Preliminary analysis of /r/ acoustics and features in three Māori speakers

Isabella Shields[1], Catherine Watson[1], Peter Keegan[1]

[1]University of Auckland, New Zealand

ishi836@aucklanduni.ac.nz, c.watson@auckland.ac.nz, p.keegan@auckland.ac.nz

Abstract

This paper presents preliminary analyses of Māori /r/ in the speech of three fluent speakers. Māori /r/ is not yet fully understood and can be a somewhat misunderstood sound for Māori learners. We consider /r/ in several vowel environments and word- and phrase-stress contexts. Duration of /r/ was found to vary across speakers and phrasal stress. Most /r/ tokens were taps, although variation was found in formant behaviour and presence of frication and release bursts. Formant assessment of /VrV/ sequences indicates different articulations for /r/ are possible. We emphasise the need for more speakers/participants to characterise the sound.

Index Terms: Māori language, acoustic phonetics, rhotics

1. Introduction and Background

This study presents preliminary findings of a speech study investigating the acoustic features of the /r/ phoneme in Māori. This sound is as yet not fully characterised nor understood; while Māori /r/ has been described, it has not been the subject of a comprehensive acoustic investigation. This paper outlines the methodology of an ongoing investigation of Māori /r/ and presents results based a small set of speakers. The overall aim of the study is to identify the acoustic characteristics of Māori /r/ and assess the impact that factors such as vowel environment and stress environment have on the sound.

Māori is a Polynesian language indigenous to Aotearoa (a popular Māori name for New Zealand). While Māori and New Zealand Sign Language are the only official languages of Aotearoa, English is the most commonly spoken language. As a result of the colonisation of Aotearoa in the 1800s, there has been significant language contact between Māori and English. This language contact, accompanied by active discouragement of Māori language use, resulted in a significant decline in the number of Māori speakers that eventually resulted in a break in inter-generational transmission of the language. There has been notable success in revitalising Māori, with efforts spearheaded by Māori language communities [1]. The present research aims to contribute to our understanding of the Māori language and its revitalisation, and as our research involves Māori researchers and Māori data, it must adhere to principles of Māori data sovereignty [2, 3].

1.1. Māori phonology

Māori has five monophthongs: /i e a o u/. These vowels are high-front, mid-front, open, mid-back, and high-back, respectively [1]. There are 10 consonants (/ p t k m n ŋ f h w r /), although some variation across dialects exists. The five vowels have phonemically distinctive long and short quantities, although this distinction is weakening [4]. Long vowels are usually denoted with a macron in modern written Māori. Diphthongs can be formed by combinations of these monophthongs. Māori is considered by some to be a mora-timed language, with a mora consisting of a short vowel and optional preceding consonant (μ = (C)V; C = consonant, V = short vowel) [5]. Biggs' stress rules put forward explanations for placement of both word and phrase stress [6]. For monomorphemic words, stress placement is dictated by a syllable hierarchy, with the highest ranked syllable bearing the stress. This hierarchy, as outlined by Bauer, is: (C)V_1V_1 >(C)V_1V_2 >(C)V_1. Biggs' rules suggest that, in a phrase that is sentence-final, phrase stress should fall on the primary word stress of the final content word. Otherwise, phrase stress should fall on the phrase's penultimate mora. For example, when the phrase [*ki te kura nei*] (meaning *to this school*) is sentence final, the phrase stress (underlined) overlaps with the stressed mora of the content word (in bold) ([*ki te* ***<u>ku</u>****ra nei*], otherwise it falls on the penultimate mora ([*ki te* ***ku****ra* *<u>nei</u>*].

