AN ACOUSTIC STUDY OF THE FRICATIVE VOWEL IN TWO ENDANGERED RYUKYUAN LANGUAGES

Shigeko Shinohara[1], Bowei Shao[2]

[1] Laboratoire de Phonétique et Phonologie, CNRS/Université Sorbonne Nouvelle, Paris, France
[2] Département d'études cognitives, École Normale Supérieure, Université PSL, Paris, France
shigeko.shinohara@sorbonne-nouvelle.fr, bowei.shao@ens.psl.eu

ABSTRACT

This study investigates the acoustic properties of the fricative vowel of two varieties of Ryukyuan languages: Aragusuku Yaeyama (AR) and Karimata Miyako (KR) Ryukyuans. Ryukyuans are endangered languages spoken in the Southwest Islands of Japan. The phonetic properties of the fricative vowel have been an issue of debate. We aim to characterize the frication in the acoustic signals of this vowel. We compared the zero-crossing rates (ZCR) between the fricative vowel and the regular vowel /i/. The results indicated that the fricative vowel has frication noise superimposed on its formant structure in the 5-8 kHz region. This noise distinguished the fricative vowel from /i/ in AR but not in KR based on ZCR. Nevertheless, the fricative vowels in both languages have a similar quantity of frication noise. These results suggest that the fricative vowels in AR and KR have a similar acoustic profile to those reported in other languages.

Keywords: Fricative vowel, Ryukyuan, Endangered language, Zero-crossing rate, GAMM

1. INTRODUCTION

Ryukyuan languages are spoken on the Southwest Islands of Japan. All six Ryukyuan languages were identified as endangered by UNESCO [15]. Some dialects of Miyako and Yaeyama Ryukyuans possess, cross-linguistically rare, fricative vowels (FVs) [6, 7, 13].

FVs are syllable nuclei composed of fricative sounds and are found in several unrelated languages [2, 13]. One well-known example is the apical vowels in Chinese languages, which is the most studied example of FVs [3, 4, 5, 11, 14, 23, 24, 25]. Acoustic analyses of FVs showed supralaryngeal frication generated at labial or alveolar place superimposed on voicing and that the frication evolves during the vowel interval. This frication could be observed following sibilant, plosive, and nasal onsets [23, 25]. Articulatory studies using ultrasound imaging showed a fricative-like tongue shape with medial grooving for the alveolar FV in Chinese languages [5, 25].

Data collection of the FV was planned in Miyako Islands for 2020 but has been suspended due to the pandemic. This situation drove us to explore the audio recordings of the Aragusuku dialect of Yaeyama Ryukyuan (AR) and the Karimata dialect of Miyako Ryukyuan (KR) made available by NII-SRC [17]. Ryukyuans are generally spoken by the elderly population. The number of speakers of each dialect is not censused. In 2022, the population over 65 years old in Aragusuku counts 8 and that of Karimata is 190. AR is the only dialect preserving an FV among Yaeyama Ryukyuan [7] whereas the FV is found in most of the dialects on the Miyako main Island but in distinct preconsonantal contexts depending on the dialect (see Sec. 2).

Figure 1: The Ryukyu Islands and the areas where the two languages are spoken: Karimata (KR) on Miyako Islands and Aragusuku (AR) Islets of Yaeyama Islands (Okinawa Travel Guide, link)

The quality of the FV in Ryukyuans has been an issue of debate in Japanese dialectology. It has been variably called an 'apical vowel' [8, 22] or a 'central vowel' [6, 20]. [20] investigated the formant structure of the vowels in the same database. The F2 values of the FVs were located midway between /i/ and /u/ in AR and KR. KR's FV was slightly more front, closer to /i/. Frication noise was observed at 3 kHz along with the presence of F1 and F2 in AR. As a complementary study to [20], we shall examine frication noise in the FV in both regular and devoiced contexts (see below).

