Change in Māori focus/topic *ko*: the impact of language contact on prosody

Sasha Calhoun, Naoko Yui, and Karena Kelly

Victoria University of Wellington

sasha.calhoun@vuw.ac.nz, karena.kelly@vuw.ac.nz

Abstract

This paper looks at the impact of contact with English on the prosody of te reo Māori. We investigate the functions and prosody of sentences with pre-predicate focus/topic marker *ko* in three groups of present-day male speakers from the MAONZE corpus (King et al. 2010): older L1, and younger L1 and L2 speakers. Older speakers show the expected pattern of main stress in the *ko*-phrase for focus *ko*, and in the main clause for topic *ko*. L1Y speakers show a similar pattern, but with a weaker distinction. L2Y speakers primarily use topic *ko*, but apparently with "focus" prosody.

Index Terms: prosody, topic, focus, information structure, Māori, Austronesian, language change, language contact

1. Introduction

Language change is a natural phenomenon, occurring when languages meet each other [1]. It is believed that language contact can impact all levels of linguistic structure [1]; however, there have been relatively few studies looking at how language contact affects prosody (cf. [2]). This needs to be addressed, as prosody is an integral part of spoken language, signalling a wide range of meanings, including information structure [3].

Te reo Māori is the indigenous language of New Zealand, a Polynesian VSO language which came into contact with the English language in the early 19th century, leading to a period of dramatic language loss in the mid-20th century, and a subsequent language revitalisation movement from the 1980s to the present. National survey results in 2013 showed 50,000 people (11% of the adult Māori population) reported speaking Māori very well or well, around 14,500 of whom were aged 55+ years, and 18,000 were aged 25-44 (http://www.stats.govt.nz/browse_for_stats/people_and_communities/maori/TeKupenga_HOTP13/Tables.aspx).

There have been a few linguistic studies exploring the effect of contact with English on modern Māori. Kelly's [4] corpus-based investigation identified a number of aspects of syntactic change occurring between 1900 and 1990. The Māori and New Zealand English (MAONZE) project has investigated sound change in Māori, showing ongoing influence from English, as well as system-internal changes, using recordings of 60 male and female Māori speakers from three different generations with birth dates spanning over 100 years [5]. Thompson [6] looked at the perception of prosodic prominence among contemporary Māori and non-Māori speakers, using recordings from the MAONZE corpus.

The information structure of traditional Māori matches other Polynesian languages, and what has been proposed to be a rule for verb-initial languages, i.e. the focus is usually sentence initial, optionally preceded by a slot for establishing topics [7,8,9]. Foci present new information in relation to the current proposition and/or carry the discourse forward, whereas topics are the entities being discussed [10,11]. There are a number of constructions which place a focused constituent in initial position, including pre-predicate *ko*-phrases. The initial focus carries the main stress, which cannot generally be moved in the sentence to mark emphasis [7], ie. less intonational plasticity. By contrast, in English the focus tends to be final, and stress can be moved freely within a sentence to mark focus [3]. Anecdotal evidence suggests younger Māori speakers are more influenced by English means of marking information structure ([7]:218, [12]:664).

In this study, we look at the prosody of the focus marker *ko* (all examples from MAONZE):

(1) Ko tō mana kei roto kē i te reo
FOC your power LOC inside INTENS PREP the language
'Your power lies within the language.'

Here the speaker has just been talking about his language, so 'your mana' is new in the proposition, and *ko* marks focus.

However, *ko* can also be used to mark topics:

(2) Ko rātou kei te rongo mai i te haunga ika
TOP they TAM sense DIR DO the odour fish
'They are smelling the fishy odour.'

'They' were just mentioned in the previous discourse, so *ko* marks the topic. *Ko*-phrases can also appear after the predicate, but here only topic *ko* is attested ([12]:181). *Ko* has a number of other functions, eg. future locative preposition; these are not considered in this study.

