ACOUSTIC PROPERTIES OF THE KAGAYANEN VOWEL SPACE

Kenneth S. Olson[1] and Jeff Mielke[2]

SIL International & University of North Dakota[1], University of Ottawa[2]
Ken_Olson@sil.org, jmielke@uottawa.ca

ABSTRACT

We present a preliminary study of the acoustic properties of the Kagayanen vowel space. We find that /ə/ has an F1 value similar to /i/ and /u/ and hence should be classified as a high vowel. The vowel /i/ has a reduced F2 value in closed syllables. For /u/, both F1 and F2 increase in word-final open syllables.

Keywords: speech production, acoustic phonetics, vowel space, Austronesian, Kayaganen.

1. INTRODUCTION

This paper is a preliminary assessment of the acoustic properties of the vowel space of Kagayanen (ISO 639–3 code cgc [4]), an Austronesian language spoken in the Philippines on Cagayan Island in the Sulu Sea and around the town of Puerto Princessa on Palawan Island.

Kagayanen has a four-vowel system that is rare in the world's languages [13] but common in Philippine languages [16]. It consists of three peripheral vowels [2] /i, u, a/ and an interior vowel /ə/ (written in the orthography as <e>) commonly referred to as the 'pepet' vowel in Philippine linguistics [1].

Harmon [5] considers /ə/ to be lower in the vowel space than the high vowels /i/ and /u/, whereas MacGregor [12] considers all three vowels to be high. In addition, both authors describe allophonic variation of the high vowels, presumably based on their auditory impressions. They state that the high front vowel /i/ becomes [ɪ] or [ɛ] before a consonant cluster and in unstressed syllables, and is [i] ~ [e] elsewhere. The high back round vowel /u/ becomes [ʊ] before a consonant cluster and in unstressed syllables, and it becomes [o] when it occurs in the final syllable of a word.

In this paper, we examine the formant structure of the Kagayanen vowels in order to clarify the height of /ə/ and corroborate the putative allophonic behavior of /i/ and /u/.

2. GENERAL PROCEDURES

2.1. Subject

A 27-year old female native speaker of Kagayanen participated in this experiment. The subject lived in Puerto Princessa, Palawan until she moved to Montana (USA) at age 22. Both of her parents are from Cagayan Island. She has completed some college. Besides Kagayanen, she also speaks English, Tagalog, and Ilongo.

We were unable to test a larger number of subjects due to logistical constraints. Testing more subjects would make our data more indicative of the larger Kagayanen-speaking population.

2.2. Procedures

The data were collected in an office familiar to the subject. The subject was seated in a high-backed office chair, facing the prompts on sheets of paper placed at eye level approximately one meter away.

Audio recordings of the subject's voice were created using an Audio-Technica PRO 49Q condenser microphone mounted close to the subject's mouth, through a single channel of a Symetrix 302 dual microphone preamplifier. The audio signals were recorded onto a Sony GV-1000 NTSC MiniDV Recorder and captured using Final Cut Express on a Mac G3 laptop.

The subject was prompted by one of the experimenters to read three repetitions of target words in the carrier phrase [amˈbaðən no ___ ˈisab] 'Say ___ again' [9]. The experimenters took turns prompting the subject.

Other stimuli were interspersed with these to address a variety of questions, and these sets served as distractors for each other.

2.3.Analysis

Since Peterson and Barney [15], vowels have typically been plotted according to their first two formants, F1 and F2. Plotting F1 vs. F2–F1 [6] or F1 vs. F2ˈ (where F2ˈ is a weighted average of F2 and F3) [11, 3] is commonly done, but we did not pursue these because plotting F1 vs. F2 was sufficient for distinguishing the vowels.

Formant frequencies were measured using the following criteria. First, we visually inspected a wide-band spectrogram of each token to verify that there was a steady state period of the vowel. We then visually identified the midpoint of the steady state. The window of analysis was centered on this midpoint. The formant measurements were made using the LPC analysis feature in PRAAT (version 4.4.16) employing its default parameters, which are appropriate for a female speaker. Because LPC calculations of F1 can potentially be influenced by a high f0 (typically about 200 Hz for our speaker), we verified the formant measurements by visual inspection on a wide-band spectrogram.