2. SOUND STRUCTURES OF AR AND KR

The phonological descriptions below are based on [12] for AR and [19] for KR and supplemented by [7] for AR and [10] for KR. AR contrasts six vowels /a i u e o I/ (/I/ denotes the FV); KR does five /a i u o I/. As we interpret the consonant segments in [12], AR and KR both have /p b t d k g ts f v s z h m n w r j/ as contrastive consonants. They have a similar basic syllable structure (N)(C)V(N). In addition, many of the segments have length contrast. /I/ appears in restricted consonantal environments in both languages. It can follow /p b k g s ts z r/ in AR and /p b k g s ts z/ in KR. In both languages, it can appear without an onset consonant (e.g., /I/ [I:] ‘rice’). The high vowels (including /I/) are devoiced between two voiceless consonants in both AR and KR. They can be devoiced phrase-finally, and the other vowels can also be devoiced in similar environments. Vowels can be devoiced after a voiceless consonant, and a following voiced consonant can also be devoiced (e.g., [kI̥m̥u] ‘liver’).

3. STUDY QUESTIONS AND METHOD

3.1. Study questions

Our study aims to answer three research questions: 1) Do the FVs /I/ in AR and KR contain any frication noise in the acoustic signals?; 2) If they do, how does the frication noise evolve during the entire duration of /I/?; 3) Is there any difference in the degree of frication between the two languages?

3.2. Zero-crossing Rate and Generalized additive mixed model

To measure the degree of frication noise in the FV, we chose Zero-crossing Rate (ZCR) as an indicator. ZCR measures the number of upward and downward crossings of zero dB per second in the waveform, without involving the detection of voicing or pitch. Previous studies on the apical vowel in Chinese languages [3, 23, 25] demonstrated the robustness of this measurement in analyses of voiced segments containing frication noise.

In our study, ZCR was obtained as *PointProcess* objects in Praat [1] within a 20 ms sliding window on the acoustic signals. The ZCR was calculated as the number of zero-crossings divided by the length of the window. 22 data points (the onset and offset of all nucleic segments and 20 data points evenly spaced on each segment) were extracted. This method allowed us to visualize and analyze each segment’s temporally dynamic ZCR patterns.

We chose Generalized Additive Mixed Models (GAMMs) [28, 29] to model the dynamic pattern of FVs, with *mgcv* package v1.8-40 in *R* v4.2.3 [21]. This method has been gaining ground in the investigation of dynamic speech patterns, including temporal and spatial aspects and their interactions [26, 27]. We chose GAMMs because ZCR was expected to change over time [3, 25], and we could also explore if the frication noise was influenced by other factors, notably by the type of onsets.

We made within-language and cross-language comparisons with GAMMs. Within-language comparisons would show the difference in ZCR between an FV /I/ and a regular vowel /i/ in a language. Cross-language comparisons were performed on the normalized ZCR scores, following [3]. The normalization took the ratio between the ZCR of /I/ and that of /s/ for each speaker (i.e., Relative zero-crossing rate (RZCR)). In the models, ZCR of the vowels was estimated over time, with factor smooths for onset and the voicing categories of the onset. Tweedie distributions were used in the model, as ZCR follows a left-skewed, long-tailed distribution. Results were visualized using *tidyverse* v1.3.2 and *tidymv* v3.3.2 [30, 31].

3.3. Speech materials

NII-SRC data contain recordings of a male Aragusuku speaker aged 83 in 1990 [12] and a female Karimata speaker aged 69 in 1996 [19]. The recording corpus included real words with five types of vowels: always-voiced long fricative vowel [I:], short fricative vowel [I] in voiced and devoiced environments, and regular vowel [i] in voiced and devoiced environments. The devoiced [I̥ i̥] tokens were treated separately from the voiced [I i]. The speakers repeated each word three times. The datasets turned out to be relatively small and uneven (AR: 37 [i], 20 [I], 4 [I:], 12 [I̥], 2 [i̥]; KR: 12 [i], 23 [I], 3 [I:], 8 [I̥], 2 [i̥]).

