Phrases, pitch and perceived prominence in Māori

Laura Thompson[1], *Catherine I. Watson*[1], *Ray Harlow, Jeanette King*[2], *Margaret Maclagan*[2], *Helen Charters*[1], *Peter Keegan*[1]

[1] University of Auckland, New Zealand
[2] University of Canterbury, New Zealand

la.thompson@auckland.ac.nz, c.watson@auckland.ac.nz

Abstract

This study explores phrase-level prosody and prominence in the Māori language. Limited existing prosodic analysis and anecdotal evidence of diachronic change have motivated the present investigation into alignment of descriptions of intonation and stress with prominence perception test results and pitch analysis of speech data. In general, we find the expected case does occur most often, but examining results across speakers with birthdates spanning a century shows conservatism in modern elders and contradictory results in younger speakers: while making 'errors' in prominence placement, they are often as faithful to the overall expected contour as their elders.

Index Terms: speech perception, Māori, prominence, prosody, intonation, pitch, stress

10.21437/Interspeech.2011-450

1. Introduction

This paper presents part of an ongoing study into prosody and prominence in Māori, and change in the language over time. In this part of the study, we focus on pitch and perceived prominence within phrasal units, using utterances from speakers whose birthdates span more than 100 years. Pitch behaviour is compared with existing descriptions of intonation contours in Māori and with prominence locations identified in a perception exercise.

2. Background

2.1. Māori language

Māori is the indigenous language of New Zealand. It has been in contact with and influenced by English for about 200 years, causing both qualitative and quantitative changes [1]. Māori has 10 consonants: /p t k m n N f h w r/, and the five-vowel system has a two-way phonemic length distinction, but this is being lost in all pairs except /a - a:/ [2]. Māori syllables may take any of the forms permitted by (C)V(V(V)).

Older Māori speakers (or 'elders') are, anecdotally, sensitive to a change in what they call the *mita* of their language. The exact definition of this concept presents difficulties: sometimes described as the 'sound' or 'accent' of a language, *mita* seems to encompass a range of elements, from regional lexical or phonetic variation to stress, rhythm and pitch [3].

There is existing description of Māori prosody [4, 5, 6, 7], but no in-depth acoustic prosodic analysis. The first detailed analysis of both syntactic and prosodic organization in Māori was that of Bruce Biggs (1921-2000) [4], revised in [5]. Subsequent analyses by Winifred Bauer [6] and Paul de Lacy [7] have provided more detail in different frameworks, while following essentially the same pattern. This section will provide a brief introduction to the 'rules' of stress and intonation in Māori based on these three analyses.

2.1.1. The Māori phrase

Māori syntactic phrases and phonological phrases (of whatever type) are recognized as being inextricably linked [4], [6]. Māori syntactic phrase structure may be described in a number of ways, but the most useful here is that used by Biggs. The Māori phrase is a unit with up to three parts: an obligatory nucleus contains one or more content words, and optional peripheries either side of it hold function words [5]. An example is found in (1) below.

(1) Māori phrase structure, with example.

[(preposed periphery)		**nucleus**	**(postposed periphery)]**
ki	*te*	*whare*	*nei*
to/towards	the-SG	house	'close to speaker'
'to this house'			

2.1.2. Māori stress assignment

Māori stress rules refer to 'word stress' and 'phrase stress'. Word stress (WS) assignment is generally predictable from rules using both the mora, (C)V, and the syllable. Monomoraic syllables are light; all others are heavy. In monomorphemic words, WS is assigned according to a syllable weight hierarchy [6], where *heavy* > *light* and *long V* > *diphthong* > *short V*. The highest available syllable type within four morae from the end of a word bears WS (e.g. *koi**raa*** 'and so'). This is also true of some long vowels or diphthongs created across morpheme boundaries, and prefixes and reduplications have different rules. In de Lacy's analysis [7], which uses more recent terms referring to levels on a prosodic hierarchy [8], WS is assigned within a prosodic word.

The next domain of stress assignment is generally known as the 'phrase', and is said to be bounded by pauses. For Biggs and Bauer, this phonological phrase (PPh) has exactly the same shape as the syntactic phrase described above, though it is allowed that in rapid speech more than one syntactic phrase may be included in a PPh. It is not clear how often this is due to enclitic pronouns. The PPh as described by de Lacy does not include the preposed periphery, because he aligns the edges of PPhs with the edges of lexically-headed syntactic phrases only, but otherwise is the same as above. All phrases have one phrase stress (PS), and the location of this PS depends on whether the phrase is sentence-final or not. In final phrases, the PS is expected to be the WS of the last content word, *not* including enclitic pronouns [9]. In non-final phrases, PS is expected on the penultimate mora of the phrase, which may depart from the content word WS position.

