Tonal Activity in Kara, an Austronesian language spoken in New Britain

John Hajek and Mary Stevens
School of Languages
University of Melbourne,
Australia
j.hajek@unimelb.edu.au, mes@unimelb.edu.au

Abstract

This paper presents the results of a small phonetic investigation of tonal activity in Kara, a little-known Austronesian language spoken in Papua New Guinea. Sketchy reports of some kind of tonal contrast in this language surfaced in the 1960s and 1970s, only to disappear in later published references to the language. Our auditory and acoustic investigations confirm the existence of contrastive tone in Kara. Native speaker intuitions also support such a conclusion. At least two tonemes (high and low) are identified. A third tone level (mid) is also noted but appears to be a variant of the low toneme.

1. Introduction

Although the Austronesian language family is the largest in the world, very few of its members are known to show evidence of contrastive tone. Those that have been confirmed to be tonal are geographically scattered. They include a small number of languages in New Caledonia, the Raja Empat languages of western Papua, Utsat in China, and Jabem and Bukawa spoken along the coast of the Huon Gulf area of New Guinea. The tone systems of these languages have been described and are well known (see Remijsen 2003, Edmondson et al 1993 for details). However, the possibility of tone in other Austronesian cannot be excluded. Little known, for instance, are very sketchy reports of tonal activity in Kara and closely related languages spoken on the island of New Ireland in eastern Papua New Guinea. Unfortunately, published information about tonal activity in Kara is restricted to brief mentions in two survey reports. Lithgow and Claassen (1968) note specifically that "..clearly contrasting tones were noted on words of a similar syllable pattern...". The only examples they give are fāvús 'meat' (mid-high) v fūwàn 'fat' (mid-low) (p.10). Capell (1971) in his survey of Austronesian New Guinea also refers briefly to tone in Kara. In a single sentence, he claims Kara to be a two-tone language (high v low) and that tone may be semantic. He gives the contrasting pair in Kara fói 'hair' (high) v fòì 'clean' (low). Later references to Kara make no mention of tone (see Hajek 1995 for details).

Based on these published sources, the status of tone in Kara can only be described as highly uncertain, and remains to be properly determined by phonological and phonetic investigation of new data. This paper attempts to address precisely this issue by presenting the results of a pilot study investigating tone in Kara. This task is now possible because of access to a small corpus of recorded materials, and to transcribed materials collected some 40 to 50 years ago.

For the purposes of this study, we relied on a range of different data types and approaches: native speaker intuitions about tone, some limited phonological analysis, as well as auditory and acoustic analysis. With respect to the latter, we inspected fundamental frequency (and its perceptual correlate, pitch), duration and intensity for vowels in a sample of Kara mono- and disyllables.

1.1. The language

Kara is a member of the Western Oceanic group of Austonesian languages. Its closest neighbours are sister languages, Tigak and Nalik. Neither of these languages is reported to show any signs of tonal activity. There appears to be some dialect variation in Kara, although the extent and implications are not known.

Kara has the following inventory of consonant phonemes: /p t q b d g ɸ s β ɣ m n ŋ l r ʔ/. In the Luaun dialect of Kara for which we have recordings, bilabial fricatives appear to be in free variation with labiodental [f v] respectively. Kara also has a marked uvular quality – seen in the absence of velar /k/, the presence of /q/ and frequent backing of /g/ and /ɣ/. /q/ is also frequently affricated.

Amongst vowels, /a/ and /ə/ are contrastive, but show significant overlap and a marked morphophonemic relationship in our recorded materials: low /a/ is frequently partly or fully raised along a phonetic continuum to mid central position.

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004. Copyright, Australian Speech Science & Technology Association Inc.

Accepted after abstract only review

With regard to phonotactics, Kara allows open and closed syllables, and medial clusters are also permitted. There are no particular restrictions on word length: words range from 1 to 5 syllables in length in Lithgow and Claassen's wordlist.