4. RESULTS

4.1. Aragukusu /I/

Fig. 2 presents three tokens of /I/: [I] preceded by a vowel, [maI] ‘rice’; [I] preceded by a voiced affricate, [adzI] ‘flavor’; and [I̥] between voiceless obstruents (i.e., a devoicing environment), [pI̥tu] ‘human.’ Frication noise is observed in the 4-6 kHz region in all three tokens. This suggests that the frication noise is not introduced by the preceding consonant. For the first two voiced tokens, the frication noise is superimposed on a clear formant structure, and the frication seems to diminish and disappear towards the end of [I]. In the case of devoiced [I̥], the frication noise seems to be more intense in the same frequency range.

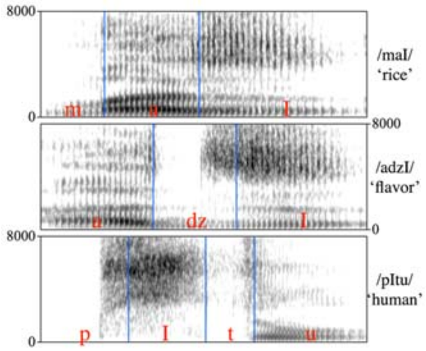

Figure 2: Spectrograms of three realizations of /I/: after a vowel, after a voiced consonant, and between voiceless consonants in Aragusuku.

The ZCR in FVs was further analyzed using GAMMs, compared with the high front vowel [i]. Fig. 3 shows the evolution of ZCR during [i], [I:], and devoiced [I̥] of AR. The devoiced segments' pattern is different from the voiced ones: devoiced segments have much higher ZCR. This is not surprising since when FVs are devoiced, their acoustic signal resembles that of a voiceless fricative consonant, having a much higher ZCR. For the voiced segments, overall, [i] has a lower ZCR and [I] has a higher ZCR, whereas [I:] has a descending pattern: it starts from a high point and gradually reaches the level of [i] and continues further down toward the end of the segment.

The [I:] and the devoiced [i̥] have larger confidence intervals compared to the other segments. This could be explained by the fact that there were fewer tokens of them in the dataset, thus, the fitted values of these two segments should be interpreted with caution.

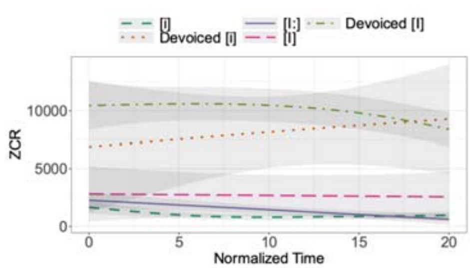

Figure 3: GAMM fitted values of Zero-crossing Rate during /i/ and /I/ in Aragusuku.

The differences among the three voiced segments [I I: i] are presented in Fig. 4. Compared to [i], [I:] has a higher ZCR in the first half of its interval; the second half is less fricated, having a similar ZCR to [i]. The ZCR of [I] is less reliable, however, it still has a higher ZCR overall than [i] does, especially in the middle portion, where the difference is significant.

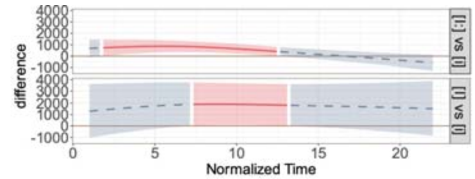

Figure 4: Difference between [i] and the fricative vowel [I, I:] in Aragusuku. The solid lines indicate portions with a significant difference.

4.2. Karimata /I/

Similar to AR, frication noise was observed in /I/ in all different contexts. Fig. 5 presents three tokens of /I/: [I] preceded by a vowel, [I] preceded by a voiced affricate, and [I̥] between voiceless obstruents. Frication noise is observed on the spectrogram of all three tokens but appears to be located at higher frequencies (specifically, in the 5-7 kHz region) than in AR. The frication noise is superimposed on a formant structure for the first two voiced tokens.