1.2. Māori /r/: existing descriptions and investigations

The focus of the present study, /r/, is most often described as an alveolar flap (e.g. in [7]). However, there is some inconsistency in its description. Trilled /r/ has been reported in some older descriptions, and approximant /r/ has also been reported by Harlow [8, 1]. He notes that outside of rapid speech and repeated sequences with intervening unstressed vowels, approximant /r/ is a direct result of English pronunciation influencing that of Māori. Harlow and Biggs both refer to lateral realisations of the phoneme, but with emphasis on different geographical locations: Harlow points to South Island dialects, and Biggs to Eastern dialects [1, 8]. The spectrographic characteristics and duration of /r/ were analysed in the speech of male *kaumātua* (elders) from the MAONZE corpus [9, 10]. The findings of this study supported the canonical designation of Māori /r/ as a flap, as well as Harlow's description of the contexts in which approximant /r/ appear. The Ngā Mahi recordings, a collection of recordings of one speaker produced for the development of Māori text-to-speech system (see [11]), were also investigated in the context of lexical stress and higher formant behaviours that could not be analysed in the MAONZE corpus as a result of audio quality limitations [10, 12]. A phonetic correlate of stress reported by Bauer is 'emphatic onset', by which /r/ should increase in duration when stressed [7]. In greater proximity to a stressed syllable, /r/ duration and intensity were found to increase and decrease, respectively. For the same speaker, notable lowering of the fourth formant (F4) in a range of segmental contexts was identified as a salient feature.

2. Methodology

2.1. *Kaikōrero* (speakers)

To date, speakers have been recruited using a *kaupapa Māori* (Māori principled) approach [13]. However, the recruitment process for the present study has been slow-moving given

COVID-19 restrictions and considerations. We observe that many Māori speakers can still feel a sense of *whakamā* (shame or embarrassment) when sharing it with others and feel reluctant to do so.

There are three *kaikōrero* (speakers) analysed in the present study: Speakers 1 and 2 are a 57 year old man and a 53 year woman, while Speaker 3 is a 17 year old woman. These speakers are fluent speakers of Māori, and make up a *whānau* (family unit). They use Māori in their daily lives, both at home and at work or school. All speakers are also fluent speakers of Standard New Zealand English, with Speaker 1 additionally reporting basic competency in Malay, the canonical /r/ of which is a trill or tap, although fricatives also appear [14]. Speakers 1 and 3 reported speaking Māori with their friends. Speakers 2 and 3 both indicated they learned Māori as children (at home and on the *marae* (tribal meeting place); and at home and in immersion schooling, respectively), while Speaker 1 learned the language post-adolescence.

2.2. Materials

Five word forms were selected which place /r/ in different positions relative to word stress (WS) as predicted by Biggs' stress rules. Of these, we report on three in the present study, with syllable boundaries denoted by /./: /ˈCV.rV/ (WS1; /r/ is in the unstressed onset), /ˈrV.CV/ (WS2; /r/ is in the stressed onset and word-initial), and /ˈCV_{long}.CV.rV/ (WS3; /r/ is unstressed and placed further from the stressed syllable). Nasal consonants were avoided in the target words to avoid potential nasalisation effects on surrounding sounds. Six immediate vowel contexts for /r/ are considered: /iri, ira, iro, ara, ari, aro/. These vowel environments were selected to cover a range of environments and tongue movements, and were based on Māori point vowels. While the vowel /u/ could have been included instead of /o/, it has become increasingly fronted [15]. The target words used are shown in Table 1, with the stressed mora in bold.

Table 1: *Target words used in the elicitation study.*

WS	/iri/	/ira/	/iro/	/ara/	/ari/	/aro/
1	**pi**ri	**hi**ra	**pi**ro	**pa**ra	**pa**ri	**pa**ro
2	**ri**po	**ra**pi	**ro**pi	**ra**pu	**ri**pa	**ro**tu
3	**tā**piri	**pā**kira	**kō**piro	**hō**para	**tō**kari	**tā**karo

In order to account for any potential influences of phrase stress, two different environments for the target word were selected. Carrier sentences are used to deliver the target words, as recommended in [16, 17]. In line with [17], the target word is shielded from the end of a phrase or sentence to avoid any boundary effects. The two carrier sentences used were:

- Sentence (1): [*Ka/I* {target} *tonu*][*a* {Name}] - E.g. *Ka **tā**karo tonu a Pita* (*Pita will still play*).
- Sentence (2): [*I kite*][*a* {Name}][*i ngā* {target} *nā*] - E.g. *I kite a Tia i te **tā**karo nā* (*Tia saw that game*).

The target word functions as a verb and as a noun in Sentences (1) and (2), respectively. In Sentence (2), Biggs' rules suggest the phrase stress falls on the primary word stress of the target word. In the case of Sentence (1), Biggs' rules suggest phrase stress does not fall on the target word, but rather on the penultimate mora of the phrase. Each sentence was randomly assigned one of six two-syllable Māori first names. In order to account for any temporal variations in speakers' production, the sentence list was repeated five times and its order randomised during each repetition. To add variation, a dummy sentence was added to the sentence list every 10 sentences.

