THE PHONETICS OF PAICĨ VOWELS[1]

[1] This work would not have been possible without the generous cooperation of the speakers who shared their knowledge of their language, and the hospitality of the community. Much essential assistance was provided by Stephen Schooling. The project was funded by the National Science Foundation.

This paper presents results of a phonetic study of the vowel system of Paicî, an Austronesian language of central New Caledonia. The Paicî vowel system is of phonetic interest for both its three-way lexical tone contrast, rare among Austronesian languages, and its relatively large inventory of both oral and nasalized vowels. The large number of nasalized vowels is rare not only from an Austronesian persepctive, but also is typologically atypical throughout the world. This paper focuses on the analysis of qualitative aspects of both the oral and nasalized vowels of Paicî. It is shown that vowel qualities posited in previous research on Paicî are phonetically differentiated, with the contrast between certain nasalized vowels being more subtle than the contrast involving their oral counterparts. In addition, the phonetic realization of the three tones of Paicî is discussed.

1. <u>Introduction</u>

Paicî is one of the twenty-five or so Austronesian languages indigenous to the French “overseas territory” of New Caledonia. It is spoken in a narrow band stretching across the center of the main island, Grande Terre, from Poindimié and Ponerihouen on the east coast to Pouembout and Koné on the west. The location is shown in the map in Figure 1. The number of speakers was estimated at approximately 5000 in the 1970’s. It is now likely to be on the order of 8000 as the indigenous component of New Caledonia’s population has shown vigorous growth in the last few decades. Paicî has the largest number of speakers of any of the languages of Grande Terre, and because of its importance it has been chosen by the territorial education authorities as the first local language to be (re-) introduced into use in primary education.

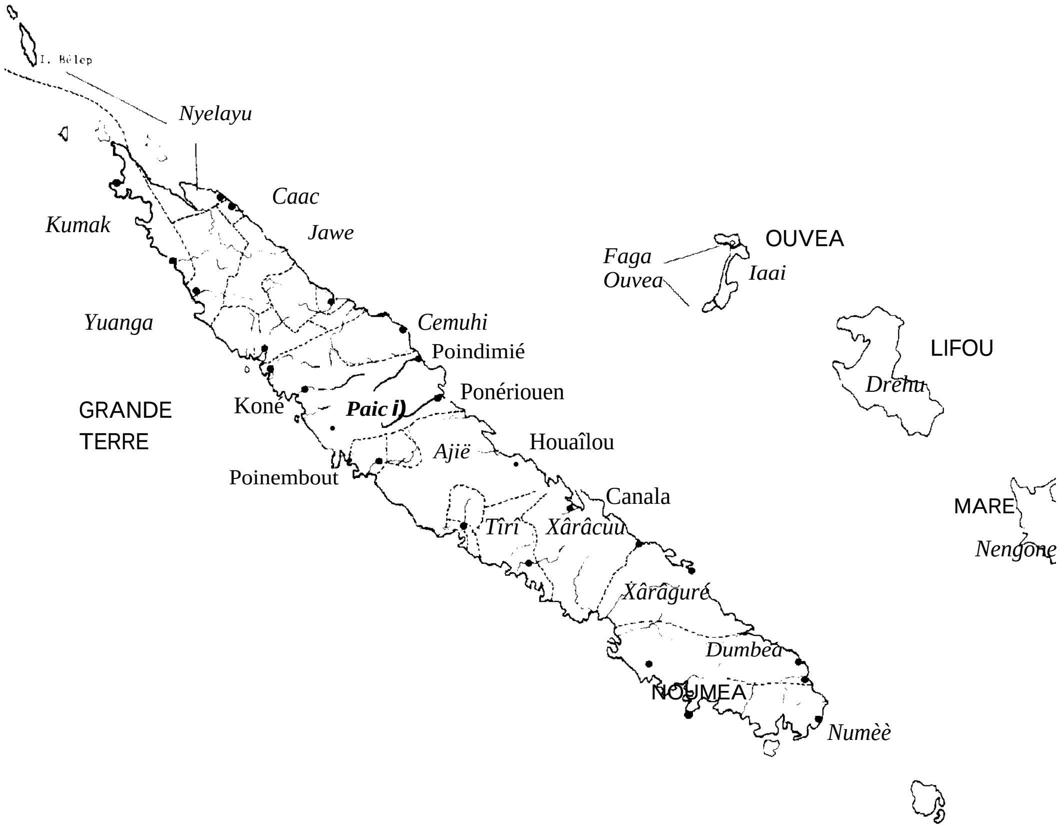

Figure 1. Map of New Caledonia showing areas in which the various indigenous languages are distributed (based on a map prepared by Ozanne-Riviere and Rivierre). Language names are italicized.

Ironically, recent administrative changes which have given the indigenous people greater influence in the political life of New Caledonia might pose a threat to the future of even the stronger languages, such as Paicî. This is because French is the main language for inter-ethnic communication, is needed for participation in the modern political and economic life of the territory, and most school-based education is only available in French. In view of these factors the number of speakers might well be expected to decline in future years even as the ethnic population grows.

From a phonetic point of view, Paicî is principally of interest for its extensive vowel system. It has ten oral vowels and seven nasalized vowels, which makes Paicî something of a rarity cross-linguistically. Only 26, or 8.1% of the languages in UPSID sample of languages of the world (Maddieson 1984) distinguish ten or more contrastive vowel qualities, and most languages with nasalized vowels contrast no more than five qualities in this category (Maddieson 1991). Paicî is also one of a relatively small number of Austronesian languages which is tonal (Rivierre 1974, 1978).

Yet, despite the exceptional nature of its vowel system and its important political status in New Caledonia, Paicî has not been the subject of any previous detailed phonetic examination. Besides the two articles by Rivierre cited immediately above, the only published general linguistic works on Paicî are a substantial dictionary published by Rivierre (1983), which we have relied on extensively, and a short grammar and word list which appeared in Leenhardt's (1946) comprehensive book on East Melanesian languages under the name *Pati*. Rivierre's dictionary is founded in part on materials collected earlier by Leenhardt, Grace, Haudricourt and others but was entirely checked in the field and much expanded.

Following a brief discussion of the overall phonological inventory the acoustic structure of the Paicî vowel system will be examined from three angles below. First, the phonetic realization of the tonal contrasts, rare among Austronesian languages, will be examined. Second, the location of the oral vowels in the acoustic space will be described, in order to add to the relatively small cross-linguistic phonetic database on large vowel inventories. By providing further data on how the vowels in large vowel systems are distributed we hope to contribute to understanding the principles governing this distribution. Finally, the acoustic patterns of the nasalized vowels will be examined. Numerous studies have shown phonetic asymmetries between oral and nasalized vowels (e.g. Wright 1986, Maeda 1993). Given its large number of nasalized vowels, Paicî provides a particularly good opportunity to examine the acoustic consequences of nasalization. This paper is only a preliminary report, and it is hoped that more detailed studies will follow.