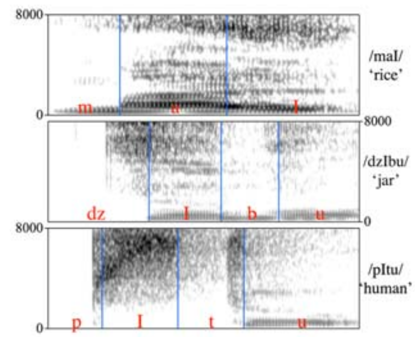

Figure 5: Spectrograms of three realizations of /I/ in Karimata. In /dzIbu/ 'jar,' the onset /dz/ is realized as [ts].

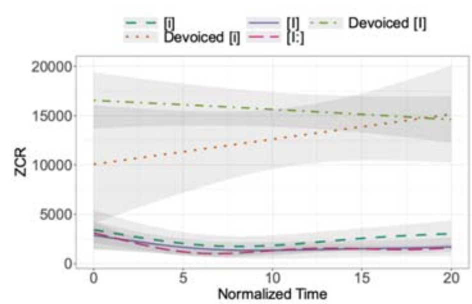

Figure 6: GAMM fitted values of Zero-crossing Rate during /i/ and /I/ in Karimata.

The ZCR of /i/ and /I/ were also analyzed using GAMMs. The first observation is, again, that the devoiced [I̥] and [i̥] pattern together differently from the voiced segments. For the voiced segments, contrary to AR [I], the KR [I] seems to pattern closely with [i], which can also be seen in Fig. 7. The overall difference between [I] and [i] is not significant, and the [I:] seems to have an even lower ZCR towards the end of its duration compared to [i]. However, the size of the dataset calls for a cautious interpretation of the results.

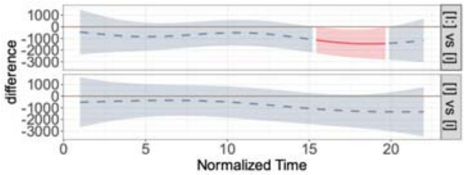

Figure 7: Difference between [i] and the fricative vowel [I, I:] in Karimata. The solid lines indicate a portion with a significant difference.

4.3. Comparison between AR and KR [I]

The results presented above may give the impression that the AR voiced [I] contains more frication noise than that in KR. The cross-language comparison (Fig. 8) indicates that although AR [I] has a slightly higher RZCR, there is no significant difference between the two languages. Fig. 9 arguably shows that the two languages' /I/s do not differ in terms of RZCR.

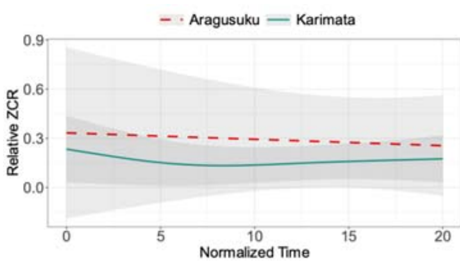

Figure 8. GAMM fitted values of Relative Zero-crossing Rate during /I/ in Aragusuku and Karimata.

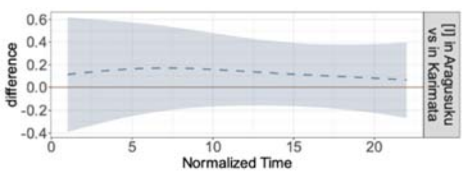

Figure 9. The difference in RZCR between the /I/s in Aragusuku and Karimata. The overall difference is not significant.

5. DISCUSSION AND CONCLUSION

This analysis is the first study conducted on the frication of the FV /I/ in two endangered languages, AR and KR. The results showed that there is frication noise on /I/ in the 5 kHz to 7 kHz region. This frication noise was observed when it was preceded by voiceless sibilants, vowels and voiced affricates. These observations show that the frication noise of /I/ could not be interpreted as being introduced by preceding sibilant consonants. When /I/ was devoiced ([I̥]) (i.e., between two voiceless consonants), the frication noise is more intense.

