# VOWEL ACOUSTICS OF NUNGON, PAPUA NEW GUINEA

Hannah Sarvasy[1,2], Jaydene Elvin[2,3], Weicong Li[1,2], Paola Escudero[1,2]

[1]MARCS Institute, Western Sydney University [2]ARC Centre of Excellence for the Dynamics of Language
[3]Department of Linguistics, California State University, Fresno
H.Sarvasy@westernsydney.edu.au

## ABSTRACT

Most reference grammars relegate vowel acoustics to minor sections, at best. Likewise, a corpus assembled to support grammatical description may lack high-quality recordings targeting vowels in all environments. Recent studies have applied multi-point acoustic-phonetic analyses for under-described languages. But these studies use corpora that were purpose-built for acoustic analysis. The present study applies state-of-the-art acoustic analysis techniques to a more general language documentation corpus. We present a preliminary comprehensive acoustic description of the vowels /i e a u o ɔ/ in the Towet dialect of the Papuan language Nungon. Duration, fundamental frequency and formant measurements (F1, F2 and F3) of the six phonological vowels in word-initial, -medial, and -final contexts were obtained. Results are compared with the description of phonetics and phonology in the Nungon reference grammar. In particular, the reference grammar's claim of a mid-back vowel /o/ with exceptionally low F2 values is evaluated.

**Keywords**: Nungon, vowel, acoustics, Papuan, fieldwork

## 1. INTRODUCTION

It is relatively uncommon for grammars of under-described languages to include detailed acoustic-phonetic analyses of the vowel systems of those languages. Grammars traditionally include only impressionistic descriptions of vowel articulation. This absence of acoustic-phonetic description likely stems from the traditional lack of focus on phonetics in grammatical research [8]. Like reference grammars, the audio corpora assembled during primary linguistic fieldwork are not tailored to in-depth acoustic analyses. They often involve relatively few speakers, non-uniform numbers of tokens, and relatively poor recording quality. However, recent studies have shown that it is possible to apply advanced, multi-point acoustic analysis techniques (applied previously to English [e.g., 2]) to lesser-described languages. For example, Kashima et al. [5] showed that data collected in the field on the under-described Papuan language Nambo was of sufficient quality for the purposes of computational analyses. This resulted in the first comprehensive acoustic description of the vowel system of Nambo and helped to substantiate the vowels' phonemic status. But the data used in Kashima et al. was still collected with a primary goal of acoustic analysis, albeit in the field [5].

In this study, we present a preliminary acoustic analysis of the six phonological vowels in Nungon, /i e a u o ɔ/, spoken by 4 speakers (2 female, 2 male), using similar acoustic analysis methods to those described by Kashima et al., but applying these to an existing corpus of free narrative speech that was originally created for the purpose of grammatical analysis. Nungon (<yuw>) is a Papuan (non-Austronesian) language of the Finisterre-Huon family spoken by 1,000 people in remote hamlets of the Saruwaged Mountains, northeastern Papua New Guinea (PNG). Nungon is actually the umbrella term for four dialects of a dialect continuum, grouped together based on morphosyntax and lexicon [11]. The best-described Nungon dialect is that of Towet village. Towet Nungon is spoken in all spheres, with limited use of Tok Pisin in some church sermons and some public addresses. The first two years of schooling are in Nungon only; subsequent schooling is in Tok Pisin. Towet Nungon phonetics and phonology are described in [12], but acoustic measurements of vowels there are based on a limited number of tokens spoken by a single speaker, measured at vowel midpoint.

The description in [12] presents three intriguing issues relating to Nungon vowels. First, the Nungon vowel inventory is typologically unusual in having more phonemic back vowels than front vowels. The vowel plot in [12] is asymmetrical, with a gap in the front of the vowel space corresponding to vowels in the back. A primary aim of the current study is to produce Nungon vowel plots using more sophisticated acoustic sampling techniques.

Second, vowel length is claimed in [12] to be contrastive, but the Nungon orthography (based on the orthography created for a different dialect, Yau, [13]) does not distinguish vowel length. Thus, another aim of the present study is to evaluate the acoustic correlates of claimed phonological vowel length.