It is claimed that in pre-predicate position, topic and focus *ko* are distinguishable prosodically, with the main stress falling within the *ko*-phrase for the focus usage, and outside the *ko*-phrase for the topic usage [10]. In Māori each phonological phrase is said to be marked with an H*L- pitch contour, with the H* associated with the lexically-stressed syllable of the head word [13]. Bauer ([12]:562) comments that the pitch peak in the first phonological phrase in a sentence can be raised to mark emphasis, however, the subsequent phrases are not dephrased or deaccented. We therefore expect the main stress to be marked by relative prominence, not accent placement or type.

While there have been anecdotal observations that the use of focus/topic *ko* is declining among younger Māori speakers, there has not been any research on this. Further, the prosodic distinction between topic v focus *ko* has not yet been examined using modern prosodic analysis, nor is it known if this distinction is maintained by younger speakers. This study sought to investigate these issues.

2. Method

2.1. Data

The data for this study was taken from the MAONZE corpus [5]. Two of the speaker groups have been analysed to date: present day male elders (kaumātua, Group K) and present day young males (Group Y) (recorded mid 2000s) (see summary Table 1). This intergenerational comparison provided a means to investigate language change. The interview data were used.

Group K speakers were raised speaking Māori in a rural Māori-speaking community, learning English after starting school ([5]:8). Group Y was divided into two sub-groups based on language experience: the L1Y sub-group were native speakers of Māori, raised in a Māori-speaking environment, and the L2Y sub-group were second-language speakers, who had learned Māori through schooling ([14]:318).

Table 1: *Summary of* ko *data*

	Speaker Group		
	K	**L1Y**	**L2Y**
Speakers	9	5	4
Year of birth	1925-38	1969-84	1969-84
Age at recording	64-79	21-35	21-35
Total *Ko* occurences	798	420	456
FOC/TOP *ko*	121	65	59
Pre-pred FOC/TOP *ko*	83	51	34

LaBB-CAT [15] was used to extract all sentences with *ko*-phrases (see Table 1). There were initially 10 K and 10 Y speakers, but one K and one L2Y speaker's recordings were excluded due to poor sound quality. The transcriptions for each recording were then exported as Praat Textgrids [16], with automatic word and phone alignments. Sentences with focus/topic *ko* were identified, glossed and translated into English. Surrounding discourse context was also translated to enable classification of each *ko* as marking topic or focus from the translation. Here we only consider pre-predicate *ko*.

2.2. Labelling topic and focus

The type and subtype of each focus/topic *ko* was then annotated following the guidelines in Skopeteas et al. [17]. All annotations were checked and agreed by all authors. Topic and focus were defined in section 1. Two subtypes of topic were identified: *aboutness topic* (TOPa), what the sentence is about (see (2)); and *contrastive topic* (TOPc):

(3) Ko tētahi i noho mai i roto o Waipū
TOP one TAM stay DIR LOC inside of Waipū
'One (of them) stayed in Waipū.'

The speaker's brothers were discussed previously, but he contrasts *tētahi* 'one of (them)' with the others.

Two subtypes of focus were also identified: *new-information focus* (FOCn), new and missing information which develops the discourse (see (1)); and *contrastive focus* (FOCc):

(4) Ko koe anake kei te mōhio ki tērā
FOC you alone TAM know DO that
'You alone know that.'

The Māori language (='tērā') is under discussion. '*Ko koe anake*' (only you) is contrasted with a semantically and/or syntactically parallel constituent in the discourse (i.e. other people) ([17]:172).

2.3. Prosodic analysis

To investigate a relative prominence distinction between focus and topic *ko* (cf. [10]), we compared the prosodic realization of the most prominent word in the *ko*-phrase, and that in the main clause. We manually segmented the word with the strongest accent in the *ko*-phrase and the main clause, including any preceding particle to capture the accentual rise, e.g. *ko tētahi* in the *ko*-phrase in (3). Accented words which were phrase final were labelled differently from non-phrase-final words, as the F0 contours of the former would be affected by the boundary tone. Mean acoustic measures and time-normalised F0 contours were automatically extracted for each labelled section using ProsodyPro tools [18].