2. Consonants

Compared to a number of other New Caledonian languages, including Iaai (Maddieson and Anderson 1995) and Ndumbea (Gordon and Maddieson 1999), Paicî has a relatively simple consonant system, which is presented in Table 1. Nasals, voiceless plosives and prenasalized voiced stops occur at four places of articulation. The segments written in the ‘alveolar’ column are in fact slightly post-alveolar. The palatal stops are released with quite strong frication and could well be considered to be affricates. Plain and labialized bilabial stops and nasals contrast before non-rounded vowels (before /ɔ/ the bilabials are noticeably labialized in pronunciation, but do not contrast with non-labialized counterparts). There is also a labial-velar approximant /w/ which is similarly restricted in its distribution, occurring only before non-rounded vowels and, occasionally, /ɔ/. For convenience, this is placed in the same column as the labialized bilabials on the chart. The apical tap /ɾ/ also has a restricted distribution, as it only occurs word-initially in one grammatical element (/ɾʌ/ 3 pl subject prefix). In a very high proportion of its occurrences the same vowel precedes and follows the tap, suggesting that a process of epenthesis may be involved. When this segment precedes a nasalized vowel it is also nasalized or is pronounced as a very brief nasal segment. The alveolar lateral approximant /l/ occurs initially only in a few loanwords, such as /lāācí/ ‘rice’, but is relatively common at the beginning of the second syllable of a word. There are no consonant clusters or syllable-final consonants in Paicî.

Table 1. Consonants of Paicî (after Rivierre 1983: 21)

	Bilabial	Labialized Bilabial	Alveolar	Palatal	Velar
Voiceless plosive	**p**	**pʷ**	**t**	**c**	**k**
Voiced pre-nasalized stop	**mb**	**mbʷ**	**nd**	**¯ɟ**	**ŋg**
Nasal	**m**	**mʷ**	**n**	**¯**	**ŋ**
Approximant		**w**	**l**		
Tap			**ɾ**		

3. <u>Vowels and Tones</u>

As noted earlier, Paicî distinguishes ten oral and seven nasalized vowel qualities. The ten oral qualities are given in Table 2, using the symbols suggested by Rivierre. As this table shows, Paicî adds a series of three back (or at least non-front) unrounded vowels to the 'standard' 7-vowel set. This type of expansion of a vowel inventory is less common than the addition of front rounded vowels (Maddieson 1984). We will consider below the question of whether the acoustic analysis supports the suggested notation, particularly with respect to the degree of backness indicated for these vowels.

Table 2.   Oral vowels of Paicî.

i   **ɯ u**
  e   **ɤ o**
    ɛ   **ʌ ɔ**
       a

The seven nasalized vowel qualities are given in Table 3, again using the symbols suggested by Rivierre. Vowels are always nasalized after nasals, and always oral after prenasalized voiced stops, but contrast elsewhere, as in the pair

/cóo/ ‘dress, cloth’ vs /cṍõ/ ‘look, watch’. In the nasalized vowel set the contrast between higher mid and lower mid qualities that is present among the oral vowels is lacking. Rivierre indicates that the remaining nasalized mid vowel is higher mid in the front unrounded and back rounded series, but lower mid in the back unrounded series. We will consider below the question of whether this reflects an articulatory difference between these vowels or could be accounted for by acoustic effects alone. Rivierre also suggests that the low nasalized vowel /ã/ varies in range between [ã] and [ɛ̃]. We will look for confirmation of the greater acoustic range of this vowel in our data.

Table 3.   Nasalized vowels of Paicĩ.

ĩ			**ɯ̃**	**ũ**
ẽ			**õ**	
		ʌ̃		
	ã			

Each of the vowel qualities in Tables 2 and 3 can occur distinctively long or short. The orthography used in Rivierre’s dictionary as well as in recent literacy and evangelical materials marks long vowels by doubling the vowel symbol. We have maintained this convention in our transcriptions. (The orthography otherwise uses only the five basic symbols of the Roman alphabet for vowels, requiring the resort to a variety of diacritics to mark further distinctions of height, rounding and nasalization.) Sequences of unlike vowels also occur quite frequently, although it is rare to find two long vowels adjacent.

Rivierre (1974) distinguishes three level tones, high [´], mid [¯] and low [`]. Following Rivierre, we leave low toned vowels unmarked. Some items are also ‘toneless’ in that their tone is predictable from the tone of the item they follow (usually they are at the same level). Immediately adjacent short vowels may bear different tones but a long vowel never bears more than one tone. A very high proportion of Paicî words with more than one syllable have the same tone on all their syllables, suggesting that the tonal patterns may be characteristics of the word. The phonetic realization of tone is discussed below in section 4.2.

4. Acoustic Analysis

4.1. Materials

A wordlist illustrating all the consonants and vowels of Paicî was prepared on the basis of the dictionary and the guide to the orthography prepared for schools. Words selected to illustrate consonants contained principally /a/ vowels, and words selected to illustrate vowels principally contained bilabial consonants (the high frequency of bilabials making this context the easiest in which to find minimal or near-minimal contrasts among the vowels). This list was recorded using a high-quality directional microphone with two groups of speakers in outdoor settings near Napoemi and Tibarama in the region of Poindimié. The participants in these recordings were eight native speakers of Paicî, five men and three women, ranging in age from their late forties to early twenties. The actual words recorded differed slightly in the two sessions, and some individual tokens were unusable because of wind noise, overlaid voices or other problems. These recordings, made in New Caledonia in February 1993 by the second author, serve as the basis for our analysis of the Paicî vowel system.

4.2. Tones

In order to analyze the phonetic properties of tone, fundamental frequency was measured for the vowel /a/ in three words which each contained a different tone: the high toned vowel in pwáa ‘white’, the mid toned vowel in kā̃ā̃ ‘cry’, and the second low toned vowel in ãwãã ‘Amoa river’. Because of difficulties often encountered with the pitch extraction algorithm for individual tokens, fundamental frequency was measured using a narrowband spectrogram. In order to increase the accuracy of the results, the center frequency of a higher harmonic, the 10th wherever possible, was measured from a narrow band spectrogram and divided by the number of that harmonic yielding a fundamental frequency value. Thus, for example, if the center of the 10th harmonc is at 1500Hz, the fundamental frequency is equivalent to 1500 divided by 10, or 150Hz.