It was not possible to formulate sentences which maintained the initial /i/ of /iri, ira, iro/ in WS2 words in Sentence (2) while also maintaining syntactically meaningful sentences. Instead, the plural article *ngā* (*the*; /ŋaː/) was replaced with the singular *te* (*the*, /te/), producing the sequences /eri, era, ero/. While not directly equivalent, these sequences were the closest approximation possible.

2.3. Recording process and speech task

Recordings for this study were approved by the University of Auckland Human Participants Ethics Committee (UAHPEC23198). Speech recordings were made using a Rode Lavalier lapel microphone and Roland OCTA-CAPTURE for preamplification and digitisation. Audio was captured at a sample rate of 44.1kHz and 16-bit bit depth. A maximum of five repetitions of the sentence lists were recorded, depending on the speed of the speaker. Digital capture and management of the display of recording materials was achieved using Python (version 3.7.2) and the *sounddevice* package (version 0.4.4) [18]. Recordings took place in a WhisperRoom Sound Isolation Enclosure (`https://whisperroom.com`), with speakers undertaking the task seated at a desk in front of a computer monitor. As a warm up task, speakers read aloud two passages of text in Māori shown on the monitor. This was followed by the central speech task, where sentences were displayed on the monitor and recorded one at a time. Recording sessions lasted a maximum of one hour. Speakers 1 and 3 completed the entire speech task and Speaker 2 completed four of the five repetition rounds within the hour.

2.4. Data preparation

Data processing, analysis, and visualisation was undertaken using R (version 4.2.0) [19]. Recordings were converted into speech databases using the package *EmuR* (version 2.3.0) [20]. Initial phonetic segmentation of recordings was achieved using WebMAUS General (language: language independent (SAMPA), otherwise default settings) [21, 22, 23]. Boundaries were then hand-corrected where necessary in *EmuR*. Start and end boundaries for /r/ were placed based on reduced amplitude of the waveform envelope and reduction in formant energies in the spectrogram and were placed at the nearest zero-crossing. Formant trajectories were estimated using the *forest* function in the *wrassp* package using default settings, and then hand-adjusted by the first author [24]. When a formant was weak enough to not be visible in a spectrogram, the formant value was excluded. This was particularly common for the fourth formant (F4), especially in the case of Speakers 2 and 3. All formant energies were also frequently not visible throughout the /r/.

2.5. Feature labelling

Labels for spectrographic/waveform features were added based on visual inspection of each /r/ token. These described the presence of a release burst ('B'), the presence of at least one formant between F1 and F4 throughout the /r/ ('R'), and the presence of frication in at least part of the /r/ ('F'). All cases of frication present were identified as a secondary articulation with the same primary manner of articulation (flap). In the rare cases where a different primary manner of articulation was identified (such as

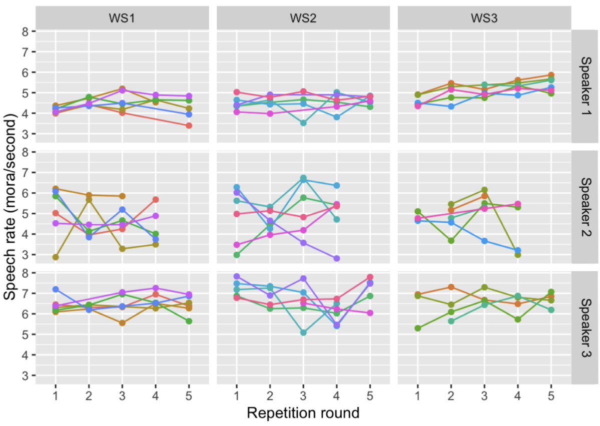

Figure 1: *Changes in speech rate (mora/s) for Speakers 1-3 in Sentence (1). Each colour corresponds to a different sentence.*

an approximant or trill), the tokens were labelled accordingly.

3. Results

In total 475 /r/ tokens were analysed. A total of 22 tokens were excluded from analysis, with exclusions made when the speaker made an error, the production was breathy, or where noise from movement interfered with the recording quality.

