions, in contrast, the differences between /C/ and /C^j/ seem at best marginal.

A mixed-effects analysis reveals that the effect of /C^j/ is significant in /k^j/ (β = 9.95, t = 16.84, $p < 0.001$), but its interaction with the consonant type of /p/ is negatively significant (β = –8.96, t = –7.92, $p < 0.001$); /C^j/ actually does not lengthen VOT in the case of /p^j/ (β = 0.99, t = 1.03, p = 0.305). Additionally, the main effect of the consonant type is significant; /k/ has longer VOT than /p/ (β = 9.11, t = 22.29, $p < 0.001$).

Fig. 7 shows the duration of fricatives. For the most part, /C^j/ is longer than /C/. Statistically, palatalization lengthens both /s^j/ (β = 6.07, t = 8.13, $p < 0.001$) and /h^j/ (β = 9.49, t = 8.24, $p < 0.001$), although its significant interaction with the consonant type indicates that the effect is stronger in the latter (β = 3.42, t = 2.53, p = 0.013). Also, /s/ is longer than /h/ (β = 23.88, t = 37.39, $p < 0.001$).

The results as a whole thus show that palatalization lengthens the VOT of stops and duration of fricatives, except in the case of labial /p^j/ where no such lengthening effect is found.

4. DISCUSSION

Acoustic analyses of plain and palatalized stops and fricatives in Japanese have shown that palatalization lengthens their release and frication noise and raises their F2 and CoG, although these effects are not observed with certain segment types. These results, including the null effects, in fact align well with cross-linguistically common patterns, and also hold some theoretical implications.

As stated above, palatalized consonants tend to be longer. [9] further reports that in Connemara Irish, palatalization lengthens stop release duration, but the effect is weaker in labials than in coronals. This is parallel to the null result regarding the VOT of

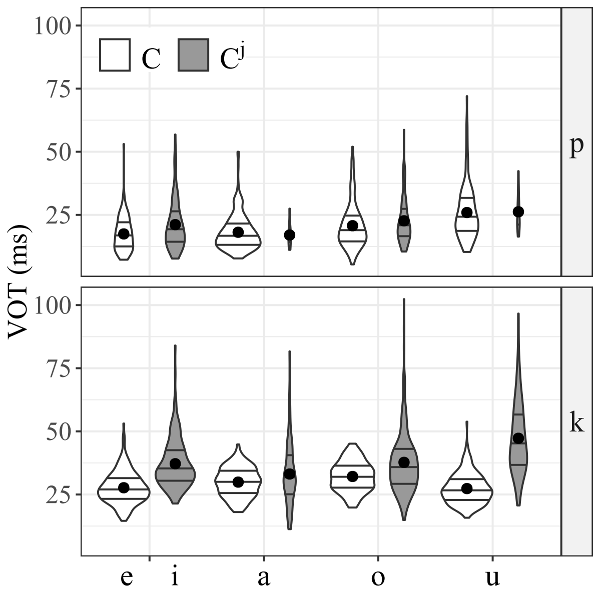

Figure 6: VOT of stops

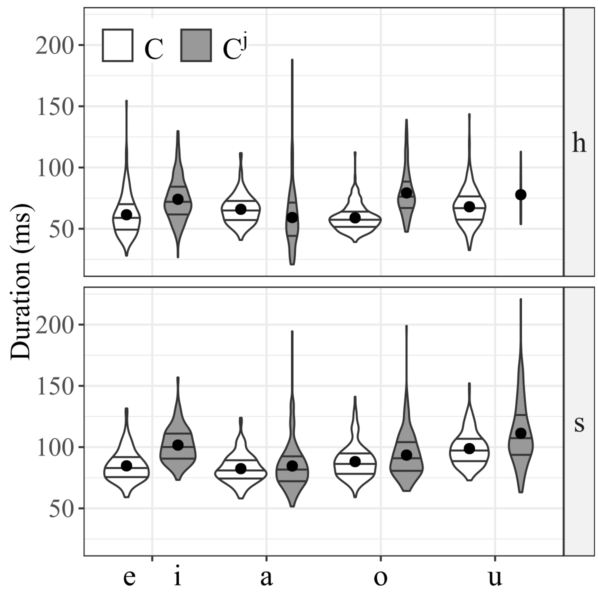

Figure 7: Duration of fricatives

/pʲ/ in Japanese (Fig. 6). Labial stops are known for already having relatively short VOT [20, 21]. The facts above suggest that they even resist the lengthening effect of palatalization.