1.2. Data Sources

For the purposes of this study, we relied on data from two sources: (1) our own recordings, discussed below; and (2) unpublished fieldnotes collected by Lithgow and Claassen in 1966 and by Capell in the early 1950s.

At the time of our own data collection in the field, we had access to a copy of Lithgow et al's original short wordlist for Kara collected in 1966. Of 62 items, 30 multisyllabic items are marked for tone. Three tone levels are indicated by means of diacritics placed above the vowel: à low (L), ā mid (M) and á high (H). Stress placement is also indicated in 23 of these items, as well as on some items unmarked for tone. Neither stress nor tone are marked on monosyllables in their list.

Quite some time after recordings were made, we were given access to handwritten fieldnotes collected by Capell in the 1950s. Two tone levels are marked: á (high) and à (low). Unfortunately, Kara items appear numbered, but without translation into English. Only a few items have been identified, but these have proven useful in some cases when our tone transcriptions differed from those noted by Lithgow and Claassen.

The phonetic data used in this study are drawn from recordings collected by the first author during a short research visit to Papua New Guinea in the mid-1990s. A Sony Walkman Professional cassette recorder with a high quality external microphone were used to make recordings of one speaker under fieldwork conditions. The speaker was an adult male in his late 40s, who had been temporarily evacuated from the Bismarck Archipelago to the mainland as a result of volcanic activity. His native dialect is the variety of Kara spoken in the village of Lauan, on the east coast and close to Lemakot, where Lithgow et al (1968) first recorded tone. The speaker is also a fluent speaker of Tigak, Tok Pisin, and has a good knowledge of other varieties of Kara.

The Kara dialects of Lauan and Lemakot are similar, but there appear to be some lexical and phonetic differences. For instance, velars appear in Lithgow and Claassen's Lemakot Kara wordlist in the place of uvulars in Lauan Kara. Where lexical items were not the same, they were excluded from further consideration.

As time was limited, and there was no opportunity for further contact, recordings focussed primarily on discussions between the first author and the speaker about possible tonal activity in Kara, as well as elicitations of items marked with tone diacritics in the wordlists collected by Lithgow and Claassen during their original fieldtrip to New Ireland in 1966. Additional examples of monosyllables, drawn from the original wordlist but not marked for tone, were also recorded.

2. Methods

2.1. Pre-experimental Procedure

During fieldwork, metalinguistic information about tone was sought directly from the Kara subject, and a little time was spent on trying to identify monosyllabic minimal pairs.

More recently for the purposes of this study, lexical items marked with tone in the Lithgow and Claassen list were typed up, and tone patterns involving low (L) mid (M), and high (H) diacritics were tabulated. These were then analysed allowing for phonological or phonetic generalizations to be made.

2.2. Experimental Procedure

Recorded materials were digitized and sound files created. For the purposes of this pilot study, only monosyllables and disyllables were included in the data corpus. Disyllables were restricted to those items in Lithgow and Claassen's list that were marked for tone. Monosyllables were items that also appeared in the same list, but for which tone and stress was never indicated. The acoustic corpus consisted of 27 words (12 monosyllables, 15 disyllables). The number of tokens for each lexical item varied from one to five, giving a total token corpus of 49 items.

We began with an auditory analysis of the corpus. The authors listened separately to each token and made independent judgments about tone level, and possibly related effects such as vowel durations. We also tried to identify the most prominent syllable in each word, which was then marked as stress-bearing. Results were later compared. Based on the ratings of the two authors, vowel tokens were then divided into two groups for subsequent acoustic comparison: (1) high and (2) low.

Acoustic analysis of all tokens followed. We used spectrographic and waveform displays, as well as data extraction functions within the Praat program to gather relevant acoustic data for each vowel in the corpus. Acoustic measures of interest were fundamental frequency (F0) over the course of the vowel, duration, and peak and average intensity levels of each vowel.