3. Results

Table 2: *Frequency of* ko *subtypes by speaker group*

	Speaker Group			
Info	**K**	**L1Y**	**L2Y**	**Total**
FOCn	15.7%	5.9%	0.0%	16
FOCc	16.9%	17.6%	2.9%	24
TOPa	37.3%	33.3%	35.3%	60
TOPc	30.1%	43.1%	61.8%	68
Total	83	51	34	168

3.1. Frequency of focus/topic *ko*

The first analysis looked at the functions of *ko* for the different speaker groups (see Table 2). A logistic mixed effects model was built using the *lme4* package in *R* [19,20] with Info Type (TOP v FOC) as the dependent, Speaker Group as the fixed effect and Speaker as the random effect (N=168). This showed that all groups were significantly more likely to use TOP *ko* than FOC (Intercept p=0.002), however L2Y speakers were significantly more likely to use TOP than the others (p=0.008); there was no difference between K and L1Y.

We can therefore see that the older K speakers, and the younger L1Y speakers, use *ko* to mark both topics and foci; though focus *ko* is very rare for the L2Y speakers. Within the subtypes of FOC, it appears L1Y speakers are more likely to use FOCc function than FOCn, while K speakers use both functions equally (see Table 2). Within the subtypes of TOP, it appears L2Y speakers, and to a lesser extent L1Y speakers, use TOPc more than TOPa, while K speakers use both equally. However, these differences could not be verified using statistical tests because of the low counts.

3.2. Relative prominence of *ko*-phrase and main phrase

For traditional Māori (as spoken by the K speakers), it is claimed that for topic *ko* the main stress is in the main clause, whereas for focus *ko* it is in the *ko*-phrase (see section 1). Therefore, for focus *ko,* we should expect to see a drop in the acoustic correlates of stress between the *ko*-phrase and the main clause; but no drop for topic *ko*. We measured two key acoustic correlates of stress, mean F0 and intensity. In each case, the difference in the measure between the most prominent word in the *ko*-phrase and the main clause is taken (see section 2.2). As this is conversational data, with huge segmental variability, it was very difficult to measure durational differences.

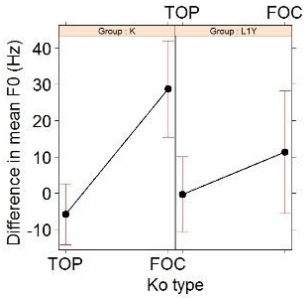

Figure 1: *Difference in mean F0 between the ko-phrase and main clause by ko Type and Speaker Group*

A linear mixed-effects model was built using the *lme4* and *lmerTest* packages [21] with the difference in mean F0 between the *ko*-phrase and the main clause as the dependent, Speaker Group (K v L1Y), *ko* Type (FOC v TOP) and their interaction as fixed effects, and Speaker as a random effect (N=120). The model showed a large drop in F0 when *ko* is a FOC (p<0.0001). The drop was much smaller for the L1Y speakers, although the

difference between the groups only approached significance (p=0.067). An ANOVA comparing models with and without the interaction was marginally significant (p=0.057), but the data set was small. The model effects can be seen in Figure 1 (a larger difference indicates a greater fall in F0 between the *ko*-phrase and the main clause).

A linear mixed-effects model was built with the difference in mean intensity as the dependent, and the same fixed and random effects as above (N=132). This showed that there was a significant drop in intensity with FOC *ko*, compared to TOP *ko* (p<0.0001); however, there was much less of a drop for the L1Y speakers than the K speakers (p=0.021). An ANOVA showed that the model with the interaction between *ko* Type and Speaker Group was significantly better than one without (p=0.021). The model effects can be seen in Figure 2.