Somewhat suspect in the Nungon vowel plot in [12] is the middle back vowel /o/. This vowel is

apparently close in height in the vowel plane (approximated by F1) to the highest back vowel /u/, but is depicted as having an anomalously low F2 value, much lower than that of either /u/ or /o/. Sarvasy [12] proposes that the extra-low F2 in this vowel is due to extreme lip protrusion, which would extend the oral cavity. The veracity of this low F2 could be questioned, however, based on the relatively few tokens from the single speaker. Further, [13] suggested that the vowel /o/ was the least commonly used vowel in the closely-related language Yau. The vowel's infrequency compounds with occasional speaker uncertainty in back vowel differentiation for certain lexical items. That is, while the distinctions between back vowels are vital to differentiate /oːhi/ 'ascend!' from /ɔːhi/ 'descend!' and /mɔro/ 'large' from /mɔru/ 'fiber string,' there is some variability in judgments about back vowel distinctions in less-frequently used words such as names of certain forest plants, e.g. the shrub /gosɔndɔng/ (variant judgments /gɔsɔndɔng/, /gosondɔng/, and /gɔsondɔng/). A third aim, then, is to evaluate the phonological and acoustic status of the back vowel /o/.

## 2. METHODS

### 2.1 Speakers

Speakers are: two women aged ~36 and ~26, and two men aged ~26 and ~24. All grew up within the Towet village community and are fluent in the Towet Nungon dialect, but the older of the two women has a parent from a different dialect community, and this is reflected in her own speech, especially her frequent use of final glottal stops in place of unreleased [k] (see [12] for further details). The men are half-brothers, sons of different wives of a polygamist; they grew up in the same household. All speakers have familiarity with other Nungon dialects and limited-to-moderate knowledge of Tok Pisin. Of them, only the younger female attended school; she left in sixth grade.

### 2.2 Recordings and transcriptions

Vowel acoustics were analysed based on tokens in narratives recorded between September 2011 and March 2013. As part of the comprehensive documentation and description of Nungon grammar, a corpus of 221 recorded and transcribed texts was created. In most of these recordings, speakers delivered monologues on themes of their own choice, and were recorded in close range using the built-in twin microphones of a Zoom H4N Handy audio recorder, at a 44.1 kHz sampling rate, in WAV format. One narrative each for the two women, two long narratives for the younger man, and five short narratives for the elder, served as source material. These were digitally transcribed together with native speakers in the field in 2011-2013. The transcriptions were then searched for words that included the target vowels in word-initial, word-medial and word-final environments, and not adjacent to nasals. This last stipulation was meant to reduce coarticulation effects, and to accord with the practice in [12]. When these words were checked with the corresponding recordings, tokens with especially poor audio quality were excluded. This yielded 692 vowel tokens for analysis.

### 2.3 Acoustic analysis

For alignment at the utterance, word and phonetic level, the Munich Automatic Segmentation System (MAUS) was applied: first the transcriptions were roughly aligned at utterance level and sent to WebMAUS [6] for grapheme-to-phoneme conversion [10]; then the "language independent" mode was applied for segmentation, since this mode does not require any language training, yet still delivers reasonable results [4, 7]. Later, the segmentation by MAUS at phonetic level was manually checked and adjusted, as there were a large number of cases in which misalignment occurred.