Narrow band spectrograms show that the tones are characterized by different contours and/or differences in frequency. For many speakers, the high tone displays a quasi hat like pattern with a rise in frequency at the beginning of the vowel followed by a relatively steep fall which occupies over half of the duration of the vowel. The mid and low tones, on the other hand, typically show a level plateau through approximately the first half of the vowel followed by a gradual decline, or alternatively, a gentler fall throughout the duration of the vowel. Representative narrowband spectrograms of the three tones are illustrated in figures 2-4. The arrow in each figure points to the 10th harmonic, the one from which measurements were made.

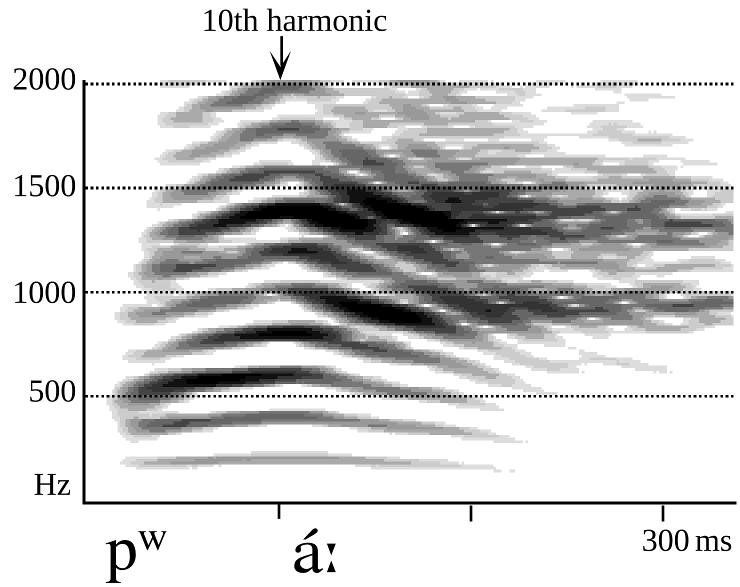

Figure 2. Narrow band spectrogram illustrating the high tone in pwáa ‘white’ (Speaker M1)

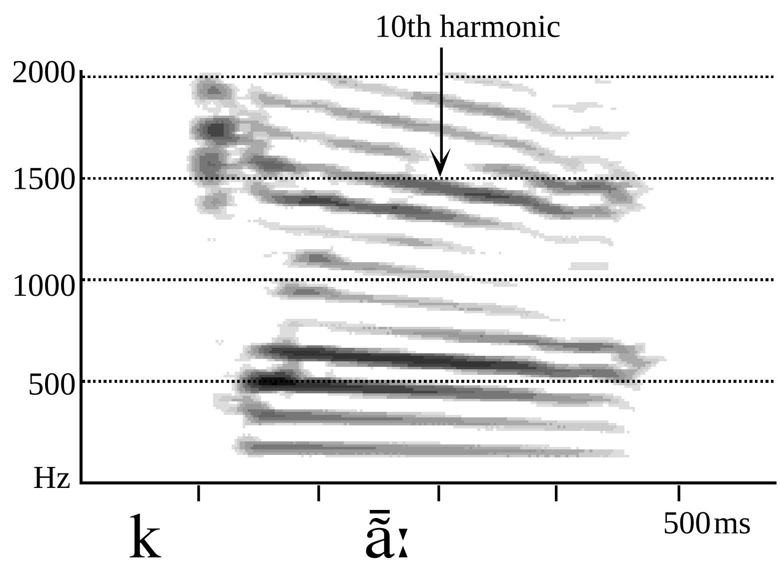

Figure 3. Narrow band spectrogram illustrating the mid tone in kā̃ā̃ ‘cry’ (Speaker M1)

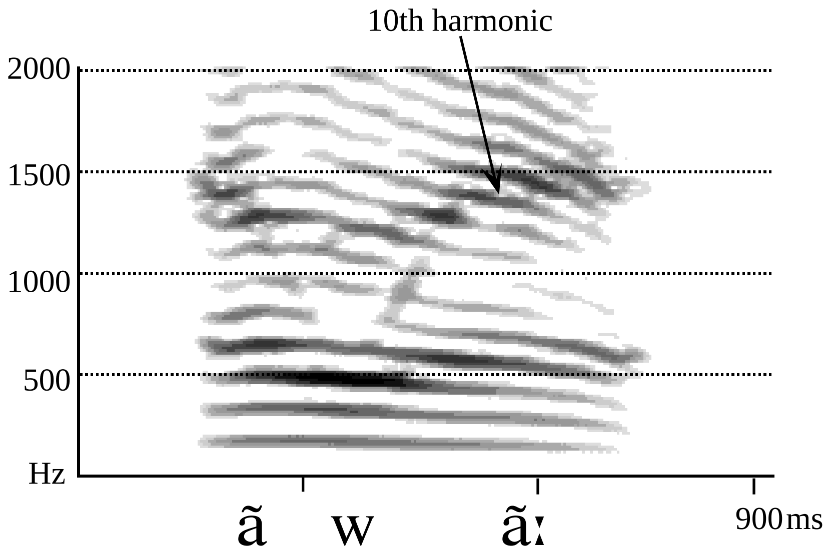

Figure 4. Narrow band spectrogram illustrating the low tone in ãwãã 'Amoa river' (Speaker M1)

The mid and low tones generally have similar shapes, consisting of a gradual fall throughout the duration of the vowel (possibly attributed at least partially to an utterance-final effect in the examined data), with the low tone having a slightly steeper fall and lower values for the fundamental: approximately 140Hz at its midpoint in figure 4 as compared to 150Hz for the mid tone in figure 3). The high tone is characterized by an initial rise followed by a steep fall; the fundamental frequency also has a much higher peak value (approximately 200Hz in figure 2) than either the mid or low tone.

4.3. Oral Vowels

In order to describe the principal acoustic characteristics of the Paicî vowels the frequencies of the first three formants were measured in a selected subset of the words recorded. The words from which the measurements were made are listed in Appendix 1. The analyses were performed using the Kay Elemetrics Computerized Speech Laboratory (CSL). The audio recordings were digitized at 20 kHz and formants obtained algorithmically using LPC analysis. Either 12 or

14 coefficients were employed, depending on the speaker's sex and the difficulty resolving the formants. A window of about 25 ms (26.5 ms) centered around the middle of the vowel's duration was chosen for the LPC. An FFT power spectrum calculated over the same window was consulted for confirmation of the values obtained from the LPC analysis. An alternative placement of the analysis window was tried when LPC and FFT analyses appeared to conflict.