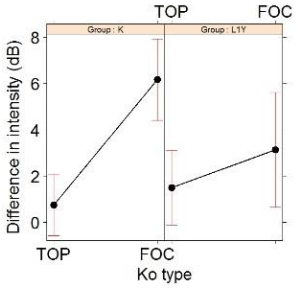

Figure 2: *Difference in mean intensity between the* ko-*phrase and main clause by* ko *Type and Speaker Group*

These results follow the expected pattern for the older K speakers: there is a large drop in mean F0 and intensity between the *ko*-phrase and the main clause for focus *ko*, consistent with the *ko*-phrase carrying the main stress. No drop is found for topic *ko*, consistent with the main stress occurring in the main clause. The L1Y speakers show the same pattern, but the F0 and intensity drops are much smaller for focus *ko*, and there is much more overlap in prominence for the two functions of *ko*. The L2Y speakers were not included as they had very few focus *ko* tokens. Interestingly, their differences for topic *ko* were roughly in between the values for topic and focus *ko* for the L1Y speakers (mean F0 difference=3.6Hz, sd=17Hz, outliers removed; mean intensity difference=1.8dB, sd=2.9dB).

3.3. F0 contours of *ko-phrase* and main clause

The last part of our analysis is exploratory. We wished to consider the F0 contour of the *ko*-phrase and the main clause as a whole, to see how the relative prominence differences found above manifest themselves for different groups, and if any of the sub-types of topic or focus are being distinguished.

Figure 3 (next page) shows time-normalised F0 contours for the most prominent word in the *ko*-phrase and the main clause, where these were phrase-final, produced by the ProsodyPro tools (see section 2.3. Non-phrase-final words are not shown for space reasons, as the phrase-final word would have to be shown separately). For the topic *ko*-phrases, sub-types are shown (blue dashed lines for TOPa, solid red for TOPc). There were no clear differences between FOCn and FOCc for the K group, and very few tokens for the L1Y group, so these are grouped.

For the K speakers, although there is variation, we can see that the contour for topic *ko* is generally either flat or rising (Fig. 3ai), while focus *ko* starts high and falls (Fig. 3bi). The main clause generally has a clear rise-fall when it is focal (Fig. 3aii), but a flat contour at a lower F0 when the *ko*-phrase is focal (Fig. 3bii). The L1Y speakers show a similar pattern for the *ko*-phrases (Figs. 3ci&di). However, apart from a few exceptions, the main clause has a flat, low pattern whether it carries the focus or not (Fig. 3cii&dii), although the drop is greater when the *ko*-phrase is focal. For the L2Y speakers, the F0 contour for the *ko*-phrase is falling, i.e. similar to the focus pattern for the other groups (Fig. 3ei); there is also an F0 drop in the main clause, although there is also often a rise-fall accent (Fig. 3eii). Interestingly, the L1Y speakers seem to distinguish TOPa and TOPc sub-types, i.e. TOPa contours are usually low and flat, while TOPc are higher and rising (Fig. 3ci). K speakers have a mix of flat and rising contours for topic *ko* as well, but these do not consistently map to the aboutness/contrastive distinction. The function of these for the K group remains to be explored.

4. Discussion

The results of this study are consistent with the descriptions of focus/topic *ko* in the literature for the older male speakers (K). When *ko* marks focus, there is a clear drop in prominence between the *ko*-phrase and the main clause, consistent with the main stress being in the *ko*-phrase; whereas for topic *ko* the strongest accent in the main clause is equally or more acoustically prominent than that in the *ko*-phrase, consistent with this carrying the main stress. For this generation of speakers, the organization of information fits the pattern expected for Polynesian and other verb-initial languages.

There were clear differences among the young male speakers depending on whether they were first (L1Y) or second (L2Y) language speakers of Māori. L1Y speakers broadly followed the patterns for K speakers, however, there were signs of change consistent with greater contact with English. Focus *ko* was usually contrastive for these speakers, and contrastive topics were more common than aboutness, so *ko* may be developing into a marker of contrast. The focus position in English is usually final, however, foci can be initial if they are contrastive [11]. There is a relative prominence distinction between topic and focus *ko*, but is much weaker. Interestingly, there are signs that L1Y speakers may be developing distinct contours to mark subtypes of topics: flat for aboutness, rising for contrastive. This type distinction is salient in English. The K speakers appear to have these two topic contours as well, but we have not yet established the functional distinction between them. It may be that their functions have been remapped for the L1Y generation, similar to what Queen observed for Turkish-German speakers [2]. The L2Y speakers primarily use *ko* to mark topics. This is consistent with the influence of their L1 English focus-final pattern. Notably, however, the prosodic pattern for their topic *ko* more closely resembles focus *ko* for the other speaker groups: a falling contour on the *ko*-phrase and lower pitch in the main clause. It may be that these speakers have acquired the more phonetically salient focus *ko* prosodic contour, but applied it to the more familiar topic function.

