Vowel Acoustics in Conversational Central Australian Aboriginal English

Yizhou Wang[a], Carmel O'Shannessy[b], Vanessa Davis[c], Rikke Bundgaard-Nielsen[a], Denise Foster,[c] Joshua Roberts[b]

[a]University of Melbourne, [b]Australian National University, [c]Tangentyere Council
wyz2014tw@gmail.com, Carmel.OShannessy@anu.edu.au, vanessa.davis@tangentyere.org.au, rikkelou@gmail.com, denise.foster@tangentyere.org.au, josh.ian.roberts@gmail.com

Abstract

This paper presents a descriptive study of the acoustic features of the monophthongs in Central Australian Aboriginal English (CAAE, Alice Springs). We extracted vowel tokens from conversational speech of six female speakers who spoke CAAE as one of their primary languages. We analysed the acoustic differences between the vowel categories using a classification analysis and a series of acoustic comparisons based on both duration and spectral information. The results suggest that the phonemic contrasts between all neighbouring monophthongs in CAAE are maintained, and that CAAE vowels display systematic allophonic variation like what has been reported for Standard Australian English.

Index Terms: CAAE, vowel, formant, acoustics, allophone

1. Introduction

Across Australia today, the majority of Aboriginal and Torres Strait Islander people speak English on a daily basis [1]. The English varieties spoken by Aboriginal people are often referred to as Australian Aboriginal English (AAE), and many of these varieties have been reported to diverge from Standard or Mainstream Australian English (SAE/MAE) in the phonetic and phonological properties of consonants and vowels. Some of these may reflect systematic differences between the varieties, and some of the differences may relate to whether an individual speaks English as a first or primary language (L1) or as a second or additional language (L2) [2]–[5]. For instance, some varieties of L2 AAE have a reduced vowel system with five vowels /ɪ ɛ ɐ ɔ ʊ/, or even three vowels /ɪ ɐ ʊ/, depending on the speaker's L1 vowel inventory, while SAE has thirteen monophthongs, including the vowels in FLEECE, KIT, SQUARE, DRESS, TRAP, NURSE, STRUT, BATH, LOT, THOUGHT, FOOT, GOOSE, and SCHWA (IPA: /iː ɪ eː e æ ɜː ɐ ɐː ɔ oː ʊ ʉː ə/) plus diphthongs [6]. In the present paper, we limit the scope to only monophthongs.

Central Australian Aboriginal English (CAAE) is a variety of English spoken by Aboriginal Australians living in Central Australia (Alice Springs/Mparntwe, Northern Territory) [7], [8]. Alice Springs is a highly multilingual community; early impressionistic reports have suggested that the phonetics of CAAE may have influence from local Indigenous Australian languages due to language contact. Among the many traditional Aboriginal languages spoken in and around Alice Springs, are Eastern and Central Arrernte, Western Arrarnta, Alyawarr, Anmatyerr, Kaytetye, Pintupi-Luritja, Pitjantjatjara, Yankuny-tjatjara, and Warlpiri.

The only previous study of the CAAE vowel system [9]—undertaken almost 20 years ago—examined the production of four multilingual speakers of English as a second language (L2); the speakers' first languages [L1s] were Eastern Arrernte, Warlpiri, or Western Desert language. The results from the study, which relied on vowels elicited in a /hVd/ context, suggested that the L2 CAAE speakers' vowel space is *phonologically* similar to SAE, as the phonemic vowels were acoustically differentiated from their neighbouring categories. No statistical tests were, however, carried out to check the significance of acoustic differences between categories.

In addition to indicating that CAAE shares the phonological system of SAE, the analysis in [9] also indicated that CAAE vowels differed *phonetically* from that of SAE speakers from South Australia. For example, the study showed that TRAP vowel /æ/ in CAAE was closer/higher in relation to the STRUT vowel /ɐ/, and the THOUGHT vowel /oː/ was lower in relation to the FOOT vowel /ʊ/ (whereas the two vowels had similar heights in SAE). Additionally, The GOOSE and NURSE vowels (/ʉː/ and /ɜː/) in CAAE seemed to be further back than SAE.

Further, since the experimental protocol in [9] relied on a /hVd/ word-reading task, it did not allow for analysis of potential allophonic variations in the realisation of the monophthongs in different phonological contexts as has been demonstrated for SAE [10]–[13]. As a consequence, it remains unclear whether this variety of AAE exhibits contextual allophony of the kind observed in SAE. At the same time, it remains unclear whether the current CAAE speakers, who use English as their primary language (L1), demonstrate similar acoustic vowel properties to the previous report based on L2 speakers.