Values for long and short vowels are generally very similar, with the only consistent and sizable difference between long and short counterparts being that between /ɯ/ and /ɯː/. For this pair, the overall mean F2 value of the short vowel is higher than that of the long vowel (1596 Hz for /ɯ/ and 1419 Hz for /ɯː/) and a difference is found for every speaker for whom at least one token each of short /ɯ/ and long /ɯː/ was measured (all speakers except M5). However, the difference is not statistically significant, perhaps because of the small number of tokens involved, and in the plots and tables which follow long and short /ɯ/ are not separated.

The measured values of the first two formants of each vowel token for the five male speakers are plotted together in Figure 5. In this and subsequent figures of the same type, the origin for both scales is in the upper right of the figure. The distances along the formant scales are denominated in Hz but proportional to intervals in Barks. Each vowel symbol on the chart represents one or more tokens of the vowel indicated by the symbol. An ellipse is drawn around each cluster of points representing a single vowel type. The ellipses have a radius of two standard deviations along the axes of the first two principal components of the distribution.

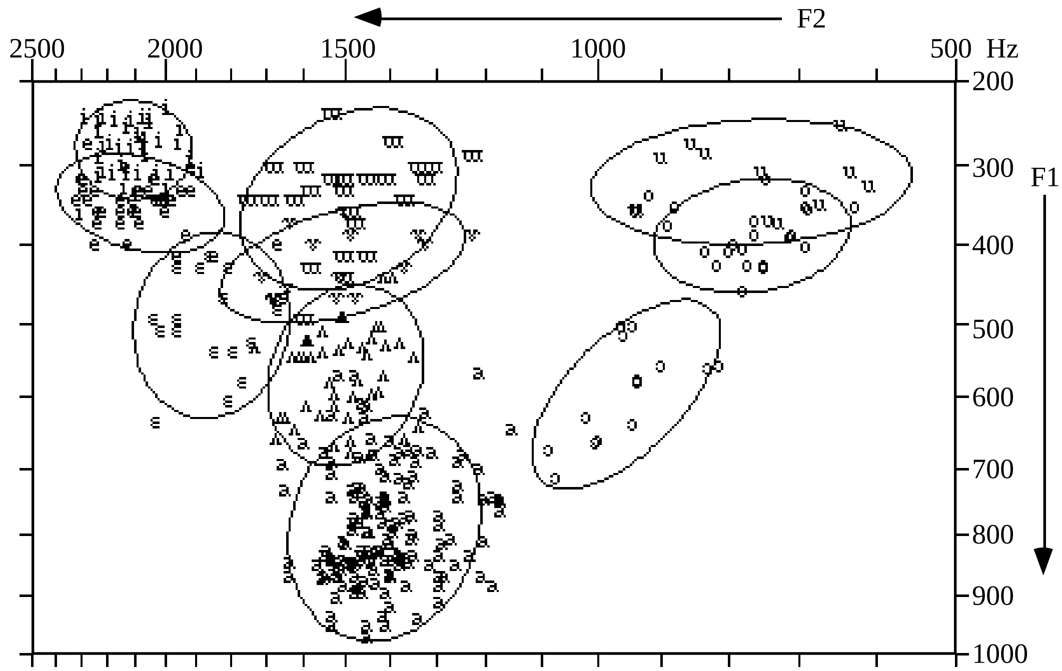

Figure 5. Values of the first two formants in oral vowels, male speakers, all tokens.

The scatter of points for each vowel in Figure 5 includes some variation attributable to differences in consonantal context, especially for the /a/ vowels where a wider range of contexts was available, and some attributable to cross-speaker differences. To show a clearer picture of the typical vowel positions, the mean formant values for each vowel for each individual speaker were calculated and each mean plotted as a single point. These data for the male speakers are shown in Figure 6. The residual scatter in this figure is mainly attributable to inter-speaker differences. The large scatter seen for /a/ in Figure 5 is much reduced in Figure 6 and is thus shown to be due largely to the contextual effects.

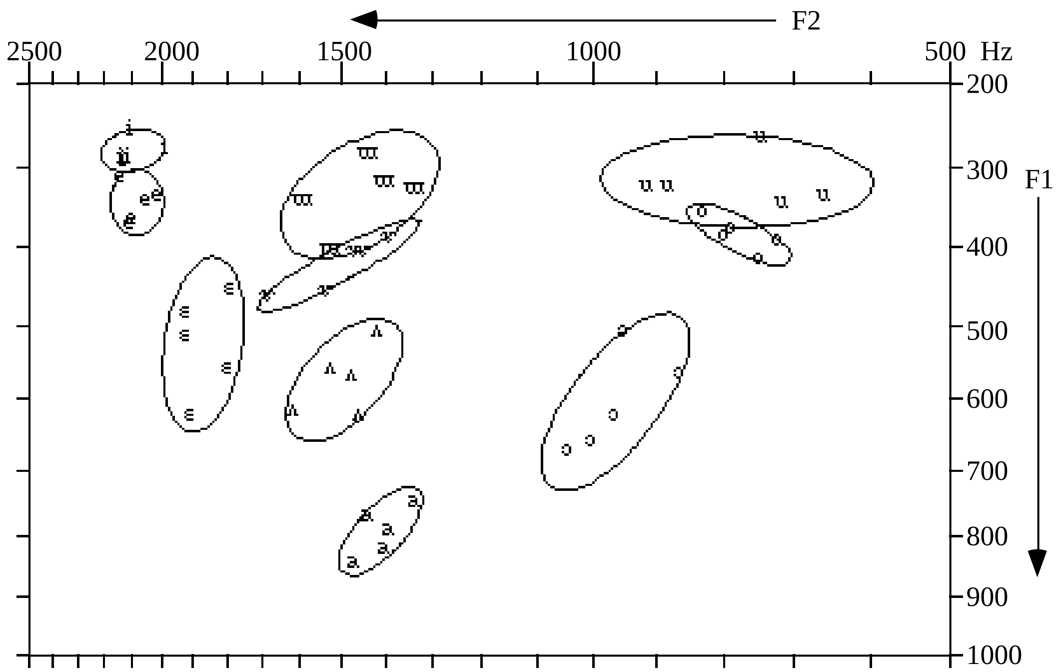

Figure 6. Values of the first two formants in oral vowels, male speakers, mean value for each vowel.

Similar displays of the distribution of the vowels of the female speakers in a two-formant space are shown in Figures 7 and 8. Figure 7 shows the scatter of points from all the tokens measured, and Figure 8 plots just the means for each vowel for each speaker.