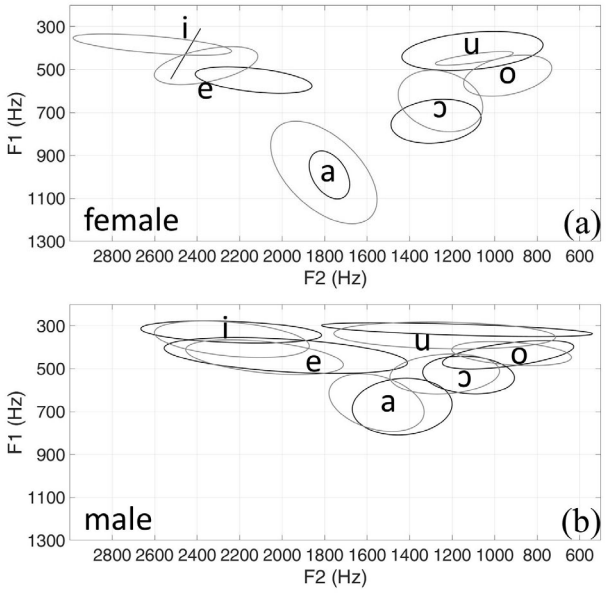
In acoustic studies, it is common for monophthongs to be measured at one static point in time (typically the midpoint), whereas diphthongs are measured at multiple time-points (usually 25%, 50% and 75% of the vowel's duration) in order to highlight their dynamic nature. However, recent studies have shown that formant trajectory measurements can also be useful for characterizing monophthongs [e.g. 2, 3, 14]. In formant analysis, using averaged values over multiple time-points could also be more reliable than only using midpoint measurements, especially for recordings with poor quality. Therefore, our acoustic measurements were extracted using the same acoustic analysis techniques as Williams & Escudero [14], which was also used in Kashima et al.'s study of Nambo [5]. First, vowel duration, fundamental frequency and formant values were extracted at 30 evenly distributed points starting from the 20% point to the 80% point of the vowel duration in Praat [1]. We then smoothed the series of formant values (or trajectories) for each vowel, using discrete cosine transforms in Matlab (DCTs, see [14] for a full description). The formant values were further averaged across the 30 time-points, then averaged across speakers and different positions. The vowel durations were also averaged across the four speakers and different positions.

## 3. RESULTS

### 3.1. Formant measurements

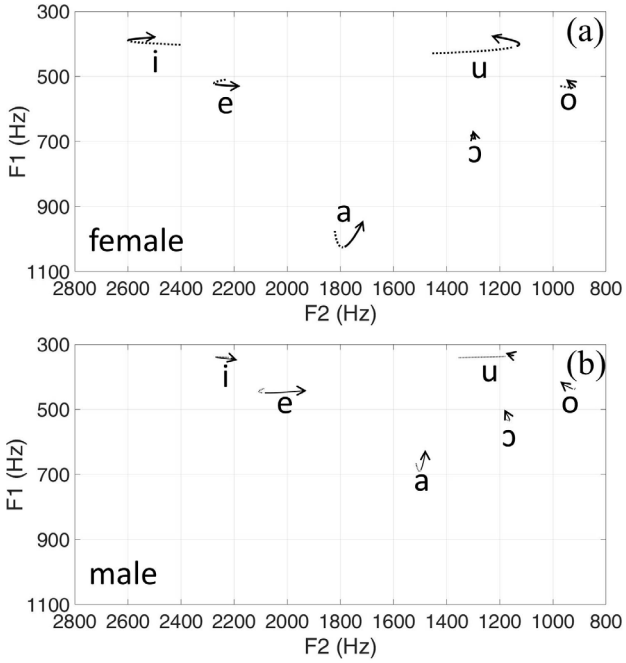
Sarvasy [12] argued for a six-vowel inventory for Towet Nungon. Indeed, our results show that the Nungon vowels form six distinct groups in terms of mean formant values. The average F1 and F2 values we obtained, shown in the vowel plots in Figure 1, are similar to those in [12]. In visual inspection of the plots, the female vowel space appears more spread out than that of the males. This difference, along with an expected difference in F1, could be due to sex-based vocal tract differences. There is also speaker variability in the formants for each group (5–94 Hz for averaged F1 and 8–224 Hz for averaged F2), but this variation is smaller than that between the two genders (50–326 Hz for averaged F1, and 9–281 Hz for averaged F2). Finally, the finding in [12] that the mid-back vowel /o/ has the lowest mean F2 value of all vowels seems to be confirmed in Figure 1, although the plots here also show a fair amount of inter-speaker variation.

**Figure 1**: Formant ellipses of the six vowels for female (a) and male (b) speakers. Different speakers are shown with grey vs. black ellipses.




In Figure 1, the typologically unusual back vowel-heavy asymmetrical array of vowels does mean a gap in the lower front parts of both plots, though this is smaller for the men. Our multi-point measurement technique enables us to further expand on the basic vowel plot in [12] with a description of formant trajectories. Figure 2 shows the average formant trajectories of the six Nungon vowels for the speakers in the present study.