The results of CoG are also compatible with cross-linguistic data. Secondary palatalization itself tends to raise CoG, but alveolo-palatal and post-alveolar sibilants such as [ɕ] and [ʃ] inherently show lower CoG than alveolar [s] (see [8, 9, 22, 23]). As we have seen, Japanese palatalized consonants also have higher CoG than their plain counterparts, except that /sʲ/, phonetically realized as alveolo-palatal [ɕ], has lower CoG than /s/ [s] (Fig. 5).

The results for fricative duration (Fig. 7) can be relevant to debates on representations in traditional Japanese phonology. For alveolo-palatal sibilants such as [ɕ] and [ʑ], some scholars posit underlying /Cʲ/, as in /sʲ/ and /zʲ/, based on the fact that they pattern with other phonemically palatalized consonants [1, 2]. Others propose representations that are closer to the surface forms, as in /ɕ/ and /ʑ/, partly because they are phonetically single sounds [11, 24]. Although arguments about phonological representations should be based on phonological evidence, the results here suggest that [ɕ] (as well as [ç]), which is transcribed as a single sound, shows common characteristics of "palatalized" consonants; it is longer, as if involving two articulatory gestures, as implied by the notation /Cʲ/.

The results for the F2 of stop release noise (Fig. 2.) also require some attention. Recall that plain /p/ has higher F2 than /k/. This is somewhat unexpected given that labials generally lower F2. It is possible that the automated measurement had errors when F2 was too low, especially in the cases of /po/ and /pu/. That said, the low F2 values of /k/ in back vowel contexts seem genuine. Impressionistically, some speakers' production of /k/ before /a, o, u/ is quite back, sometimes sounding like uvular [q]. This may be interpreted as a kind of enhancement effect, as seen in Russian [25]; not only are palatalized consonants fronted, but non-palatalized ones are backer.

Another caveat is that the phonological contexts and the frequencies of the target segments are not uniform. In principle, /p(ʲ)/ occurs word-medially as a geminate or after a moraic nasal [26] (e.g. /happa/ 'leaf', /kaɴpai/ 'toast'). The occurrence of /h(ʲ)/ is also limited to word-initial position in the native vocabulary. Even though the analyzed corpus data include recent loans where these restrictions do not hold (e.g. /paɴ/ 'bread'), the results above may still have been affected by the distributional and frequency differences. Conducting a production experiment using nonce words with more control on these factors would solve the issue.

There are other possible acoustic features of palatalization which this study has not explored. For example, previous studies show that the intensity of stop release and F2 of following vowels could also differentiate palatalized and plain consonants. Intensity was especially difficult to address using speech data of different speakers recorded in various environments. Future experimental work should examine these other acoustic properties with respect to palatalized consonants in Japanese.

5. ACKNOWLEDGMENTS

I thank two anonymous reviewers for their comments. This study was partially supported by JSPS Kakenhi Grants 20K13019 and 22K13106.

6. REFERENCES

[1] J. Ito and A. Mester, *Japanese Morphophonemics*. Cambridge, MA: MIT Press, 2003.

[2] L. Labrune, *The Phonology of Japanese*, ser. The Phonology of the World's Languages. Oxford: Oxford University Press, 2012.

[3] D. N. S. Bhat, "A general study of palatalization," in *Universals of Human Language*, J. Greenberg, Ed. Stanford: Stanford University Press, 1978, vol. 2: Phonology, pp. 47–92.