Figures 5-8 show substantial differences in the degree to which vowels are separated in the two-formant space. The three high vowels and the three higher mid vowels form three pairs whose members are not far apart, especially in the case of the pair /ɯ/ and /ɤ/ for the female speakers. The three lower mid vowels on the other hand are more clearly separated from their higher mid counterparts. The relative crowding of the upper part of the acoustic vowel space may be what permits the quite large inter-speaker differences in the first formants of these lower mid vowels seen in figures 6 and 8; since the higher-mid vowels are relatively high, speakers are free to choose a realization of the lower mid vowels anywhere within a rather broad region.

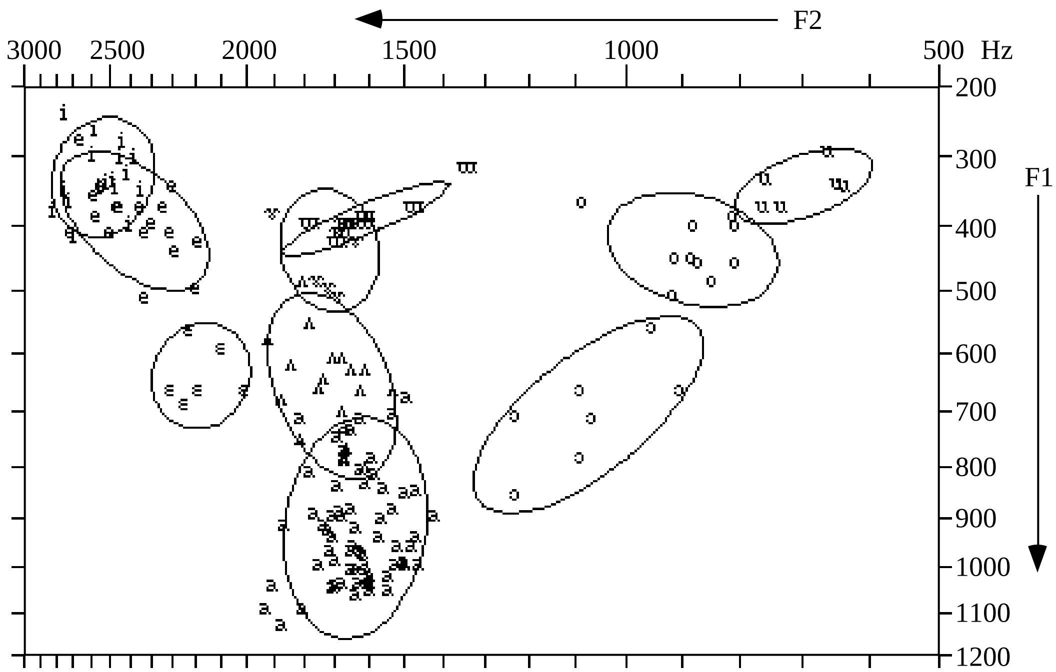

Figure 7. Values of the first two formants in oral vowels, female speakers, all tokens.

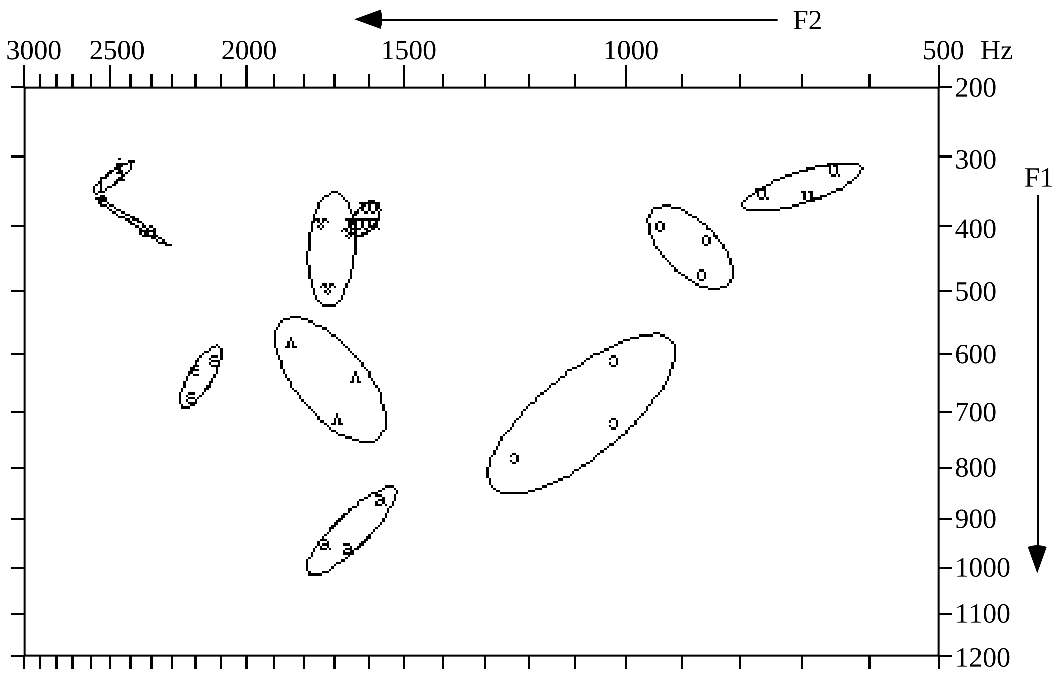

Figure 8: Oral vowels, female speakers, mean value for each vowel.

All the speaker means for the first three formants of each vowel, as well as the number of tokens measured, are shown in Appendix 2.

As noted in connection with Figure 5, the high and higher mid vowels are not greatly separated in the F1/F2 space, but the front pair i/e have quite distinct third formant values. The F3 for /i/ is about 300 Hz higher than that for /e/, the mean of the speaker means for male speakers being 2948 for /i/ and 2646 for /e/ (significant at the .0015 level in an unpaired t-test). This formant presumably assists in differentiating these two vowels, but for the other two vowel pairs concerned there is no reliable difference in F3 in our data.

One of the main features of interest in the Paicî vowel system are the acoustically ‘interior’ vowels transcribed /ɯ, ɤ, ʌ/. For the majority of speakers studied the F2 values of these vowels suggest that these may not be truly back in their articulation — lower second formants would be expected — but might better be considered as central vowels. For the sake of consistency, however, we will continue to use Rivierre's transcription.

4.4. Nasalized Vowels

The values of the first three formants of the nasalized vowels were also estimated, using the same procedure outlined above for the oral vowels. We are aware of the problems involved in modeling nasalized vowels with an LPC analysis, but note that our method involves cross-checking with FFT spectra. We feel that a reasonable characterization of some of the most important acoustic properties of these vowels can be obtained in this way. The speaker-means of the resulting measurements are included in Appendix 2. Due to the absence of long vs. short pairs for certain nasalized vowels and an overall paucity of tokens containing nasalized vowels, values for long and short nasalized vowels are collapsed together. Differences in the nasalized vowels dependent on length are discussed later in this section.