2.1.3. Māori intonation

The description of intonation here will be restricted to the characteristics of simple declaratives and questions, since these are relevant for the materials analysed in the study.

Copyright © 2011 ISCA

Each (phonological) phrase in a declarative is expected to have a default 'rise-fall' contour, or, in de Lacy's analysis, a H*L⁻ tune [6], [7]. The 'peak' or high target should be reached over the 'most prominent' syllable in the phrase; this is expected to align with the PS. The low target should be reached at the end. Question intonation varies greatly: speakers may use a high rising terminal, a fall, a rise on the question word, higher pitch throughout, or retain declarative intonation and rely on lexical means to mark the question [6]. The present study is intended to investigate whether or not pitch behaviour and listeners' perceived prominence results bear any of this out.

According to Bauer, there is no apparent intonation 'peak' at the sentence (or, intonational phrase) level in Māori. There is sometimes a general pattern of descending pitch across the utterance, but this is not consistent enough to be called a rule. The right boundary of an intonational phrase is often marked by devoicing, though this may also occur elsewhere.

Biggs' analysis, from the mid C20th, is based on speakers from his parents' generation and before [4]. Bauer's and de Lacy's analyses, though more recent, are based on adult speakers and at least in part related to Biggs' analysis. If the sound of Māori has indeed changed, then it is logical to test these analyses against both historic and modern speech, and see how often the described patterns occur. This is made possible by recorded material from the MAONZE Database (see 3.1).

3. Methodology & Analysis

3.1. Material

The MAONZE (Māori and New Zealand English) Database [10] includes recordings of read and continuous speech in both Māori and English from speaker groups with birthdates spanning about 100 years. Here, recordings of Māori L1 males from three of these groups are used: 'historical' elders (HE, born ca.1880); present-day elders (PE, born ca.1920~1940), and young speakers (PY, born ca.1970~1990). To put this in context with the above, the HE group were 'elders' to Biggs; the PE group slightly younger; and he would have been considered an elder by the PY group.

Thirty-two recordings of Māori sentences were taken from interviews with several different speakers in the database: five HE, and four each of PE and PY. Not including the buffer sentences used at the beginning and end of the perception test, there are ten complete sentences from each of the three groups. Sentence selection criteria included clarity of recording, content, and length: most are 3-5 phrases long.

3.2. The prominence perception study

3.2.1. Delivery

Participants accessed a web interface, where they entered basic information (age, gender, language experience, etc.), and then listened to the sentences in the exercise. Each sentence was visible on the screen with normal punctuation, and syllabified underneath according to Māori syllabification rules (see e.g. [9]). Each syllable had a checkbox (Fig. 1). Participants played each recording, then checked the boxes of syllables that 'stood out' to them. The instruction wording was deliberate, to avoid influencing responses. Recordings could be played multiple times, but not paused or restarted partway through. This method has been used successfully in previous work [11], showing that participants will identify prominences consistently.

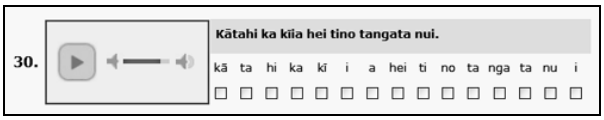

Figure 1: *Screenshot from the online survey interface.*

3.2.2. Participants

The ongoing prominence perception experiment currently has 68 participants obtained through networks. The results here represent a subset of 40. There were 30 females and 10 males, aged 16 to 65+. In perception studies of this type, 'native' speaker perception is crucial. However, NZ Māori presents two difficulties here. First, L1 speakers are hard to find. Second, all NZ Māori speakers, L1 or otherwise, are also exposed to (and often speak) NZ English on an extensive, daily basis, and their 'native' Māori cues must inevitably be influenced by that. Because of this, in this study we analyse all participants with experience parsing Māori – a natural result of NZ residency of any duration. Māori proficiency in this set was self-rated as between 'very good' (14/40) and 'words and phrases' (26/40). All participants had high English proficiency.