[4] N. Bateman, "On the typology of palatalization," *Language and Linguistics Compass*, vol. 5, no. 8, pp. 588–602, 2011.

[5] A. Kochetov, "Palatalization," in *The Blackwell Companion to Phonology*, M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice, Eds. Malden: Wiley-Blackwell, 2011, vol. 3: Phonological Processes, pp. 1666–1690.

[6] ——, *Production, Perception, and Emergent Phonotactic Patterns: A Case of Contrastive Palatalization*. New York: Routledge, 2002.

[7] ——, "Testing licensing by cue: A case of Russian palatalized coronals," *Phonetica*, vol. 63, no. 2-3, pp. 113–148, 2006.

[8] ——, "Acoustics of Russian voiceless sibilant fricatives," *Journal of the International Phonetic Association*, vol. 47, no. 3, pp. 321–348, 2017.

[9] M. Ní Chiosáin and J. Padgett, "An acoustic and perceptual study of Connemara Irish palatalization," *Journal of the International Phonetic Association*, vol. 42, no. 2, pp. 171–191, 2012.

[10] L. Spinu, I. Vogel, and H. T. Bunnel, "Palatalization in Romanian – Acoustic properties and perception," *Journal of Phonetics*, vol. 40, no. 1, pp. 54–66, 2012.

[11] A. Nogita, "Arguments that Japanese [Cj]s are complex onsets: Durations of Japanese [Cj]s and Russian [C^j]s and blocking of Japanese vowel devoicing," *Working Papers of the Linguistics Circle of the University of Victoria*, vol. 26, no. 1, pp. 73–99, 2016.

[12] M. F. Matsui, "A case study of the temporal features of Japanese palatalized consonants [Nihongo yōon no jikan tokusei ni kansuru jirei kenkyū]," *Theoretical and Applied Linguistics at Kobe Shoin*, vol. 24, pp. 434–443, 2021.

[13] ——, "The temporal features of Japanese palatalized consonants [Onsei jizoku jikan kara mita Nihongo yōon no seishitsu]," *Proceedings of the 38th Meeting of the Japanese Cognitive Science Society*, vol. 38, pp. 434–443, 2021.

[14] National Institute for Japanese Language and Linguistics, "The Corpus of Spontaneous Japanese," 2012, https://ccd.ninjal.ac.jp/csj/index.html.

[15] K. Crosswhite, "formant-logger," 2009, Praat script.

[16] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer," 2022, Computer program. Version 6.3.03. http://www.praat.org/.

[17] D. Bates, M. Maechler, B. Bolker, and S. Walker, "lme4: Linear mixed-effects models using 'Eigen' and S4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48., 2015.

[18] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: Tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, no. 13, pp. 1–26., 2017.

[19] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2022.

[20] L. Lisker and A. S. Abramson, "A cross-language study of voicing in initial stops: Acoustical measurements," *Word*, vol. 20, no. 3, pp. 384–422, 1964.

[21] T. Cho and P. Ladefoged, "Variations and universals in VOT: Evidence from 18 languages," *Journal of Phonetics*, vol. 27, pp. 207–229, 1999.

[22] A. Jongman, R. Wayland, and S. Wong, "Acoustic characteristics of English fricatives," *The Journal of the Acoustical Society of America*, vol. 108, pp. 1252–1263, 2000.

[23] M. Gordon, P. Barthmaier, and K. Sands, "A cross-linguistic acoustic study of voiceless fricatives," *Journal of the International Phonetic Association*, vol. 32, no. 2, pp. 141–174, 2002.

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

[25] J. Padgett, "Contrast dispersion and Russian palatalization," in *The Role of Speech Perception in Phonology*, E. Hume and K. Johnson, Eds. New York: Academic Press, 2001, pp. 187–218.

[26] J. Ito and A. Mester, "Japanese phonology," in *The Handbook of Phonological Theory*, J. Goldsmith, Ed. Cambridge, MA: Blackwell Publishers, 1995, pp. 817–838.