Motivated by the scarcity and limitations in previous research, the present study provides an acoustic phonetic description of vowels in CAAE using naturalistic conversational speech data, including vowels spoken in a range of phonological contexts that have been reported to give rise to systematic allophonic variations in SAE [10]–[13]. Notably, /æ/ potentially has at least two allophones in SAE, namely TRAP and BAN (pre-nasal, tense); the /ɜː/ vowel has two allophones, NURSE and GIRL (pre-lateral). Finally, the /ʉː/ vowel has three allophones, the canonical GOOSE vowel, the TOO vowel (post-coronal) and the POOL vowel (pre-lateral).

The research questions of this paper are as follows:

(1) How does each CAAE phonemic vowel differ acoustically from other vowels?

(2) How do CAAE vowels occupy the acoustic vowel space, and whether CAAE vowels show significant differences in neighbouring pairs?

(3) Do CAAE vowels show contextual allophonic variations?

We approach the first question using a classification analysis, and for the second and third questions, we use acoustic phonetic analyses.

2. Methods

2.1. Participants

We recorded six female speakers of CAAE from Alice Springs, Northern Territory of Australia. The speakers ranged in age from 26-38 years (M = 32.7). All had learned English from childhood and reported speaking CAAE as their primary language. One woman additionally reported speaking the traditional Indigenous Australian languages of Eastern Arrernte and Alyawarr, and one also speaks the traditional Indigenous Australian languages of Anmetyerr and Pitjantjatjara regularly.

All participants were recruited by a senior English-Arrernte bilingual researcher involved in the *Little Kids Learning Languages* project (https://little-kids-learning-languages.net/), who visited community centres and residential areas in Alice Springs with another researcher from the project. All participants received a $AUD50 supermarket voucher for their participation.

The participants were recorded either in their homes (e.g., on the veranda) or at their local community centre, participating in conversational interviews of approximately 30 minutes duration, focused on everyday activities related to children and child-rearing and led by an Arrernte researcher fluent in Arrernte and CAAE. All interviews were recorded using a Sennheiser EW112 PG4-GB Portable Wireless Lapel Microphone System.

2.2. Data preparation

All recordings were transcribed and annotated in English orthography, and 1836 vowel tokens were manually labelled in the phonetics software Praat [14], using Standard Australian English (SAE) orthography [12]. We acknowledge that all orthographic transcriptions impose phonological distinctions on a dataset, but none-the-less take this approach under the assumption that SAE orthography captures the phonological distinctions of CAAE reflects a null hypothesis of sorts. Importantly, it allows us to make statistical comparisons between the formant and duration associated with each purported phonological category to reveal whether each SAE vowel contrast is also contrastive in CAAE.

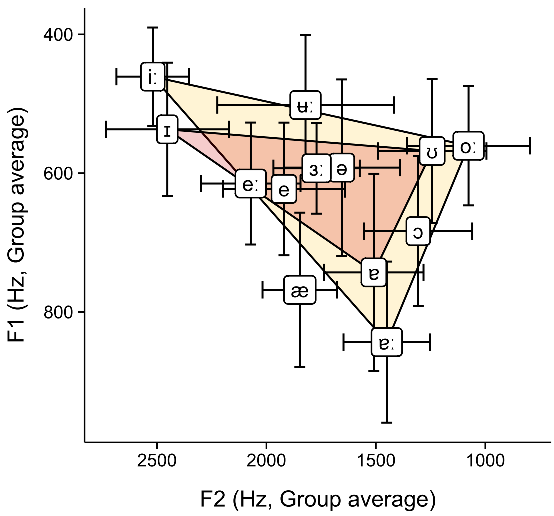

Figure 1. *Mid-point CAAE vowel qualities*

After annotation, we extracted spectral and durational measurements from each vowel, including the first four formants (F1-F4) and vowel duration using a customised *Praat* script. Formant values were estimated at the 10%, 30%, 50%, 70%, and 90% time points in the interval, to capture dynamic change over time. Erroneous data (e.g., zeroes) were excluded. For the vowel classification analysis, all formant data was included, but for the analysis of the acoustic vowel space, we included only the midpoint (50%) values which we assume represents the spectral quality of each vowel's steady state, defined by the first two formants, F1 and F2. Figure 1 presents the midpoint formant values for all CAAE phonemic vowels at the group level, where triangles indicate peripheral tense vowels (yellow) and lax vowels (red). Error bars indicate Standard Deviations. The mean duration values (in ms) are reported in Table 1.

Table 1. Mean durations (Dur.) in ms in CAAE vowels.

Vowel	Dur.	Vowel	Dur.	Vowel	Dur.
iː	106	eː	143	oː	136
ɪ	71	e	78	ʊ	67
ɐː	190	æ	108	ʉː	134
ɐ	80	ɔ	85	ɜː	153

2.3. Data analysis

We first conducted a phonemic vowel analysis (Section 3.1) to determine whether each pair of adjacent vowels also differ in CAAE, using a supervised learning algorithm (Random Forest, RM). Similar methods have been used in the other research of a similar kind, e.g., Linear Discriminant Analysis (LDA) as in [15], but RM has the additional advantage of allowing both numerical predictors (e.g., formants, duration) and categorical variables (e.g., individual speakers). Using the statistical software JASP, we fitted ten RM models based on the dataset, and then asked the models to perform predictions (70-30 split in training and testing, five-fold cross-validation).