Means of the first two formants for each of the nasalized vowels for the male speakers are plotted in figure 9, and for each of the female speakers in Figure 10. The scales are the same as in Figures 6 and 8.

The nasalized high vowels /ĩ, ũ/ have similar values of the first two formants to their oral counterparts, suggesting that they do not differ much from them with respect to their oral articulations. On the other hand, /ɯ̃/ has a substantially lower second formant than /ɯ/ among the female speakers. It is possible that this vowels is produced by some speakers with a more retracted position than its oral counterpart. But as only a single token of this vowel, and that following a labialized consonant, was available, any interpretation should be cautious.

Among the mid nasalized vowels the front vowel /ẽ/ has a substantially higher F1 than the oral vowel /e/, and compares more closely with the formants of /ɛ/. For the male speakers at least the first formant of /õ/ is substantially higher than that of /ẽ/. Compared to its oral counterpart /o/, this vowel has a higher F1 and F2, placing it more in the perceptual center of the vowel space. Formant values for /ʌ/ and /ʌ̃/ do not differ in a consistent direction, though there is a tendency for F1 to be slightly higher for the oral vowel /ʌ/. The nasalized low vowel /ã/ has a considerably lower F1 than its oral counterpart /a/, reducing the distance between this vowel and /ʌ̃/.

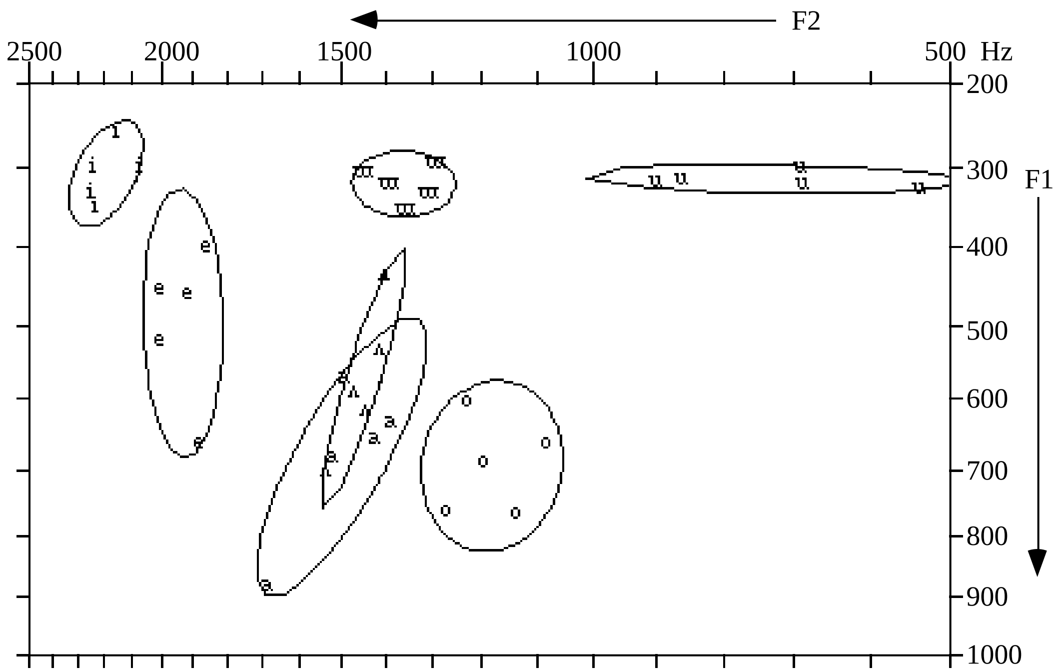

Figure 9. Mean values of first two formants of nasalized vowels for male speakers.

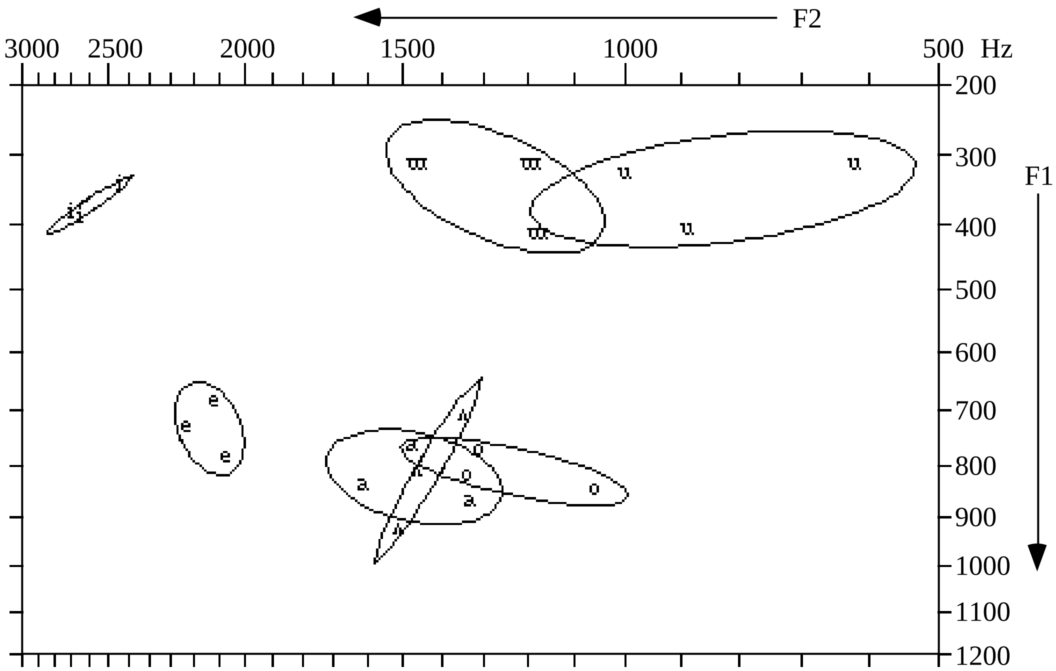

Figure 10. Mean values of first two formants of nasalized vowels for female speakers.

In summary, the high nasalized vowels remain high in the acoustic space, the mid nasalized vowels appear to be acoustically closer to the lower mid vowels in the oral set rather than the higher mid vowels, and the low nasalized vowel is acoustically raised.