All measurements were tabulated and statistical tests were, unless otherwise indicated, conducted on each of the measurement sets using single-factor ANOVA tests within Excel. Although F0 was calculated at regular 15 ms. intervals in each vowel, for the purposes of this study we concentrated on F0 values at 4 different points: soon after vowel onset, mid-point, before vowel offset, as well as the peak F0 value.

Accepted after abstract only review

3. Pre-experimental Results

From discussions with our subject, it is clear that Kara speakers are aware of tone in their language. They have terms to describe pitch movements, including 'tone errors' made by speakers of other dialects. In the limited time available, the subject was also able to provide an example of a word pair with shared segmental structure but differing tones: Lauan Kara /qól/ (high) v. Lemakot Kara /qòl/ (low) 'down the beach'. Eventually, we were also able to identify a minimal pair in his own dialect: /ŋót/ 'louse' (high) v /ŋòt/ 'to weed' (low). He was not able, however, to identify the number of different tones in Kara.

With respect to our phonological analysis of Lithgow and Claassen's wordlist transcriptions, a number of generalizations about tone patterns can be made. Mid tone occurs most frequently, and high most rarely. Only one tone mark can appear on a syllable.

There are evident restrictions on tone distributions. L occurs only at right-edge, and appears to mark a phonetic process of word-level declination, e.g. MHL, ML. Apart from a single example of antepenult H, H always occurs on the last or second to last syllable - but only appears in 1/3 of lexical items in the corpus. M is the only tone that can appear on more than one syllable in any word, e.g. MM, MML, and MMHL.

With respect to the marking of stress by Lithgow and Claassen, there are clear interactions with tone. Final syllables can only be stressed if H. L is never stress-bearing. H is always stress-bearing, e.g. M$^{\text{ˈ}}$HL and $^{\text{ˈ}}$HL. In words without H, only M can be stressed. In MM sequences, stress is on the first M, e.g $^{\text{ˈ}}$MM, $^{\text{ˈ}}$MML.

Analysis of Lithgow and Claassen's transcriptions point to a kind of pitch accent system (see Donohue 1997). There is a single pitch contour over a word, with a stress-marked peak on one syllable (there is one exception involving MLM). A change in pitch level occurs with each new syllable, with the exception of disyllabic MM sequences.

The phonological status of H is not in doubt. When present, it is always most prominent. M and L appear to be for the most part in complementary distribution, and it is tempting to agree with Capell (1971) that Kara has two tones: high (H) and low (= M and L). However, Lithgow and Claassen transcription of disyllables shows potential for a three-way contrast:

M$^{\text{ˈ}}$H	[fɑ̄$^{\text{ˈ}}$vús] 'meat'
$^{\text{ˈ}}$MM	[$^{\text{ˈ}}$fɅ̄t$^{\text{h}}$ɑ̄] 'fire'
$^{\text{ˈ}}$ML	[$^{\text{ˈ}}$khɅ̄nʌ̀] '(his) foot'

It is possible that the transcribed difference between final M and L reflects segmental perturbation of F0 (see Gandour 1978, Hombert 1978). We return to this point below.

4. Phonetic Results

4.1. Auditory Analysis

There was very high inter-rater agreement between the two authors about tone and stress placement in tokens. In the first instance, a general distinction between high pitch and non-high pitch (pitch being the perceptual correlate of F0) was easily identified. We also confirmed the tonal distinction in the previously suggested minimal pair /ŋót/ 'louse' v. /ŋòt/ 'to weed'.