3. HEIGHT OF SCHWA

We first set out to identify the height of /ə/ with respect to the high vowels /i/ and /u/. We recorded nine tokens of each vowel (three tokens of three words) in a stressed word-initial open syllable immediately following either a labial or alveolar stop ˈCV.CV(C) [8]. (Stress usually occurs on the penultimate syllable in Kagayanen.) One word in our corpus lacked an initial consonant. We avoided adjacent nasal consonants in order to minimize the influences of nasalization [9]. All words in our data were disyllabic. The mean values of F1 and F2 are shown in Table 1.

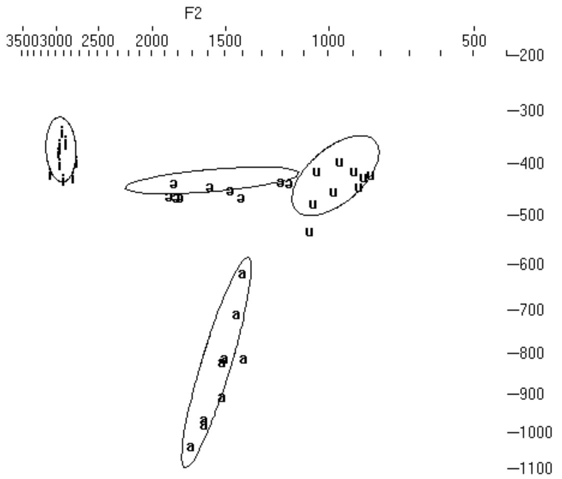
Table 1: Mean values of F1 and F2 for each vowel. Units are Hertz.

vowel	F1	F2
i	373	2930
ə	432	1598
u	423	972
a	826	1552

Figure 1 shows a plot of F1 vs. F2 created using the Windows version of the UCLA PlotFormants program (version 4.0). A vowel symbol is given for each individual token. The axes are marked in Hertz, but scaled on the Bark scale, which reflects the ear's sensitivity to differences in pitch [7, 17]. The ellipses are centered on the mean for each vowel and have radii of two standard deviations.

There is reasonable separation between the vowels, indicating that the values of F1 and F2 are sufficient acoustic properties for distinguishing the vowels. The vowel /i/ has a slightly lower F1 value than /u/, which is a common crosslinguistic tendency. A comparison of F1 means for /i/, /ə/, and /u/ using ANOVA provided strong evidence of difference [$F(2,24) = 9.40$; $p < 0.001$]. The vowels /ə/ and /u/ have comparable values of F1, both slightly greater than /i/, but post-hoc comparisons showed them not to be significantly different from each other (one-tailed t-test, $t(10) = 0.626$, $p = 0.27$). This suggests that if /u/ is categorized as a high vowel, then /ə/ should be as well, which supports MacGregor's interpretation of the vowel space. This also suggests that the central vowel would be better transcribed as /ɨ/.

Figure 1: Kagayanen vowels in stressed word-initial open syllables following labial or coronal stops.



4. ALLOPHONES OF /i/

Harmon [5] and MacGregor [12] both state that the high front vowel /i/ becomes [ɪ] or [ɛ] before a consonant cluster and in unstressed syllables, and is [i] ~ [e] elsewhere. To test this, we recorded nine tokens of /i/ in a stressed word-initial closed syllable ˈCiC.CV(C) (three tokens of three words), and six tokens of /i/ in an unstressed word-initial open syllable Ci.ˈCV(C) (three tokens of two words). We then compared these results with the default values from section 3. The means values of F1 and F2 are shown in Table 2.

Table 2: Mean values of F1 and F2 for default, closed syllable, and unstressed values of /i/. Units are Hertz.

environment	F1	F2
default	373	2930
closed σ	383	2689
unstressed σ	328	2971

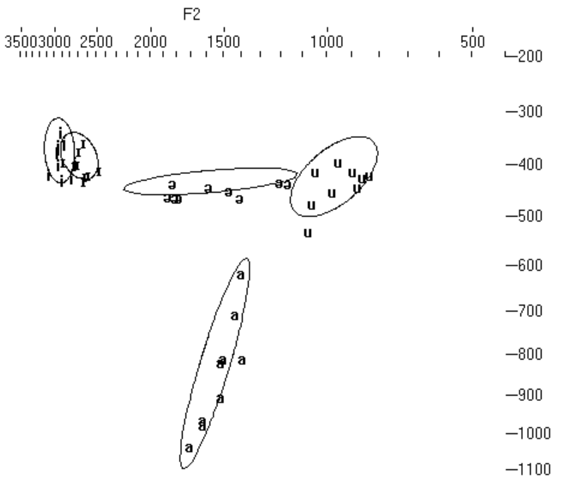
We did not observe any tokens in which the phonetic value of /i/ was in the [e] or [ɛ] range in the vowel space. One source of evidence for this is that every measurement of F1 for /i/ in our data was lower in value than the means of F1 for /ə/ and /u/.