3.1. Speech rate

We considered speech rate on a speaker-by-speaker basis to determine if there was variation throughout the speech task speech Speech rate was calculated in terms of morae per second, with the number of morae in each sentence determined manually and divided by the duration of each utterance as extracted from the corrected EmuR database. The speakers did not show a noticeable tendency to increase/decrease their speech rate during the speech task, save for in Speaker 1's productions of Sentence (1)/WS3 sentences, which showed a slowly increasing speech rate as the task progressed. Speech rate for each repetition of each Sentence (1) sentence is shown in Figure 1. These results are comparable to Sentence (2) results. Visual inspection of Figure 1 shows Speaker 2, with a range of 3.9 morae, demonstrated more variation in speech rate across the different sentences and repetitions than Speakers 1 and 3 who had a speech rate range of 1.7 morae and 2.7 morae, respectively. This may indicate unease with some aspect of the experiment. Speaker 2 did report that she felt Sentence (2) sentences were odd to speak aloud due to the final word (*nā*, a locative particle). Speaker 3 tended to produce the sentences at a higher speech rate. This may be a result of comfort with the speech task or age, with younger speakers reported to have higher speech rate [25].

3.2. /r/ type and spectrographic features

All instances of /r/ observed were taps/flaps, save for three examples of an approximant and one trill. We conclude that the approximant and trill observations were mostly likely speaker error. Furthermore, no lateral productions of /r/ were observed. A majority of tokens (91.8%) had at least one formant (F1-F4) visible throughout. Frication was identified in 45.3% of observations, and release bursts in 12.6% of observations. On a speaker-by-speaker basis, the occurrence of these features was largely similar. Speaker 2 produced a release burst in 25.4% of their /r/ productions, compared to 6.5% and 8.8% for Speakers 1 and 3, respectively. Instances of /r/ with frication were observed in around half of Speaker 1 and 2 tokens (50.0%, and 49.3%, respectively), but in 37.4% of Speaker 3 tokens. We also note that tokens with frication appeared more frequently in the vicinity of a front vowel than not: frication appeared in 58.3% of tokens preceded/followed by a front vowel, compared to 19.5% of tokens in solely open/back environments.

3.3. Duration of /r/

Table 2 summarises the mean and standard deviation duration of /r/ tokens for word stress and sentence environment. Overall, mean /r/ duration was observed to be greater in the Sentence (2) environment than in Sentence (1), albeit to differing degrees. In Sentence (1) and (2), both Speaker 1 and Speaker 2 produced /r/ with greater mean duration in the WS2 context than in WS1 or WS3. These observations appear to support the observations made by Bauer regarding increased /r/ duration when in the onset of the stressed syllable [7]. Speaker 3 did not follow this trend, with the WS3 context (where /r/ is not in the onset of the stressed syllable) having the greatest mean /r/ duration in both Sentence (1) and (2). We also note there was a tendency for /r/ duration to increase when in the vicinity of a front vowel (such as in the sequence /iri/ or /ira/), but not in the sequence /iro/. As indicated above, /r/ tokens with frication occurred more frequently in these sequences than in others.

Table 2: *Mean duration and standard deviation of /r/ (ms) by stress environments (n>21 samples in each category).*

Mean /r/ duration and S.D. (ms) in Sentence (1)			
Speaker	WS1	WS2	WS3
1	29.0 (6.9)	35.3 (8.8)	29.3 (6.5)
2	24.3 (7.2)	35.2 (11.8)	28.8 (7.8)
3	26.5 (10.4)	26.7 (8.0)	30.2 (8.9)
Mean /r/ duration and S.D. (ms) in Sentence (2)			
1	30.0 (4.1)	36.7 (7.4)	33.3 (5.2)
2	29.8 (7.3)	41.8 (16.5)	37.7 (10.0)
3	29.8 (11.8)	29.9 (11.8)	37.1 (12.3)

3.4. Formants

The formant trajectories of the vowels preceding and following /r/ in the target words were considered in each sentence and word-stress environment. For the three speakers analysed, there did not appear to be any notable visual discrepancies between the formant trajectories of the two phrase stress environments, save for greater variation in estimated formant values in the Sentence (2) environment. In Figure 2 we include formant trajectories of /ari/ ((a), (c), and (e)) and /iro/ ((b), (d), and (f)) sequences in the Sentence (1)/WS1 environment, however the present discussion is not limited to these sequences.