High tone was consistently identified in monosyllables and disyllables. It was also always marked as most prominent or stressed in the latter group. Stress placement was always agreed, regardless of tone level. Non-high tone was almost always marked as low – in both monosyllables and disyllables. Amongst disyllables, one word was identified as L$^{\text{ˈ}}$M by the second author and as uncertain L$^{\text{ˈ}}$L ~ L$^{\text{ˈ}}$M by the first author. Amongst monosyllables, all non-high tones were marked as low, with the exception of three words that were marked as M by the second author, and as uncertain M or L by the first author. All items marked as having possible M on the stressed syllable ([fəʔ] 'tooth', [maʔ] hand', and [qaʔ] 'foot', and [və$^{\text{ˈ}}$taʔ] 'cloud') were the only items to share a final glottal stop. This particular distribution suggests that M, if it exists, is an allophone of L before final glottals. Given the small number of tokens involving M, its evident allotonetic distribution, and uncertainty about its identification as M or L, it was decided to treat all such cases as L for the purposes of subsequent general analysis in this study.

Overall the following tone patterns were noted: H and L in monosyllables (with uncertainty about M or L in some tokens), and L$^{\text{ˈ}}$H, $^{\text{ˈ}}$LL, L$^{\text{ˈ}}$L (or possibly L$^{\text{ˈ}}$M in one example) in disyllables.

Our tone and stress transcriptions did not fully align with those given by Lithgow and Claassen. Where they marked M, we almost always marked L. Stress placement and identification of high tone also differed on occasion, e.g. their [$^{\text{ˈ}}$rārùm] 'water', for which we both gave [rà$^{\text{ˈ}}$rúm]. In this case we agree with Capell (1971) on tone value, but not on stress value where he instead offers [$^{\text{ˈ}}$ràrúm]. A small number of items which we considered to be L$^{\text{ˈ}}$L are marked by Lithgow and Claassen as $^{\text{ˈ}}$ML or $^{\text{ˈ}}$MM. It is not clear whether these differences in transcriptions are the result of varying perceptual responses by different listeners, or of dialect differences between speaker subjects.

4.2. Acoustic Results

The existence in Kara of a tonal minimal pair in /ŋót/ 'louse' (high) v. /ŋòt/ 'to weed' (low) is confirmed by acoustic analysis. Figure 1 plots the F0 trace over the

major data points and shows a clear separation of approximately 50Hz over the course of the trace:

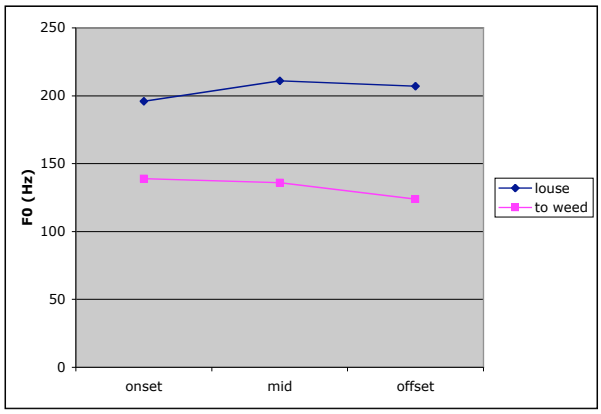

Figure 1. *F0 measurements at the onset, midpoint and offset of the minimal pair* [Noèt] *'louse'* (H); [Noòt] *'to weed'* (L).

Acoustic inspection also helped us to understand our perceptual identification of a possible M tone in the four words listed above ([fəʔ] 'tooth', [maʔ] hand', [qaʔ] 'foot', and [vəˈtaʔ] 'cloud'). Table 1 shows the average F0 values for these four words, and average F0 values for all other vowels marked as L and stress-bearing:

Table 1. Average F0 values for vowels marked as bearing stress and low pitch, according to whether they have a glottal coda ('L /_ʔ/) or not (Other 'L).

	Onset	Mid	Offset	Peak	No.
'L /_ʔ/	130	131	133	133	5
Other 'L	124	116	110	124	17

The value of F0 across the final vowel before [ʔ] was 5-20Hz higher than on other stressed vowels marked as low, and the trajectory of F0 also shared the general property of rising towards vowel offset. In contrast, for words with a final stressed tone that was clearly identified by both authors as L, there is a general pattern of F0 decline during the course of the vowel. Our results are consistent with Hombert's (1978) finding that rising F0 is a characteristic perturbation effect of syllable-final glottal stops.