3.2.3. Perceived prominence analysis

For each syllable, a percentage score was calculated using the number of participant 'votes', and the scores were charted for each sentence. Perceived prominences were identified, the main criterion being a 'peak' relative to surrounding syllables; these peaks were naturally of differing sizes. Syntactic phrases within the sentences were marked on the charts according to the rules described above (2.1.1). The predicted word and phrase stress locations were also noted on the charts. Again, see Fig. 2 for an example. Vertical lines are phrase boundaries; capitals mark expected word stress; underlines mark expected phrase stress and circles mark prominences. The angled line illustrates the way 'peaks' were identified.

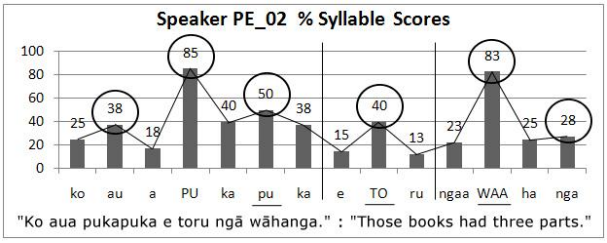

Figure 2: *Sentence analysis for a PE group speaker.*

Using the charts, the alignment of perceived prominences (pps) with the predicted stresses was assessed. The number of phrases in each category was divided by the total number of phrases to produce further percentage scores, both overall and separately for each of the three speaker groups. The results can be found in section 4.1.

3.3. The pitch study

While pitch is by no means the sole relevant acoustic cue to prominence in Māori or any other language, it plays a strong role in the descriptions given above. Biggs, Bauer and de Lacy, while mentioning other acoustic and phonetic features that may affect perception of prominence in Māori, all say that the highest *pitch* point or 'peak of intonation' should correlate with the 'most prominent' syllable in the phrase. As described in 2.1.3, this point is also expected to align with the PS syllable. With that, and constraints of space, in mind, this part of the investigation focuses on pitch, incorporating the results

from the perception study into the analysis. It is based around three questions:

(1) Is the H*L⁻ tune present in the phrases analysed here? (2) Is the H*, as visible on the pitch plot, found on the predicted PS syllable, and if not, where is it? (3) Is the H*, as visible on the pitch plot, found on the syllable considered to be the most prominent in the phrase by the listeners in the perception exercise (3.2.3), and if not, where is it? And further, how often do H*, expected phrase stress, and the greatest perceived prominence in the phrase coincide?

Pitch plots were created for each of the 30 utterances (buffer sentences excluded) using Emu (http://emu.sourceforge.net) and R (http://www.R-project.org). The phrases were marked on the plots, as on the charts in the perception study. So were perceived prominences, graded according to strength. Two assessments were made for each phrase: the shape of the contour and the location of H*, taking into account segmental effects such as pitch increases resulting from nasals and voiceless stops. An annotated plot for the same sentence as in Fig. 2 may be seen in Fig. 3. Thick black lines are between phrases; 'H*' marks the syllable with the highest pitch point; P1/2/3 mark perceived prominences by level; expected WS/PS locations are marked as before.

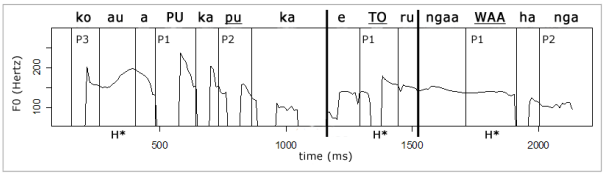

Figure 3: *Pitch plot for a PE group speaker.*

Three category-based analyses addressed the three questions above. In all cases, results are examined both overall and split over the three speaker categories (HE, PE, PY) in order to look for any indication of change over time. The details may be found in the following section (4.2-4.4).

4. Results & Discussion

This section presents the details and results of all four category analyses, along with one cross-analysis. It will also discuss some of the departures from expected behaviour and any possible evidence of change over time.

4.1. Alignment: perceived prominence

This analysis tests the alignment of perceived prominences (pp) with the expected word and phrase stress positions. The categories are: (a) 'match' (pps occurred exactly where expected), (b) 'match + extra' (all pps were as expected, but with one/more additional prominent syllable/s), or (c) 'mismatch' (pps did not match one/some/any expected positions; are either missing or in an unexpected position). The results are shown in Table 1.