In AR, [I I:] contains significantly more frication than the regular vowel [i] at least in the first half of its duration. This pattern matches what was reported for apical vowels in Chinese languages [11, 25]. The pattern could be arguably explained by the aerodynamic adjustment [18]: the voicing and the frication are difficult to achieve simultaneously, and a trade-off has to happen during the fricative vowel, making the first half more fricative-like and the second half more vowel/approximant-like. In KR, [I I:] was not significantly more fricated than the regular vowel [i] as measured in ZCR. However, we note again that the limited number of tokens and speakers invites a careful interpretation. Recall that the AR speaker was a male and KR was a female. Female speakers might have a different physical structure and different socio-linguistic tendencies from male speakers influencing the results to some extent.

The interaction between voicing, frication, and duration among [I I: I̥] is complex. The long [I:] contains generally less frication compared to [I], and the devoiced [I̥] can be considered a fricative. Without knowledge of the articulatory characteristics of [I I:], it is difficult to determine how the frication noise is generated and superimposed on voicing. Similarly, understanding the acoustic shape of [I̥] requires an understanding of its supra-laryngeal articulation.

Finally, there was no significant difference in RZCR between [I] in the two languages, which might suggest that the difference in ZCR in each language was not solely caused by the frication noise present in [I I:], but also by the acoustic shape of [i].

Our study revealed that AR and KR have a strikingly similar type of fricative vowel, despite belonging to two different Ryukyuan languages, Yaeyama and Miyako. Historically, Ishigaki, the main island of Yaeyama (see Fig. 1) had contact with Okinawa, where the capital of the Ryukyu Kingdom was located in the past while Aragusuku islets were left out from the influence of the mainstream language changes [12]. However, historical language contact might not be the only reason for the similarity between AR and KR. Phonological environments may contribute to the preservation of the frication noise of /I/. /I/ in some of the Miyako Ryukyuan dialects (e.g., Hirara, Karimata, Bora) are reported to present frication noise in wider consonantal environments than in others (e.g., Ikema, Ogami) [16, 20]. And in the latter languages, the frication noise is reported declining. It would be desirable to conduct an articulatory study (e.g., ultrasound tongue imaging) across dialects of Miyako Ryukyuan along with acoustic and also phonological analyses to investigate the dialectal differences and the articulatory nature of the fricative vowel, which would contribute to a further understanding of the FV as a group of atypical sounds.

6. ACKNOWLEDGEMENTS

We would like to express our gratitude to the National Institute of Informatics' Speech Resources Consortium for providing us with access to the Aragusuku Dialect (Aragusuku) and Oogami Dialect (Oogami) speech databases. The study was supported by Labex EFL.