A comparison of F1 means for the three environments (default, closed syllable, and unstressed syllable) using ANOVA provided evidence of difference [$F(2,21) = 9.05$; $p < 0.002$]. The one variable showing significant movement was a decrease in F1 in an unstressed syllable. However, this was in the opposite direction to the predicted result (which would be an increase in F1). This may have been due to the limited amount of data (only six tokens for the unstressed environment).

A comparison of F2 means for the three environments using ANOVA also provided evidence of difference [$F(2,21) = 17.59$; $p = 0.000$]. Post-hoc comparisons showed a decrease in F2 in a closed syllable. Compared with the default case, this decrease was very highly significant (one-tailed t-test, $t(16) = 4.69$, $p = 0.000$), taking into account the Bonferroni correction.

Figure 2 shows a plot of F1 versus F2 for the vowel /i/ in stressed word-initial closed syllables superimposed on the vowels from Figure 1. For convenience, we have transcribed this allophone as [ɪ].

Figure 2: The vowel /i/ in stressed word-initial closed syllables (transcribed as [ɪ]), superimposed on the vowels from Figure 1.



5. ALLOPHONES OF /u/

In this section, we examine the putative allophones of /u/. Harmon [5] and MacGregor [12] both state that the high back vowel /u/ becomes [ʊ] before a consonant cluster and in unstressed syllables, and is /u/ elsewhere. To test these claims, we recorded nine tokens of /u/ in a stressed word-initial closed syllable ˈCuC.CV(C) (three tokens of three words), and nine tokens of /u/ in an unstressed word-initial open syllable Cu.ˈCV(C) (three tokens of three words). In addition, both Harmon and MacGregor state that /u/ is realized as [o] in the final syllable of a word. To test this, we recorded nine tokens of /u/ in an unstressed word-final open syllable ˈCV.Cu (three tokens of three words). We then compared these results with the default values from section 3. The mean values of F1 and F2 for all of these environments are shown in Table 3.

Table 3: Mean values of F1 and F2 for default, closed syllable, unstressed word-initial, and unstressed word-final values of /u/. Units are Hertz.

environment	F1	F2
default	423	972
closed σ	442	944
unstressed word-initial σ	426	926
unstressed word-final σ	494	1130

Comparisons of the F1 means and of the F2 means for the four environments (default, closed syllable, unstressed word-initial syllable, unstressed word-final syllable) using ANOVA provided strong evidence of difference [F1: $F(3,32) = 9.01$; $p = 0.000$, F2: $F(3,32) = 5.13$; $p < 0.006$]. However, the only environment in which the differences are significant is the unstressed word-final environment, as shown by post-hoc one-tailed t-tests [F1: $t(16) = -3.69$, $p < 0.001$, highly significant with Bonferroni correction; F2: $t(16) = -3.43$, $p < 0.002$, significant with Bonferroni correction].

Figure 3: The vowel /u/ in an unstressed word-final open syllable (transcribed as [o]), superimposed on the vowels from Figure 1.

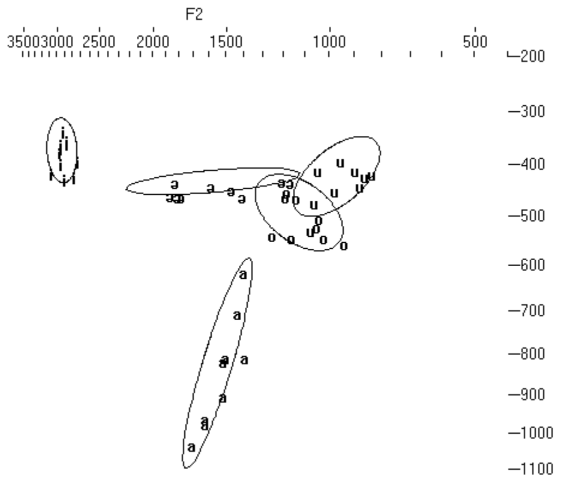


Figure 3 shows a plot of F1 versus F2 for the vowel /u/ in unstressed word-final open syllables superimposed on the vowels from Figure 1. (Following Harmon and MacGregor, we transcribe this allophone as [o] in the figure.) Crosslinguistically, [o] typically has a lower F2 value than [u]. The fact that both F1 and F2 increase in this case suggests that this allophone should be transcribed as [ʊ] rather than [o].