**Figure 2**: Formant trajectories of the six vowels for female (a) and male (b) speakers.



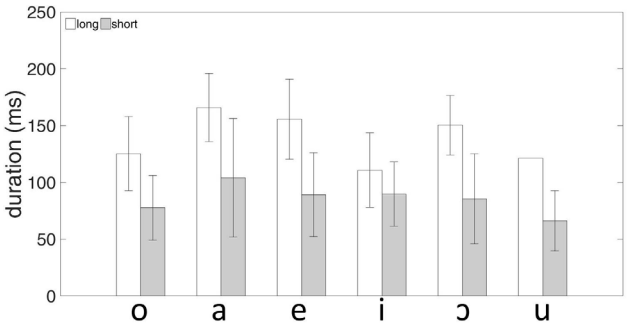

The formant trajectory plot implies that only /o/ and /ɔ/ are produced as steady state vowels. The rest of the vowels appear to have some degree of movement in either F1 or F2. Both men and women produce /a/ with an initial increase in F1, immediately followed by a decrease in F1; both groups also produce /u/ with initial decrease in F2, followed by a final increase; the final increase is more pronounced for women than for men. Female and male trajectories diverge most for the two front vowels. Women's production of /i/ and /e/ is characterized by an initial increase in F2 before a decrease in F2. Men's /i/ and /e/ also show F2 movement, but rather a unidirectional decrease.

### 3.2. Vowel duration

Sarvasy [12] argued for a phonological vowel length distinction in Nungon based on groups of lexical minimal pairs. Figure 3 confirms that the vowels considered to be phonologically long and short indeed group phonetically into separate clusters; male and female speech are combined here. Note that only one token of long /u/ served as the basis for Figure 3, hence the absence of an error bar.

We performed two-sample t-tests for each long/short vowel pairing. Durations of the phonologically long and short tokens of the vowels /o/, /a/, /e/, and /ɔ/ all differ significantly with $p$ <1e−3 ($d$ = 1.23-1.82). Long and short /u/ also differed significantly in duration ($p$ = 0.045, $d$ = 2.10). For /i/, the duration difference between phonologically long and short tokens was less significant than for the other vowels, at $p$ < 0.159, $d$ = 0.74.

**Figure 3**: Duration of long and short vowel tokens.




In all, the duration measurements and t-test results further confirm the phonetic length differences between phonologically long and short vowels. Formant values for long and short vowels were roughly the same, indicating that the distinction between long and short vowels is in duration only.

### 3.3. Back vowel distinctions

Our third target area for investigation was the Nungon back vowels. We performed two-sample t-tests for all the Nungon back vowels to evaluate the significance of differences in F1, F2, and F3 among the three back vowels. For both women and men, all three formants of /o/ and /ɔ/ differ significantly ($p < 1e{-}3$, $d$ = 1.87-15.66), and the same is the case with /ɔ/ compared with /u/ ($p < 1e{-}3$, $d$ = 2.94-9.55). For /o/ vs. /u/, F2 differed significantly (women, $p$ = 0.019, $d$ =1.58; men, $p < 1e{-}3$, $d$ = 1.99), however, the difference in F1 was much less significant for women ($p$ = 0.191, $d$ = 0.64) and not significant for men ($p$ = 0.558, $d$ = 0.12). Finally, the difference in F3 between /o/ and /u/ was significant for men ($p$ = 0.006, $d$ = 0.60), but not for women ($p$ = 0.401, $d$ = 0.36).

The two-sample t-test results, along with the vowel plots in Figures 1 and 2, indicate that the back vowel tokens (/o/, /ɔ/ and /u/) indeed form separate F1 and F2 groupings.

## 4. DISCUSSION

Overall, our study offered more sophisticated, acoustic-based support for the phonetic and phonological description of the Nungon vowel inventory in the Nungon reference grammar [12]. Measurements of F1 and F2 yielded six distinct groupings corresponding to the six phonological vowels (Figure 1). The average F1 and F2 values for the six vowels are similar to those given in [12]. Further, duration data (Figure 3) and two-sample t-tests confirmed that vowels thought to be phonologically long and short form distinct duration clusters. Finally, the contrasts among the three back vowels were confirmed with two-sample t-tests.