This investigation of focus/topic *ko* has given us a window into information structure in Māori and its prosodic realization, and how this is being influenced by contact with English over time. The area is fascinating, as traditional Māori and English have opposite information ordering strategies and differing intonational plasticity. The research is just beginning: to date we have only considered one focus/topic marking construction with a limited data set. The prosodic analysis is also exploratory. The kinds of prosodic contour distinctions discussed in section 3.3 may be differences in gradient prominence or phonological tonal type. However, the evidence is that the impact of English contact is complex, with the same syntactic and prosodic resources used for different functions between generations of Māori speakers.

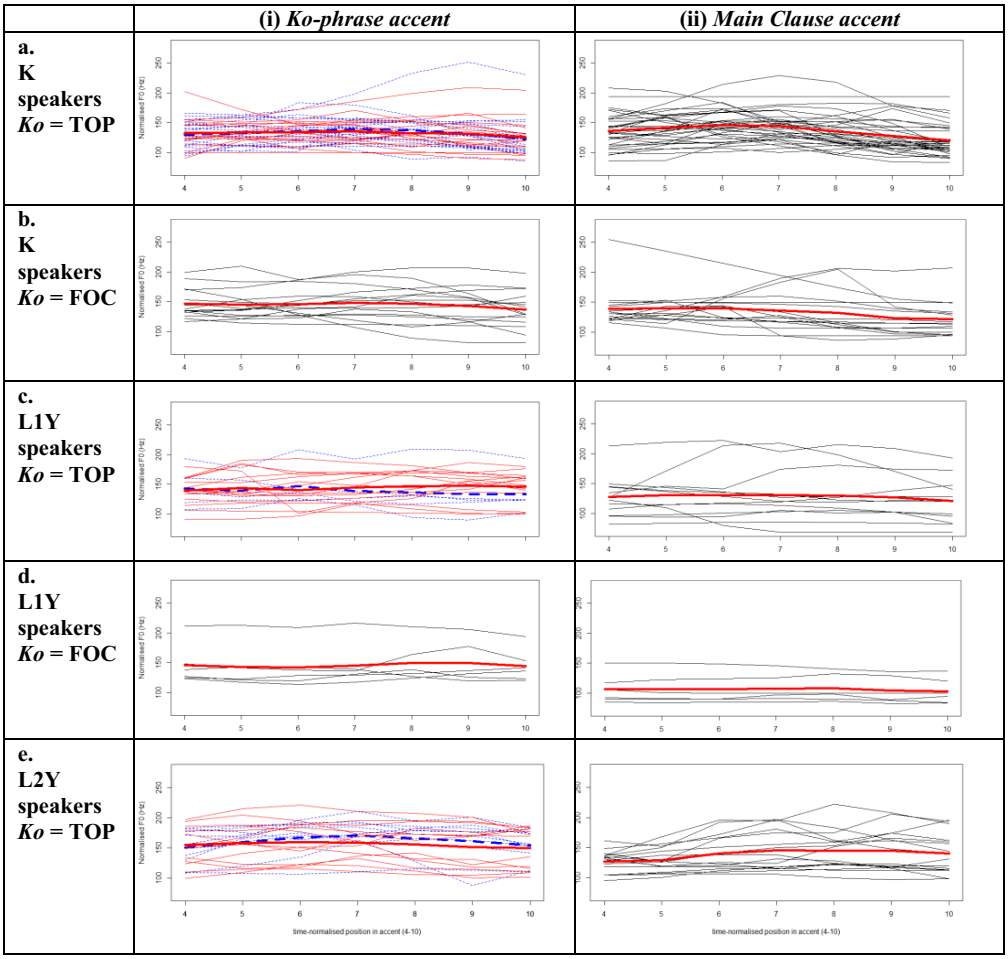

Figure 3: *Time-normalised F0 contours for phrase-final accented words in the* ko*-phrase and main clause by Speaker Group and* ko *Type. F0 level was sampled at 10 equally spaced points, points 4-10 are shown. The thick line (red or blue) shows the mean over all contours. For the TOP contours, dashed blue lines show TOPa, and solid red TOPc. Pitch tracking errors were removed.*