The movements of the first and second formants (F1 and F2) provided some insight into the movements of the tongue before and after /r/ was articulated, and indicated there are different articulatory approaches to the various /VrV/ sequences. Preparatory movement associated with /r/ appears to occur to differing degrees depending on the /VrV/ sequence. For example, in the sequence /ari/ (pictured in Figure 2 (a), (c) and (e)), Speaker 2 and 3 show a steady raising of F2 accompanying the forward movement of the tongue towards the /r/ target, followed

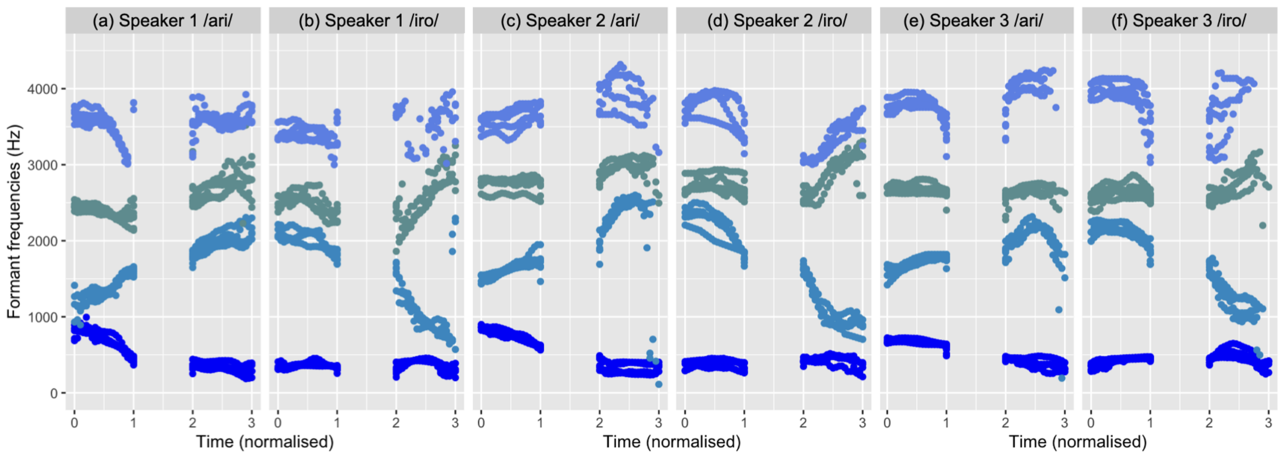

Figure 2: *Formant trajectories (F1-F4) of vowels preceding and following /r/ in /ari/ and /iro/ sequences in WS1/Sentence (1). Time 0-1 shows the formant trajectories of the preceding vowel, Time 1-2 corresponds to the /r/, and Time 2-3 to the following vowel.*

by more abrupt movement towards the /i/ target. Conversely, F2 for Speaker 1 appears to change at much the same rate both before and after /r/. Relatively little movement in F1 in observations of /iro/ demonstrated there was not much change in tongue height in the transition from /i/ to /r/, and /r/ to /o/. Conversely, there was notable lowering of F1 for Speaker 1 and 2 in the /a/ of /ari/, indicating gradual raising of the tongue to reach the alveolar ridge to produce /r/, before remaining at much the same height in preparation for articulating /i/.

Inference regarding articulation based on the third and fourth formants (F3 and F4) was more complicated. Similar to F1 and F2, preparatory changes leading into /r/ began at various points in the preceding vowel. Lowering of F3 is often associated with rhotic sounds [26]. For all speakers, we note /r/ accompanied by lowered F3 occurred in the presence of /o/. Māori /o/ is lip-rounded, contributing to the lowering of F3 (and the other formants). Outside of these contexts, F3 lowering was observed in various segmental environments for Speaker 1, less frequently for Speaker 2, and sporadically for Speaker 3. Notable movement of F4 preceding /r/ (lowering prior to /r/ and/or raising following /r/) occurred across all vowel environments for Speaker 1 and 2, albeit to differing degrees. For Speaker 3, visually salient lowering of F4 appeared to only occur in /iro/ (see Figure 2 (f)), /ero/, and /aro/ vowel environments. The source of this F4 lowering is not immediately evident, and may indicate a particular articulatory approach to /r/ in certain environments. This feature will be investigated further in future.

4. Discussion and Conclusions

The present study aimed to identify some of the acoustic characteristics of Māori /r/ and to investigate the potential impacts of segmental and stress contexts. To our knowledge, this study is the first of its kind formulating and analysing a speech task focusing on Māori /r/. Given the study is ongoing and we are presently reporting on a small number of speakers, we have focused on description rather than statistical analyses.