A new contribution here is the measurement of vowel trajectories (Figure 2). Our results indicate that most of the Nungon vowels have more formant movement than previously known. They also show sex-based differences in front vowel trajectory directions, and relative height of /a/. Given the small number of speakers in the current study, trajectory variation could be due to individual differences in vocal tract size. Alternatively, they could stem from the relatively uncontrolled nature of the environments in which vowel tokens occurred; the only environments that were excluded from the study were vowel sequences or diphthongs, e.g. /ua/ and /ai/, and adjacent nasals. Finally, the relatively poor quality of recordings could mean imprecise formant detection by Praat.

Our study further provided new insights into an idiosyncrasy suggested by the reference grammar. Sarvasy [12] described the mid-back vowel /o/ as being produced with extremely-protruded lips, more so than /u/, and took that to explain the low F2 values of /o/. Without further articulatory data, we cannot confirm this here. However, we do show here that there are significant differences in F2 between /o/ and the other two back vowels for speakers of both sexes.

## 5. CONCLUSION

In this study we used a corpus of free speech audio recordings of variable quality, created for documentation and description of the Nungon language, as the basis for multi-point vowel acoustic analysis. Results support the Nungon reference grammar's analysis of the vowel inventory as including six vowels with additional contrastive vowel length, and also confirm the grammar's mid-back vowel /o/ with lower F2 value than /u/ and /ɔ/. We further contributed new data on vowel trajectories, showing that these vowels are less steady state than originally assumed.

## ACKNOWLEDGMENTS

This study was supported by a Language Documentation grant to Sarvasy through the ARC Centre of Excellence for the Dynamics of Language, CE140100041, and MARCS Institute funds for Li. We would also like to thank Jason Peed for help with the manual segmentation and adjustment of the data.

## REFERENCES

[1] Boersma, P., Weenink, D. 2018. Praat: Doing phonetics by computer [Computer program]. Version 6.0.43, retrieved 8 September 2018 from http://www.praat.org/

[2] Elvin, J., Williams, D., Escudero, P. 2016. Dynamic acoustic properties of monophthongs and diphthongs in

Western Sydney Australian English. *J. Acoustic. Soc. Am.*, *140*(1), 576–581.
[3] Jacewicz, E., Fox, R.A., Salmons, J. 2011. Vowel change across three age groups of speakers in three regional varieties of American English. *J. Phon*, *39*(4), 683–693.
[4] Jones, C., Demuth, K., Li, W., Almeida, A. 2017. Vowels in the Barunga variety of North Australian Kriol, *Proc. Interspeech 2017* Stockholm, 219–223.
[5] Kashima, E., Williams, D., Mark Ellison, T., Schokkin, D., Escudero, P. 2016. Uncovering the acoustic vowel space of a previously undescribed language: The vowels of Nambo. *J. Acoustic. Soc. Am.*, *139*(6), EL252–EL256.
[6] Kisler, T., Reichel, U.D., Schiel, F. 2017. Multilingual processing of speech via web services, *Computer Speech & Language*, 45, 326–347.
[7] Kisler, T., Schiel, F., Sloetjes, H. 2012. Signal processing via web services: The use case WebMAUS, in *Digital Humanities Conference* Hamburg.
[8] Maddieson, I. 2002. Phonetics in the field. *Berkeley Linguistic Society Proceedings.*
[9] McElhanon, K. 1973. *Toward a Typology of the Finisterre-Huon Languages.* Canberra: Pacific Linguistics.
[10] Reichel, U.D. 2012. PermA and Balloon: Tools for string alignment and text processing, *Proc. Interspeech 2017*, 346.
[11] Sarvasy, H. 2013. Across the great divide: How birth-order terms scaled the Saruwaged Mountains, Papua New Guinea. *Anthropological Linguistics* 55:3, 234-255.
[12] Sarvasy, H. 2017. *A Grammar of Nungon: A Papuan Language of Northeast New Guinea.* Leiden: Brill.
[13] Wegmann, U. 1994. Orthography paper: Yau. Unpublished ms. Ukarumpa: Summer Institute of Linguistics.
[14] Williams, D., Escudero, P. 2014. A cross-dialectal acoustic comparison of vowels in Northern and Southern British English. *J. Acoustic. Soc. Am.*, *136*(5), 2751–2761.