5. Discussion

It is a common cross-linguistic observation that nasalized vowels tend to be more centralized than their oral counterparts. Wright (1980) suggests that the addition of a nasal formant in the vicinity of F1 triggers a perceptual shift in the height of vowels, as the 'center of gravity' of low-frequency energy in the vowel is shifted. Because the first nasal formant has very similar values across all vowel qualities (Maeda 1993), its presence has different perceptual effects on vowels of different heights. For non-low vowels, the first nasal formant is typically higher than the first oral formant, resulting in an upward shift in the central focus of low frequency energy. The perceptual consequence of this shift is that non-low nasalized vowels may be perceived as lower than their oral counterparts; over time, this may lead to an articulatory adjustment, as in the history of French. The apparent low position of the nasalized mid vowels in Paicĩ may be a reflection of this tendency.

In low vowels, by contrast, the first nasal formant is lower than the first oral formant; the result is a downward shift in energy when the vowel is nasalized and, hence, creates the perception of a raised vowel. Our measurements of /ã/ in Paicĩ as having a lowered F1 probably reflect just the acoustic effect anticipated, and not any articulatory adjustment.

The effect of nasalization on pereceived vowel backness is less well understood with effects observed in natural language at odds with those predicted by perception models (Beddor 1993: 182). Beddor suggests that the nasalization of back vowels might lead to “perceptual retraction” due to a decrease in the separation of F1 and F2 resulting from the addition of a nasal formant between the first two oral formants. This effect of nasalization should not be as great in front vowels due to the relatively great distance between the first two oral formants. Wright, however, found that the greatest effect on perceived backness occurred in the front vowel pair /i/ and /ĩ/. He also found that nasalization created the percept of fronting in /õ/, which is in line with the Paicĩ findings.

The effect of nasalization on natural speech vowels other than peripheral vowels has not been examined in the literature, due to the paucity of languages with ‘interior’ nasalized vowels. Our data contain few tokens of these vowels, but we may make some interpretative suggestions. Recall that in Paicĩ, nasalization did not appear to have a consistent effect on /ʌ/, while F2 of /ɯ̃/ was found to be often lower than F2 of /ɯ/. It may be that /ʌ̃/ has an F2 close to that of a nasal formant and hence its central frequency is not modified, but it could also be that articulatory efforts are made to keep values for /ʌ̃/ similar to those of its oral counterpart because of the proximity of nasalized /ʌ̃/ to both /õ/ and /ã/. Were /ʌ̃/ shifted either downward or backward in the vowel space, it would perceptually merge with these neighboring vowels. /ɯ̃/ on the other hand, has more room for backing, since there are no particularly nearby vowels with which it would merge. The nearest back vowel, /u/ is substantially more backed than /ɯ̃/.

6. Summary

In summary, we have examined a number of phonetic features relevant to the Paicî vowel system. The three phonemic tones of are distinguished primarily in terms of fundamental frequency level rather than contour shape, with the difference between high tone and both of the other tones being more robust than the distinction between the mid and low tones. The higher mid vowels overlap more in quality with the high vowels than with the lower mid vowels. Second formant values for the non-front unrounded vowels are suggestive of a central rather than back tongue position. Nasalization tends to trigger lowering in the mid vowels, particularly /ẽ/ and /õ/, and raising in the low vowel resulting in a compacted vowel space for the non-high nasalized vowels relative to their oral counterparts.

References

Beddor, Patrice S. 1993. Perception of nasal vowels. In *Nasals, Nasalization and the Velum*, ed. by M. Huffman and R. Krakow. New York: Academic Press. 171-196.

Gordon, Matthew and Ian Maddieson. 1999. The phonetics of Ndumbea. *Oceanic Linguistics* 38: 66-90.

Leenhardt, Maurice. 1946. *Langues et Dialectes de L'Austro-Mélanésie.* Paris: Institut d'Ethnologie, Musée de l'Homme.

Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

Maddieson, Ian. 1991. Testing the universality of phonological generalizations with a phonetically-specified segment database: results and limitations. *Phonetica* 48: 193-206.

Maddieson, Ian and Victoria B. Anderson. 1995. The phonetics of Iaai. *UCLA Working Papers in Phonetics* 87: 163-182.

Maeda, Shinji. 1993. Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In *Nasals, Nasalization and the Velum*, ed. by M. Huffman and R. Krakow. New York: Academic Press. 147-170.

Rivierre, Jean-Claude. 1974. Tons et segments du discours en langue paici (Nouvelle Calédonie). *Bulletin de la Société de Linguistique de Paris* 69: 325-340.

Rivierre, Jean-Claude. 1978. Accents, tons et inversion tonale en Nouvelle Calédonie. *Bulletin de la Société de Linguistique de Paris* 73: 415-443.

Rivierre, Jean-Claude. 1983. *Dictionnaire Paicî-Français*. SELAF, Paris.

Wright, James T. 1986. The behavior of nasalized vowels in the perceptual vowel space. In *Experimental Phonology*, ed. by J. J. Ohala and J. J. Jaeger. Orlando: Academic Press, 45-67.

<u>Appendix 1.</u> Wordlist used for analysis of vowel formants
(´ = high tone,¯=mid tone, `=low tone)

i	pī	crab, fish eggs
	mbīmbīī	wet, soaked
	kīī	key
	ápí	yam stake
	ã́mbíi	(be in) mourning
	ã́ndí	traditional money
ĩ	mĩ̄	liana, tropical creeper
	mĩ̄ĩ̄	red, yellow
	mĩ́mĩ́	gnat
	kĩ́ɾĩ́	little
e	pé	ray
	mbēē	companion
	mbwēē	buttocks
	ŋgē	feast, ceremony
	ācēē	mosquito
	ãmbwé	flying fish
	mbéré	abandoned settlement
ẽ	mẽ̄ẽ̄	eyes
	pwẽ̄ẽ̄	his mouth
	kẽ̄ẽ̄/kēē	small shell
ɛ	mbɛ̄ɾɛ̄	brother
	kɛ̄	crab (species)
a	pāɾāwē	skin
	mbā	wall
	pwáa	white
	mbwátɔ̃́	short
	ndāaé	his spine
	ndā	spear
	lāací	rice
	cá	oven