With respect to our data set, we then calculated F0, duration and intensity values for each vowel, according to the two main tone categories High and Low, and to word length. The second tone category was further sub-divided into stressed and unstressed Low to allow for more detailed analysis of possible effects of the presence or absence of stress – within and across categories.

In Table 2 we summarise F0 values across a range of syllable and stress conditions. A substantial separation of average F0 values between H and L is evident in all comparisons, which can also be seen clearly in Figure 2.

Table 2. Average F0 values (and st. dev.) in Hz at the 4 measurement points for vowels according to pitch label (H/L), word i.e. mono- (σ) or disyllable (σσ), and stress i.e. stressed (') or unstressed (*) for disyllables.

Label	Word	Onset	Mid	Offset	Peak
H	σ	169(21)	187(20)	181(29)	193(20)
L	σ	131(15)	125(10)	120(14)	133(14)
'H	σσ	162(24)	173(25)	167(29)	187(43)
'L	σσ	119(7)	128(10)	111(10)	131(11)
*L	σσ	114(15)	112(12)	113(10)	115(15)
'L&*L	σσ	116(14)	113(11)	113(10)	116(13)
H	(σ)σ	165(22)	180(24)	173(30)	190(34)
L	(σ)σ	118(15)	115(12)	114(11)	119(15)

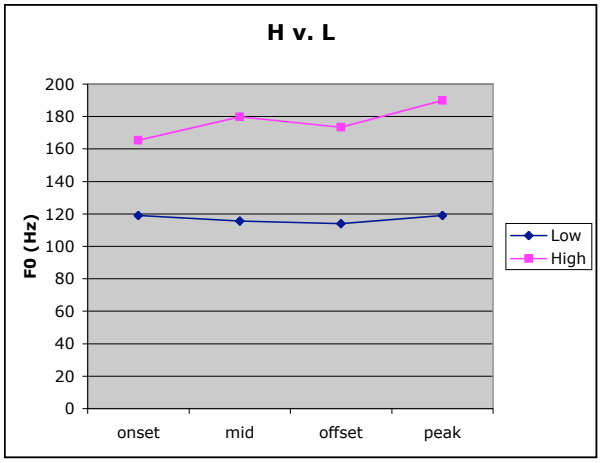

Figure 2. *Averaged F0 measurements at the onset, midpoint, offset and peak for all vowels identified as bearing High or Low tone.*

Statistical analysis of F0 values at the four measured points (onset, mid, offset and peak) shows highly significant differences between H and L at all points across all words, as well as monosyllables and disyllables ($p < 0.001$). Within disyllables the same significant effect was apparent when H was compared with stressed ˈL and unstressed *L. However, there was no significant difference between ˈL and *L at any point of comparison ($p = 0.347 \sim 0.823$).

Intensity levels were also investigated, since its perceptual correlate, loudness, is often referred to across languages as important in determining relative syllable prominence or stress position in a word. We were also interested in knowing if and how it might interact with F0, and what role if any intensity played in our own

Accepted after abstract only review

assessment of tone and stress placement. We calculated the peak intensity value in each vowel as well as the average intensity value over the duration of the vowel. Results are given in Table 3.

Table 3. Mean peak intensity and average overall intensity values (dB) for vowels according to tone label and word length.