Table 1: *Results from the alignment score analysis.*

Category	Overall	HE	PE	PY
Match	77 (52%)	21 (48%)	25 (55%)	31 (53%)
Match + Extra	20 (14%)	8 (18%)	7 (16%)	5 (9%)
Mismatch	50 (34%)	15 (34%)	13 (29%)	22 (38%)
Total phr.	**147**	**44**	**45**	**58**

Overall, more phrases have the expected alignment between predicted stress and perceived prominence than have some mismatch. When we look by speaker group, there appears to be a slight difference between the historical (HE) and the present-day (PE/PY) speakers, though it is not really possible to say that one group is obviously perceived as matching more often, or having more or less extra prominence than another. The PE speakers do not seem to be greatly different from the PY group. However, it is the PE group who are commenting on change in the language with regard to the younger speakers. The PE group also perform better in low-pass filtered English/Māori language discrimination tasks than young Māori speakers [12]. Perhaps they are particularly conservative, or aware of something else that is not immediately evident in the alignment results here.

4.2. Contour: H*L⁻ tune

This analysis assessed the presence or lack of the documented H*L⁻ tune in each phrase. Phrases were therefore categorized as being either (a) H*L⁻, or (b) No H*L⁻ (e.g., a rising tune).

Across the entire dataset, the expected H*L⁻ tune is prevalent (Table 2). Where it does not occur, this is chiefly due to a rise being present instead. The dataset includes only one question, whose final phrase does indeed show a rise, even though this is not strictly required in Māori. However, the rise is also present in several declarative phrases, in the form of the high rising terminal present in NZ English, Māori English, and increasingly in Māori itself [6]. This is particularly noticeable in the PY group.

Table 2: *Results from the contour analysis.*

Category	Overall	HE	PE	PY
H*L⁻	123 (84%)	37 (84%)	41 (91%)	45 (78%)
No H*L⁻	24 (16%)	7 (16%)	4 (9%)	13 (22%)
Total phr.	**147**	**44**	**45**	**58**

Across the three speaker groups, the default tune is also dominant. The elder speakers (HE/PE) appear to fit the documented pattern slightly more often than the younger speakers do, with the PE group showing the most conservative results.

4.3. H* Location: alignment revisited

Here, the location of the highest pitch point in each phrase was marked, and it was established whether or not this point, labelled H*, aligned with the expected PS position as predicted by the descriptions. The categories are (a) PS position, (b) WS position (if the two are different), or (c) other syllable position. The results are presented in Table 3. Overall, the H* does align more often with the PS position than not.

Table 3: *Results from the H* location analysis.*

Category	Overall	HE	PE	PY
PS Syll.	81 (55%)	20 (45%)	29 (64%)	33 (57%)
WS Syll.	13 (9%)	2 (5%)	7 (16%)	4 (7%)
Oth. Syll.	53 (36%)	22 (50%)	9 (20%)	21 (36%)
Total phr.	**147**	**44**	**45**	**58**

Where there is a departure from PS position, it is usually not to a WS syllable as one might expect, but to another syllable entirely. Among the different speaker groups, there are two points of note. First, the oldest (HE) speakers have the lowest

number of H*/PS alignments. Second, the PE speakers have the highest number of H*/PS alignments, and are more evenly split between WS and other syllables for non-PS alignment than are the other two groups. This may suggest greater awareness of expected stress locations in the PE speakers, targeting WS over PS when they do not coincide, or simply show focus on a different level.

4.4. H*: perceived prominence match

This analysis looked at whether the syllable bearing H* was also the syllable considered most prominent by the listeners in the perception task. The categories are (a) P1 (match with biggest pp in phrase), (b) P2 (match with second biggest pp in the phrase), or (c) none (H* matches no pp / no pp to match).

Table 4: *Results from the H*/pp match analysis.*

Category	Overall	HE	PE	PY
P1	75 (51%)	18 (41%)	25 (55%)	32 (55%)
P2	11 (7%)	3 (7%)	3 (7%)	5 (9%)
None	61 (42%)	23 (52%)	17 (38%)	21 (36%)
Total phr.	**147**	**44**	**45**	**58**

Overall, H* matched P1 almost as often as it did not. However, when it did not, in the vast majority of cases it matched a non-prominent syllable rather than P2, suggesting the listeners were attuned to some other cue than pitch when identifying those prominences. Across the three groups, the HE speakers had the lowest match of H*/P1. In contrast, the results for the PE and PY groups were remarkably similar.

4.5. 'Canonical' contour?

This section crosses three analyses (contour, H* location and H*/pp match), in order to see just how common the 'canonical' phrase intonation, H^*L^- with highest pitch point on PS and most prominent syllable, is in these data.