7. REFERENCES

[1] Boersma, P., Weenink, D. 2022. Praat: doing phonetics by computer [Computer program]. Version 6.1.16.
[2] Dell, F. 1994. Consonnes à prolongement syllabique en Chine. *Cahiers de linguistique – Asie orientale* 23(1), 87–94.
[3] Faytak, M., Shao, B. 2022. Fricative vowels are noisier in crowded vowel spaces. *LabPhon18*. Online.
[4] Faytak, M. 2018. *Articulatory uniformity through articulatory reuse: insights from an ultrasound study of Sūzhōu Chinese*, PhD dissertation. University of California, Berkeley.
[5] Faytak, M., Lin., S. 2015. Articulatory variability and fricative noise in apical vowels. *Proc. 18th ICPhS*, Glasgow.
[6] Hirayama, T (ed.). 1983. *Ryukyu Miyako shotou hougen kiso goi no sougouteki kenkyuu*. Tokyo: Ouhuu.
[7] Kajiku, S. 2016. Minami Ryukyu hougen gaisetsu, In: *Ryukyu no hougen, 40*, 107–200.
[8] Karimata, S. 1996. 'Miyako Hōgen no On'in Henkani Tsuite no Oboegaki: Kūki–Rikigakutekina Kanten kara Mite'. *Gengogakurin*, 1996–1997:709–22.
[9] Karlgren, B. 1915. *Études sur la phonologie chinoise*. Uppsala: K. W. Appelberg.
[10] Kinuhata, T., Hayashi, Y. 2014. Ryukyugo Miyako Karimata hougen no onin to bunpou, *Ryukyu no hougen*, 38, 17–49.
[11] Kong, H., Wu. S., Li, M. 2019. Héféi Huà Shéjiān Yuányīn de Mócā Xìngzhi jí Gǎnzhī Zēngqíang Lǐlùn Jiědú [The Frication Property of Apcial Vowels in Hefei Chinese and Its Perceptual Enhancement Account]. *Yuyan Yanjiu* 39(1), 23–33.
[12] Kuno, M., Kuno, M., Oono, M., Sugimura, T. 1993. *Zan o jounoushita hitotachi-Minami Ryukyu Aragusuku hougen no onsei to Jiraba*, Kokugakuin University, Institute for Japanese Culture and Classics.
[13] Ladefoged, P., Maddieson, I. 1996. *The sounds of the world's languages*. Oxford & Malden, MA: Blackwell.
[14] Lee-Kim, S-I. 2014. Revisiting Mandarin apical vowels: An articulatory and acoustic study. *J. Int. Phon. Assoc.* 44(3), 261–282.
[15] Moseley, C. 2010. *Atlas of the World's Languages in Danger* (3rd ed.). Paris: UNESCO Publishing.
[16] Nakahara, J, 2001. Okinawa Miyakojima Bora hougen no onin, *Ryukyu no hougen* 21, 105-123, Hosei University.
[17] NII-SRC (Speech Resources Consortium) The National Institute of Informatics (Last accessed 27 July 2022: http://research.nii.ac.jp/src/index.html).
[18] Ohala, J. J. 1983. The origin of sound pattern in vocal tract constraints. In MacNeilage, P. F. ed., The Production of Speech,189–216. Springer, New York.
[19] Oono, M., Kuno, M., Kuno, M., Sugimura, T. 1998. *Miyako Ogamijima no onsei – tango to bunpou (Appendix Karimata hogen)*, 1996-7 Grants in aid for Scientific Research Report.
[20] Oono, M., Kuno, M., Sugimura, T., Kuno, M. 2000. Phonetic substance of neutral vowels in the South Ryukyu Dialects, *Onsei Kenkyū* 4:1, 28–35.
[21] R Core Team. 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
[22] Sakiyama, O. 1963. The Comparative Phonology of Miyako Dialects in the Ryūkyūs, *Kokugogaku*, 54, 6–21.
[23] Shao, B. 2020. *The apical vowel in Jixi-Hui Chinese: phonology and phonetics*, PhD dissertation. Université Sorbonne Nouvelle.
[24] Shao, B., Ridouane, R. 2021. 'Apical vowels' are not vowels: acoustic and ultrasound evidence from Jixi-Hui Chinese. *Proc. of the 12th International Seminar on Speech Production*, 110–113.
[25] Shao, B., Ridouane, R. accepted. On the nature of apical vowel in Jixi-Hui Chinese: Acoustic and articulatory data. *J. J. Int. Phon. Assoc.*
[26] Soskuthy, M. 2021. Evaluating generalised additive mixed modelling strategies for dynamic speech analysis. *J. Phon.*, *84*, 101017.
[27] Wieling, M. 2018. Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *J. Pho.*, *70*, 86–116.
[28] Wood, S. 2006. *Generalized additive models: an introduction with R*. CRC Press.
[29] Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, 73(1), 3–36.
[30] Wickham H, Averick M, Bryan J, Chang W, McGowan L. D, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. "Welcome to the tidyverse." *Journal of Open Source Software*, 4(43), 1686.
[31] Coretta, S. 2022. tidymv: Tidy Model Visualisation for Generalised Additive Models, r package, version 3.3.2.