Label	Word	Peak	Average
H	σ	79(3)	76(3)
L	σ	79(3)	76(3)
'H	σσ	78(3)	75(3)
'L	σσ	72(5)	75(4)
*L	σσ	75(5)	73(5)
'L &*L	σσ	75(5)	73(4)
H	(σ)σ	79(3)	75(3)
L	(σ)σ	76(5)	73(4)

Overall results (H v. L) for peak intensity values show a relatively weak effect ($p = 0.02$) between H and L. However, lower level comparisons between categories often show no effect: there was no significant difference between H and L in monosyllables ($p = 0.676$), while in disyllables, the difference in peak intensity between H and 'L was just beyond significance ($p = 0.051$). An effect was found between H and *L ($p = 0.019$), but there was no significant differencee between stressed and unstressed L ($p = 0.824$). A slightly different pattern emerged when average overall intensity, rather than peak intensity, was analysed. Again a small overall effect was noted ($p = 0.026$). However, in monosyllables there was no significant difference in overall intensity between H and L ($p = 0.753$), whilst in disyllables a small effect was noted when H and 'L were compared ($p = 0.026$). However, there was no difference between H and *L ($p = 0.084$) or between 'L and *L ($p = 0.652$).

Finally we calculated the duration of all vowels, as seen in Table 4. Average values show H tones to be much longer than L tones, albeit with high standard deviations.

Table 4. Average vowel duration (ms.) and standard deviations according to tone label, word length and stress, and across all words

Label	Word	Duration
H	σ	159(78)
L	σ	123(34)
'H	σσ	122(60)
'L	σσ	92(48)
*L	σσ	81(34)
'L&*H	σσ	84(39)
H	(σ)σ	139(70)
L	(σ)σ	92(40)

Statistical analysis shows that overall there was a strongly significant difference in vowel duration between H and L tones ($p < 0.001$). Closer inspection again showed more variable effects: there was no duration effect in monosyllables ($p = 0.186$), nor between H and 'L ($p = 0.152$) in disyllables. In disyllables H vowels are significantly longer than unstressed L vowels ($p = 0.006$), but there is no difference in duration between stressed and unstressed L in disyllables ($p = 0.418$).

5. Discussion

Our results confirm earlier claims of possible contrastive tone or pitch in Kara. We have been able to adduce a range of different kinds of positive evidence in support of such a position. In the first instance, native speaker intuitions, phonological analysis of earlier transcriptions, and our own auditory assessment strongly support a two-way distinction between H and L, as suggested by Capell (1971).

Auditory inspection on our part found the separation between perceived high and non-high to be relatively easy. More problematic was the possible identification of a mid tone, but this could be accounted for as a predictable contextual effect of word-final [ʔ]. Given our results, we suggest that the frequent marking of mid tone by Lithgow and Claassen simply reflects their greater perceptual sensitivity than ours to local perturbation effects and the F0 differences that can result (see e.g. Umeda 1981 for such effects in English). We noted in §3.1 that low and mid were largely in complementary distribution in their transcriptions.

We also had no problems in identifying the most prominent syllable in disyllables. It always involved H if it was present. When H was absent, we also agreed on the identification of stressed L. However, it remains unclear what acoustico-perceptual cue was relied on to make such judgments: statistical analysis shows no significant effects to do with F0, duration or intensity between 'L and *L. Further investigation on this point is clearly needed.

Acoustic investigation shows quite clearly the importance of differences in F0 for the prosodic differentiation of syllables and words in Kara. Statistical analysis gives highly significant values for differences between H and L in all comparisons. Our results for Kara are, therefore, consistent with longstanding experimental results that show that F0 is the most reliable acoustic measure for the perception and identification of tone (see Gandour 1978, but see also Rose 1988 for some expression of caution).

Overall comparison of all H and L tones regardless of stress shows a significant effect for all acoustic measures (F0, peak and overall intensity and duration). However, it is clear that F0 is by far the most reliable

indicator of difference between vowels identified as bearing H and L tone. Closer inspection shows the effects of intensity and duration to be much less significant, and often inconsistent at lower level comparisons.