Table 5: *Occurrence of 'canonical' intonation.*

Category	Overall	HE	PE	PY
H^*L^- + PS + P1	53 (36%)	14 (32%)	17 (38%)	22 (38%)
Other	94 (64%)	30 (68%)	28 (62%)	36 (62%)
Total phr.	**147**	**44**	**45**	**58**

The results (Table 5) show that overall, this contour accounts for about a third of the phrases analysed here, with similar figures reflected across the three speaker groups. This suggests, at least, that it is not lack of this particular intonation which is causing the comments about change from elders.

5. Conclusion & Directions

Overall, in all the analyses, the expected case from existing descriptions occurs more often than not. Across the different speaker groups, it is often difficult to discern particular patterns, though it is noticed that the HE group are not, as one might predict, always the most conservative. The PE speakers often show little difference from the PY speakers, despite commenting on 'change' in the language with reference to younger speech. Possibly, the change is in the other direction as well: between the HE group and the PE group. More likely is that there are other cues at work here. Finally, the lack of canonical contour does not appear to be a factor in the impression of change. The ongoing study will continue, incorporating wider investigation of cues that may be relevant in perception of prominence in Māori. The goal is to build up a thorough, multi-level description. Further analysis will involve comparing results from Māori and non-Māori familiar participants in the study, since native language and language familiarity are known to affect prominence perception. Early indications here are that the two groups are not especially or predictably different. More Māori participants will be required to assess this with any accuracy, but their high exposure to English complicates matters.

The notions of 'correctness' and 'rules' are almost unavoidable when comparing material with existing descriptions. Biggs [13] does underline the importance of teachers of Māori being at least *consistent* in pronunciation and intonation for their students to emulate. Whether or not they are 'correct' is part of the eternal prescriptive-progressive struggle in any language, and it is not for the present project to give either side ammunition. Initially, this study simply seeks to compare what appears to be occurring with existing description, and look for possible change over time. It is hoped that the information gained might be useful in discussions about teaching models, revitalization efforts and general language maintenance.

6. Acknowledgements

This research was supported by grants from the Marsden Fund of the Royal Society of NZ and the University of Canterbury, and a scholarship from the University of Auckland. Thanks go to the speakers and the participants for their time and effort, and the four reviewers for their valuable comments.

7. References

[1] C. Watson, M. Maclagan, R. Harlow, et al., "Ka conversion - the changing sound and rhythm of Māori?" in *Laboratory Phonology 11, Book of Abstracts*, 2008, pp. 157-158.

[2] M. Maclagan, et al., "New Zealand English influence on Māori pronunciation over time," *Te Reo*, vol. 47, pp. 7-27, 2004.

[3] M. Maclagan, et al., "Investigating changes in the rhythm of Māori over time," in *Proc. 10th Ann. Conf. International Speech Comm. Assoc.: Interspeech 2009*, Brighton, 2009, pp. 1531-1534.

[4] B. Biggs, "The structure of New Zealand Maaori," *Anthropological Linguistics*, vol. 3, no. 3, Mar. Indiana University: Archives of Languages of the World, 1961.

[5] B. Biggs, *Let's learn Māori: a guide to the study of the Māori language*. Wellington: A.H. & A.W. Reed, 1969.

[6] W. Bauer, *Maori*. London: Routledge, 1993.

[7] P. de Lacy, "Constraint universality and prosodic phrasing in Māori," in *Papers in Optimality Theory II,* UMOP 26, A. Carpenter, A. Coetzee and P. de Lacy, Eds. Amherst: GLSA, 2003, pp. 59-79.

[8] S. Shattuck-Hufnagel and A. Turk, "A prosody tutorial for investigators of auditory sentence processing," *J. Psycholing. Research,* vol. 25, no. 2, 1996, pp. 193-247.

[9] R. Harlow, *A Māori reference grammar*. Auckland: Pearson Education, 2001.

[10] J. King, M. Maclagan, R. Harlow, et al., "The MAONZE Project: changing uses of an indigenous language database," *Corpus Ling. and Ling. Theory*, vol. 7, 2011, pp. 37-57.

[11] L. Thompson, C.I.Watson, H. Charters, et al., "An experiment in *mita*-reading: investigating perception of rhythmic prominence in the Māori language," presented at SST 2010, Melbourne, Australia, 2010.

[12] M. Maclagan, C.I. Watson, J. King, et al., "Identification of Māori and English from minimal pitch/rhythm cues," presented at NWAV 39, San Antonio, Texas, 2010.

[13] B. Biggs, *Pause and stress in Māori speech*. University of Auckland Working Papers in Anthropology, Linguistics, no. 20, Auckland: University of Auckland, 1973.