6. DISCUSSION

We were able to corroborate some but not all of the previous claims concerning the allophonic behavior of Kagayanen vowels. We found that a closed syllable only affects the quality of /i/ (and then only in the F2 dimension), while lack of stress by itself does not appear to be an independent indicator of allophonic movement for any sound. Having said that, lack of stress in conjunction with position in the word does affect the formant values for /u/.

Additional research could provide further light on these questions. Our study covered only one speaker; a larger number would improve the odds that the results accurately reflect the speech community at large. Ladefoged [9] suggests testing a half-dozen speakers of each sex. In addition, measuring more tokens of each target vowel in each environment (perhaps thirty each) might allow the statistical measures to tease out additional evidence for allophonic behavior.

Other follow-on studies would be appropriate. For example, both pitch and duration have previously been correlated with vowel height [14, 10], so studies of pitch and duration may be pertinent to this study. Also, in the orthography of Kagayanen, word-final /u/ is in fact written as <o>. It would be a natural follow-on study to examine how or if the orthographic representation influences the production of the sound in this environment, particularly looking at possible variation between literate and non-literate speakers.

7. ACKNOWLEDGMENTS

We wish to thank to Fraser Bennett, Steve Parker, Steve Walter, and an anonymous reviewer for helpful comments. Thanks to SIL Academic Affairs for funding, to Diana Archangeli and The Arizona Phonological Imaging Laboratory for use of equipment (James S. McDonnell Foundation grant #220020045 BBMB to Diana Archangeli and Bob Kennedy), and ERA Realty in Great Falls, Montana, for use of their office space. All errors are our own responsibility.

8. REFERENCES

[1] Conant, C.E. 1912. The pepet law in Philippine languages. *Anthropos* 7, 920–947.

[2] Crothers, J. 1978. Typology and universals of vowel systems. In: Greenberg, J. (ed), *Universals of Human Language, 2: Phonology*. Stanford, Calif.: Stanford University Press, 93–152.

[3] Fant, G. 1973. *Speech Sounds and Features*. Cambridge, Mass.: The MIT Press.

[4] Gordon, R. (ed) 2005. *Ethnologue: Languages of the World*, 15th edition. Dallas: SIL.

[5] Harmon, C. 1977. Kagayanen and the Manobo subgroup of Philippine languages. University of Hawaii dissertation.

[6] Ladefoged, P. 1982. *A Course in Phonetics*, 2nd edition. New York: Harcourt, Brace, Jovanovich.

[7] Ladefoged, P. 1996. *Elements of Acoustic Phonetics*, 2nd edition. Chicago: University of Chicago Press.

[8] Ladefoged, P. 1997. Instrumental techniques for linguistic phonetic fieldwork. In: Hardcastle, W., Laver, J. (eds), *The Handbook of Phonetic Sciences*. Oxford: Blackwell, 137–166.

[9] Ladefoged, P. 2003. *Phonetic Data Analysis*. Oxford: Blackwell.

[10] Lehiste, I. 1970. *Suprasegmentals*. Cambridge, Mass.: The MIT Press.

[11] Liljencrants, J., Lindblom, B. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48, 839–862.

[12] MacGregor, L. 1995. Kagayanen. In: Tryon, D. (ed), *Comparative Austronesian Dictionary: An Introduction to Austronesian Studies, Part 1: Fascicle 1*. (Trends in Linguistics. Documentation, 10.) Berlin: Mouton de Gruyter, 363–368.

[13] Maddieson, I. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

[14] Ohala, J., Eukel, B. 1987. Explaining the intrinsic pitch of vowels. In: Channan, R., Shockey, L. (eds), *In Honor of Ilse Lehiste*. (Netherlands Phonetic Archives, 6.) Dordrecht: Foris, 207–215.

[15] Peterson, G.E., Barney, H.L. 1952. Control methods used in a study of the vowels. *J. Acoust. Soc. Am.* 24, 175–184.

[16] Reid, L. 1973. Diachronic typology of Philippine vowel systems. In: Sebeok, T. (ed), *Current Trends in Linguistics 11: Diachronic, Areal, and Typological Linguistics*. The Hague: Mouton, 485–506.

[17] Zwicker, E., Terhardt, E. 1980. Analytical expressions for critical-band rate and critical bandwidth as a function of frequency. *J. Acoust. Soc. Am.* 68, 1523–1525.