Inter- and intra-speaker variation is a common trait of rhotic sounds in several languages (such as Portuguese [27], Dutch [28], German [29], and Spanish [30]). In the present study, we find relatively little variation in /r/ type, but variation in other features like frication, presence of release bursts, duration, and formant behaviours. While we did not observe a consistent indication that word stress was influencing /r/ duration, it was generally shown to increase in Sentence (2) contexts. In Sentence (2), Bigg's stress rules suggest the phrase stress falls on the same mora as the word stress of the target word, placing more prominence on the mora compared to Sentence (1). It is possible that phrase stress is indeed influencing the duration of /r/, or that some combined interaction of phrase stress and word stress results in increased /r/ duration.

The majority of /r/ tokens (91.8%) had formant energy present throughout, indicating that the closure in the oral cavity produced during flap articulation is usually not sufficient to fully block airflow. This is not uncommon in other languages with flap sounds (e.g. American English dialect's flap allophone [31]). We report a rate of frication (45.3% of tokens) that is surprisingly high given it is not often mentioned in existing descriptions of Māori /r/. Given the majority of these observations occurred near a front vowel, we speculate there may be an interaction between /r/ and this environment by which the proximity of the articulatory targets of /i/ and /r/ results in slightly reduced contact of the tongue and alveolar ridge in the flap movement, allowing for a sustained flow of air with introduced turbulence.

In addition to further speaker recordings, an articulation-focused investigation would be useful to determine the exact source of frication alongside further acoustic investigation, and would also tease out complex articulatory information not easily inferred from higher formant behaviours (i.e. F3 and F4).

4.1. Conclusions

Detailed investigation of the acoustics of Māori /r/ is necessary if the sound is to be well understood. The preliminary results presented here represent the first steps in achieving this goal. Our findings contribute to the wider understanding of the characteristics of /r/; in particular, we identify variation in its duration in different stress and segmental environments, as well as indication of different articulatory approaches to producing the sound based on acoustic features and formant behaviours. We hope that with further investigation the source and breadth of these variations will become evident. The results of this preliminary investigation indicate there is a need to further home in on these various contexts with a wider speaker/participant group to enable a more precise description of /r/ in modern Māori.

5. References

[1] R. Harlow, *Māori: A Linguistic Introduction*, Cambridge, U.K.: Cambridge University Press. 2007.

[2] T. Kukutai and J. Taylor, Eds., *Indigenous Data Sovereignty: Toward an agenda*. Canberra, Australia: ANU Press, 2016.

[3] M. Hudson, T. Anderson, T. K. Dewes, P. Temara, H. Whaanga, and T. Roa, "'He Matapihi ki te Mana Raraunga' - Conceptualising big data through a Māori lens," in *He Whare Hangarau Māori: Language, culture & technology*, H. Whaanga, T. T. Keegan, and M. Apperley, Eds. Kirikiriroa/Hamilton: Te Whare Wānanga o Waikato/University of Waikato, 2017, pp. 64–73.

[4] J. King, R. Harlow, C. Watson, P. Keegan, and M. Maclagan, "Changing pronunciation of the Māori language: implications for revitalization," *Indig. Lang. Revital. Encourag. Guid. lessons Learn.*, pp. 85–96, 2009.

[5] W. Bauer, "Hae.re vs. ha.e.re: a note," *Te Reo*, vol. 24, pp. 31-36, 1981.

[6] B. Biggs, *Let's learn Maori: a guide to the study of the Maori language*, 3rd ed., Auckland, N.Z.: Auckland University Press, 1998.

[7] W. Bauer, *Maori*, London/New York, N.Y., U.K./U.S.A.: Routledge, 1993.

[8] B. Biggs, "The Structure of New Zealand Maaori," *Anthropological Linguistics*, vol. 3, no. 3, pp. 1-54, Mar., 1961. [Online]. Available: https://www.jstor.org/stable/30022302

[9] J. King, Maclagan, M., Harlow, R., Keegan, P., Watson, C. 2011. The MAONZE Project: Changing uses of an indigenous language database. *Corpus Ling. and Ling. Theory*, vol. 7, no.1, pp. 37-57, Apr., 2011.