ã	mwã̄tō	share of food
	mwãɾã	again
	mwã̄nĩ̄	money
	wã	clan, house
	tã̄ã̄	family
	kã̄ã̄	cry
	ãwãã	The Amoa river
	ã̄mbʌ̄ī	shark
ʌ	pʌ̄ʌ̄	foundation (of house)
	mbʌ̄	death
	mbwʌ̄	remainder
	ɾʌ	they
	kʌ̄kʌ̄	crow
	ātʌ́	pandanus basket
ʌ̃	pwʌ̃̄	tortoise
	pwʌ̃̄ŋge	your (sg.) mouth
ɤ	mbwɤ̄/ mbwē	banyan tree
	kɤ̄ɤ̄/ kɯ̃̄ɯ̃̄kɯ̃̄/ kẽ̄ẽ̄	deaf
	kɤ̄ɾɤ̄/kēɾē	spider
ɯ	mbɯ̄ɾɯ̄	war token
	mbwɯ́ɯ́	green
	ndɯ̄ɯ̄ɾɯ̄	mud
ɯ̃	mwɯ̃	drown, flood
ɔ	kɔ̄ɔ̄	egret
	kɔ̄rēe	grasshopper
	ŋgɔ̄ɔ	crab
o	kóó	cold, frost

¯ɟā	year, season
¯jāā	mountain
kākẽ̄	manatee
ŋgāandāaé	his forehead
ŋgāɾā	taro terrace
ápí	yam stake
āpwɯ̄	shoulder
ārāo	my front
ámbwá	limestone, chalk
ācēē	mosquito
āŋgā	parrot (species)
a¯ɟ̄iʌ̄	Ajië (a New Caledonian language)
ákāē	yes, ok
alala	touché!
ātʌ́	pandanus basket

	kōōnẽ	Koné (place)
	pópáa	rain
	mbōo	descend
õ	cố́ố́	look at, see
	mɔ̃̄tɔ̄	grass, “bush”
	mɔ̃̄ɔ̃̄	cold
u	mbū	1 pl incl pn (“we 2”)
	mbūɾɛ̄	seashell
	kúrú	“bougna”
ũ	pũ̄ɾũ̄	head
	mũ̄	flower

Appendix 2. Mean values of the first three formants of each vowel for each speaker. Speakers are identified by a letter indicating gender (M or F) and a number.

Speaker & vowel		**F1**	**F2**	**F3**	**Speaker & vowel**		**F1**	**F2**	**F3**
M1	n				**M5**	n			
i	13	288	2129	3008	i	6	291	2139	2743
ĩ	5	365	2240	3038	ĩ	4	309	2248	3200
e	12	364	2111	2636	e	6	340	2064	2533
ẽ	4	477	2006	2693	ẽ	2	567	1927	2623
ɛ	4	513	1930	2534	ɛ	2	558	1807	2489
a	38	789	1397	2369	a	26	820	1409	2470
ã	8	609	1492	2122	ã	4	734	1523	2116
ʌ	11	556	1526	2388	ʌ	7	621	1459	2481
ʌ̃	3	564	1407	2660	ʌ̃	1	657	1442	2294
ɤ	3	454	1544	2162	ɤ	3	386	1397	2607
ɯ	7	403	1534	2139	ɯ	2	320	918	2503
ɯ̃	3	367	1356	2032	ɯ̃	1	344	1310	2275
ɔ	4	562	870	2429	ɔ	2	672	1048	2416
o	8	390	727	2456	o	3	354	833	2436
õ	3	717	1161	2185	õ	3	742	1198	2323
u	3	335	662	2315	u	2	324	918	2395
ũ	2	338	537	2052	ũ	3	327	862	2152
M2					**F1**				
i	8	288	2152	3101	i	8	344	2658	3160
ĩ	3	346	2251	3002	ĩ	4	375	2720	3214
e	8	368	2115	2796	e	7	409	2400	3026
ẽ	2	551	2010	2726	ẽ	3	644	2180	3301
ɛ	2	620	1917	2538	ɛ	2	628	2207	3027
a	32	841	1479	2454	a	27	958	1656	2720
ã	4	961	1686	2235	ã	4	713	1615	2407
ʌ	9	615	1621	2399	ʌ	8	713	1812	2966
ʌ̃	2	757	1535	2362	ʌ̃	2	664	1488	2458
ɤ	3	460	1692	2225	ɤ	3	497	1726	2647
ɯ	6	340	1600	2241	ɯ	4	399	1641	2410
ɯ̃	1	331	1393	2317	ɯ̃	1	399	1337	2551
ɔ	3	658	1011	2259	ɔ	3	721	1021	3000
o	5	413	754	2321	o	4	476	864	2808
õ	3	710	1082	2276	õ	2	721	1456	2386
u	3	343	719	2363	u	2	357	690	2099
ũ	1	331	689	2220	ũ	1	303	813	2068

M3					**F2**				
i	7	279	1998	3022	i	6	325	2544	3409
ĩ	7	326	2079	3065	ĩ	5	371	2761	3370
e	9	334	2018	2523	e	6	363	2643	3271
ẽ	3	419	1857	2669	ẽ	3	735	2141	3066
ɛ	7	453	1796	2421	ɛ	2	676	2221	3138
a	29	746	1342	2245	a	16	951	1732	3063
ã	4	678	1387	2288	ã	3	781	1738	2636
ʌ	10	507	1418	2206	ʌ	5	583	1838	3090
ʌ̃	1	458	1396	2347	ʌ̃	1	757	1597	3076
ɤ	3	404	1477	2093	ɤ	2	782	1662	2089
ɯ	6	319	1406	2149	ɯ	3	372	1595	2556
ɯ̃	1̃	303	1296	1834	ɯ̃	1	303	1599	2275
ɔ	3	508	955	2094	ɔ	2	782	1228	3179
o	3	376	792	2175	o	3	402	937	2510
õ	3	823	1271	1848	õ	3	765	1483	2437
u	2	322	888	2143	u	2	320	649	2347
ũ	4	334	897	1976	ũ	1	413	937	2298
M4					**F3**				
i	6	256	2114	2867	i	6	316	2561	3111
ĩ	5	263	2162	2960	ĩ	6	331	2544	3027
e	6	312	2149	2742	e	6	408	2432	2887
ẽ	3	483	1914	2694	ẽ	1	684	2285	2768
ɛ	2	482	1931	2710	ɛ	2	614	2124	2737
a	26	767	1447	2430	a	24	860	1567	2523
ã	4	704	1423	2476	ã	4	807	1478	2169
ʌ	7	566	1475	2460	ʌ	5	641	1630	2779
ʌ̃	1	629	1466	2551	ʌ̃	1	855	1640	2591
ɤ	3	404	1454	2194	ɤ	3	398	1748	2716
ɯ	4	283	1444	2251	ɯ	3	399	1595	2459
ɯ̃	1	317	1448	2303	ɯ̃	1	303	1351	2220
ɔ	2	621	970	2315	ɔ	2	613	1023	2669
o	4	383	801	2488	o	3	421	853	2921
õ	3	828	1138	2204	õ	3	789	1226	2502
u	2	264	750	2384	u	2	354	762	2364
ũ	2	309	692	2287	ũ	2	275	656	2382