5. Acknowledgements

Mei kore ake te tīma MAONZE, i ora ai te kaupapa nei, me te karahipi a VUW hei tautoko i a Naoko.

6. References

[1] Sankoff, G. 2001, Linguistic Outcomes of Language Contact. In J.K. Chambers, P. Trudgill & N. Schilling-Estes (eds), *Handbook of Language Variation and Change*. UK: Blackwell, 638-668.

[2] Queen, R. 2012. Turkish-German bilinguals and their intonation: Triangulating evidence about contact-induced language change. *Language*, *88*(4), 791–816.

[3] Ladd, D. 2008. Intonational phonology (2nd ed), UK:CUP

[4] Kelly, K. 2015, *Aspects of change in the syntax of Māori – a corpus based study*, Ph.D. thesis, Victoria University of Wellington.

[5] King, J., Maclagan, M., Harlow, R., Keegan, P. and Watson, C. 2010. The MAONZE corpus: Establishing a corpus of Maori speech, *New Zealand Studies in Applied Linguistics* 16(2):1-16.

[6] Thompson, L. 2015. *Eliciting and analysing perceptions of prosodic prominence: a Māori case study*, PhD thesis, U. of Auckland.

[7] Bauer, W. 1993. *Maori*, Routledge, London.

[8] Calhoun, S. 2015. The interaction of prosody and syntax in Samoan focus marking. *Lingua* 165: 205-229.

[9] Herring, S. 1990. Information structure as a consequence of word order type, in *Proc. of BLS* (pp 163-174).

[10] Bauer, W. 1991, Maori *ko* again, *Te Reo* 34:3-14.

[11] Gundel, J. & Fretheim, T. 2005. Topic and Focus. In R. H. Laurence & G. Ward (eds), *The Handbook of Pragmatics*, UK: Blackwell, 175-196.

[12] Bauer, W. 1997. *The Reed reference grammar of Māori*, Reed, Auckland.

[13] De Lacy, P. 2003. Constraint universality and prosodic phrasing in Māori. In A. Carpenter, A. Coetzee & P. de Lacy (eds), *Papers in Optimality Theory II*, GLSA, 59-79.

[14] Watson, C., MacLagan, M., King, J., & Harlow, R. 2006. Are there L1 and L2 effects in the speech of young speakers of Mäori?. In *Proc. of SST* (pp 317-322). Auckland.

[15] Fromont, R. & Hay, J. 2012. LaBB-CAT: an Annotation Store. http://www.aclweb.org/anthology/U12-1015

[16] Boersma, P. & D. Weenink 2016. Praat: doing phonetics by computer [Computer program], http://www.praat.org/.

[17] Skopeteas, S., Fiedler, I., Hellmuth, S., Schwarz, A., Stoel, R., Fanselow, G., Féry, C., Krifka, M., 2007. Questionnaire on information structure: reference manual. In Ishihara, S., Schmitz, M. (eds), *Working Papers of the SFB 632*.

[18] Xu, Y. 2013. ProsodyPro – A Tool for Large-scale Systematic Prosody Analysis", in *Proc. of TRASP 2013*, Aix-en-Provence.

[19] Bates, D., Maechler, M., Bolker, B., Walker, S. 2015. Fitting Linear Mixed-Effects Models Using *lme4*. *Journal of Statistical Software*, 67(1), 1-48.

[20] R Core Team 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

[21] Kuznetsova, A., Brockhoff, P., and Christensen, R., 2016. lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-30. https://CRAN.R-project.org/package=lmerTest