With specific reference to monosyllables, F0 is the only measure that reliably allows for H and L items to be distinguished. In disyllables, F0 is the only consistent measure of difference between H on the one hand and stressed and unstressed L on the other. Duration and peak intensity do not separate H from ˈL, but nor do they separate ˈL from *L. Average intensity has some effect, but this is much weaker than that found for F0. To distinguish between H and *L, duration differences are important, peak intensity has a relatively weak role, and average intensity has none, while once again F0 differences are highly significant.

With regard to the kind of prosodic system within which tone might be organized in Kara, our analysis of Lithgow and Claassen's transcriptions pointed to a pitch accent system (see Donohue 1997 for an overview of pitch accent and other tone systems in New Guinea). Results of subsequent auditory and acoustic analysis suggest that our initial transcription-based characterization of such a system for Kara needs to be modified. In particular, we note that final L can be stress-bearing. Based on our data, the system can be stated relatively simply: contrast in pitch (/H/ v /L/) is only possible on accented syllables. Otherwise, unstressed syllables are always low in pitch. Whether the location of the accented or stress-bearing syllable is predictable or not is still to be determined.

6. Conclusion

Although much remains to be investigated, our pilot study confirms the existence of contrastive tone in Kara. We can now definitively add Kara to the small list of Austronesian languages that are known to have this property. At this stage, we are certain that Kara has two tonemes (high and low), based on minimal pairs we have been able to find, as well as on native speaker intuitions and our own auditory, and acoustic analysis. There is some evidence of mid tone, but it appears to be a predictable variant of L before final glottal stops in our corpus.

F0 and its perceptual correlate, pitch, play a critical role in the prosodic system of Kara. However, further investigation of many aspects of this system is still needed. One particular puzzle that remains is the identification of the acoustico-perceptual cues used by listeners to assign stress or relative prominence in words that contain only low tones.

7. Acknowledgments

The authors wish to thank Hywel Stoakes for his assistance.

8. References

Capell, A. (1971). The Austronesian languages of Australian New Guinea. In T. A Seboek (ed) *Current Trends in Linguistics.* Vol. 8: *Linguistics in Oceania.* The Hague: Mouton.

Donohue, M. (1997). Tone systems in New Guinea. *Linguistic Typology* 1: 347-386.

Edmondson, J. & K. Gregorson (eds.) (1993) *Tonality in Austronesian languages.* Honolulu: University of Hawaii Press.

Fromkin, V.A. (1978). *Tone: a linguistic survey.* New York: Academic Press.

Gandour, J.T. (1978). The perception of tone. In Fromkin (ed), pp. 41-76.

Hajek, J. (1995). A mystery solved: the forgotten tone languages of New Ireland, *Working Papers in Linguistics* 14: 9-14.

Hombert, J.-M. (1978). Consonant types, vowel quality and tone. In Fromkin (ed), pp. 77-112.

Lithgow, D. & O. Claassen (1968). *Languages of the New Ireland district.* Port Moresby: Dept. of Information and Extension Services.

McCawley, J.D. (1978). What is a tone language? In Fromkin (ed), pp. 113-132.

Remijsen, B. (2003). *Word-prosodic systems of Raja Ampat languages*. Utrecht: LOT.

Rose, P. (1988). On the non-equivalence of fundamental frequency and pitch in tonal description. In D. Bradley, E.J.A. Henderson & M. Mazaudon (eds) *Prosodic analysis and Asian linguistics: to honour R.K. Spriggs.* Canberra: the Australian National University. Pacific Linguistics C 104, pp.55-82.

Ross, M. (1993). Tonogenesis in the North Huon Gulf Chain, In J. Edmondson & K. Gregorson (eds.) *Tonality in Austronesian languages.* Honolulu: University of Hawaii Press, 133-145.

Umeda, N. (1981). Influence of segmental factors on fundamental frequency in fluent speech. *Journal of the Acoustical Society* 70: 350-355.

Yip, M. (2002). *Tone.* Cambridge: Cambridge University Press.

Accepted after abstract only review