[10] I. Shields, C. Watson, P. Keegan, and M. Maclagan, "A preliminary investigation of the acoustics of Māori /r/," New Zealand Linguistics Society Conference, Hamilton, New Zealand, Dec., 2021. [Presentation]

[11] I. Shields, C. Watson, P. Keegan, R. Berriman and J.James "Creating a Synthetic Te Reo Māori Voice," International Conference on Language Technologies for All, Paris, Dec 2019. [Online]. Available: https://lt4all.org/media/papers/P1/136.pdf.

[12] I. Shields, C. Watson, P. Keegan, "Acoustics of te reo Māori /r/: First impressions and corpus building for fourth formant and stress analysis," 'r-atics 7 Conference, Lausanne, Switzerland, Nov., 2021. [Presentation].

[13] L. T. Smith, *Decolonizing Methodologies: Research and Indigenous Peoples*, 3rd ed., London, U.K.: Zed Books. [Online]. Available: https://bloomsbury.com/9781786998132.

[14] D. Deterding, I. A. Gardiner, and N. Noorashid, *The Phonetics of Malay*, Cambridge, U.K.: Cambridge University Press, 2022.

[15] M. Maclagan, C. I. Watson, R. Harlow, J. King, and P. Keegan. "/u/ fronting and /t/ aspiration in Māori and New Zealand English," *Language Variation and Change*, vol. 21, no. 2, pp. 175-192, Jul., 2009, doi: doi:10.1017/S095439450999007X.

[16] J. Harrington, *Phonetic Analysis of Speech Corpora*, Chichester, U.K.: Wiley-Blackwell, 2010.

[17] T. Roettger and M. Gordon, "Methodological issues in the study of word stress correlates," *Linguist. Vanguard*, vol. 3, no. 1, 2017, doi: 0.1515/lingvan-2017-0006.

[18] M. Geier. *sounddevice*. [Online]. Available: https://github.com/spatialaudio/python-sounddevice.

[19] R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

[20] R. Winkelmann, K. Jaensch, S. Cassidy and J. Harrington. *emuR: Main Package of the EMU Speech Database Management System*, R package version 2.3.0.

[21] F. Schiel. "Automatic Phonetic Transcription of Non-Prompted Speech," in Proc. of the ICPhS, San Francisco, C.A., U.S.A., 2015, pp. 607-610.

[22] F. Schiel. "A Statistical Model for Predicting Pronunciation," in Proc. of the ICPhS 2015, Glasgow, U.K., 2015, paper 195.

[23] T. Kisler, U. D. Reichel and F. Schiel, "Multilingual processing of speech via web services," *Computer Speech & Language*, vol. 45, no. 1, pp. 326–347, Sept., 2017.

[24] L. Bombien, R. Winkelmann and M. Scheffers. *wrassp: an R wrapper to the ASSP Library R package*, R package version 1.0.1.

[25] J. Bóna,"Temporal characteristics of speech: The effect of age and speech style," *J. Acoust. Soc. Am.*, vol. 136, no. 2, pp. EL116-EL121, Aug., 2014, doi: 10.1121/1.4885482.

[26] P. Ladefoged and I. Maddieson, *The sounds of the world's languages*, 1st ed. Oxford/Cambridge, M.A., U.K./U.S.A.: Blackwell, 1996.

[27] I. E. Rennicke, "Variation and Change in the Rhotics of Brazilian Portuguese," Ph.D. dissertation, Faculty of Arts, Federal University of Minas Gerais, Belo Horizonte, Brazil, 2015.

[28] K. Sebregts, "The Sociophonetics and Phonology of Dutch r," Ph.D. dissertation, University of Utrecht, Utrecht, Netherlands, 2015.

[29] R. Wiese, "The Unity and Variation of (german) /r/," *Zeitschrift für Dialektol. und Linguist.*, vol. 70, no. 1, pp. 25–43, Jul. 2003, [Online]. Available: http://www.jstor.org.ezproxy.auckland.ac.nz/stable/40504887.

[30] E. W. Willis, "An acoustic study of the 'pre-aspirated trill' in narrative Cibaeño Dominican Spanish," *J. Int. Phon. Assoc.*, vol. 37, no. 1, pp. 33–49, 2007, doi: 10.1017/S0025100306002799.

[31] N. Warner, A. Fountain, and B. V Tucker, "Cues to perception of reduced flaps," *J. Acoust. Soc. Am.*, vol. 125, no. 5, pp. 3317–3327, 2009, doi: 10.1121/1.3097773.