The classification results provided by the learning algorithm were used to generate an averaged classification matrix (Figure 2). We also analysed and visualised the CAAE vowels spoken by the speakers on a F1-F2 plane (Figure 3) in Section 3.2. Following conventional methods in acoustic phonetics, we then compared the acoustic parameters of the CAAE vowels focusing on five critical contrasts, including FLEECE-KIT (/iː/-/ɪ/), BATH-STRUT (/ɐː/-/ɐ/), SQUARE-DRESS (/eː/-/e/), THOUGHT-FOOT (/oː/-/ʊ/), and TRAP-DRESS (/æ/-/e/). This approach of course assumes that all such contrasts are maintained in CAAE.

Finally, we analysed the acoustic differences between the default allophones of CAAE vowels with their contextual allophones, e.g., TRAP-BAN, NURSE-GIRL, and GOOSE-TOO-POOL. We used linear mixed-effects modelling (LMM) for all statistical inferences (random intercepts for participants and all observed lexical items).

3. Results

3.1. Classification matrix

The RM classification models were trained with all acoustic measurements (F1-F4 at 10%, 30%, 50%, 70%, and 90% time points; duration) and speaker identity as predictors, and their average performance is summarised in Figure 2, where CAAE

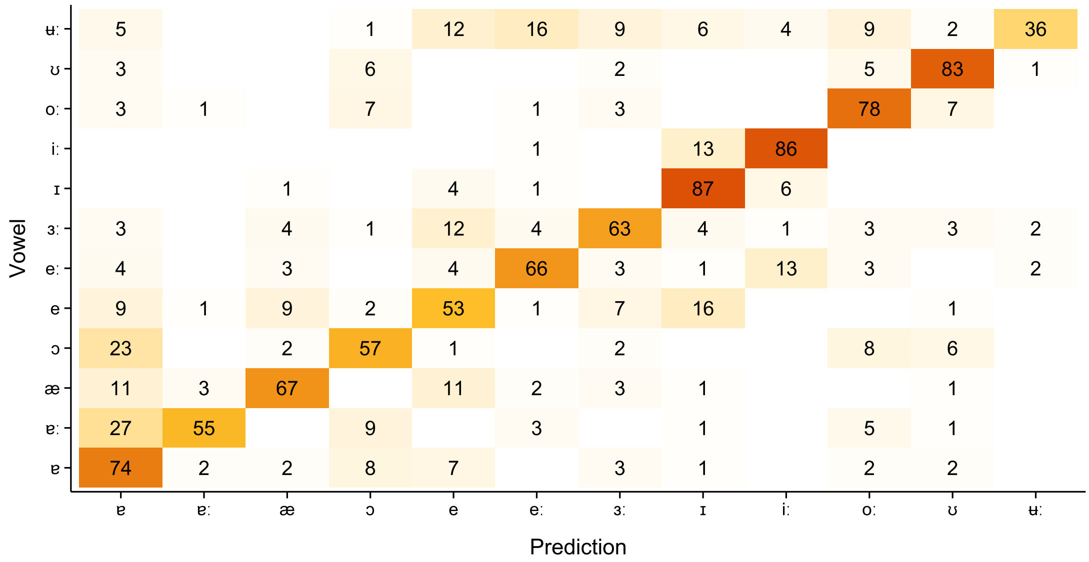

Figure 2. *Classification matrix of CAAE vowels (Random Forest models)*

vowel targets are shown on the y-axis, and the predicted categories are shown on the x-axis. Only phonemic vowels were included in this analysis, and the contextual allophones (BAN, GIRL, TOO, and POOL) were excluded.

The cells on the diagonal line display percentages of time when correct classification was performed, i.e., the predicted category matched the target category, where the chance-level is 1/12 (8.3%). Among the 12 phonemic vowels (excluding the unstressed vowel /ə/), high classification accuracies were observed for the peripheral vowels in KIT (87%), FLEECE (86%), FOOT (83%), THOUGHT (78%), and STRUT (74%), indicating that these five vowels have low levels of acoustic overlap with neighbouring CAAE vowels. Relatively lower accuracies were seen for TRAP (67%), SQUARE (66%), NURSE (63%), LOT (57%), BATH (55%), and DRESS (53%), indicating that these vowels potentially show a certain level of acoustic overlap with neighbouring vowels (see also Figure 1). Finally, we observed the lowest accuracy measure for GOOSE (36%), as the GOOSE targets were often misclassified into a wide range of vowel categories. However, other categories are seldom misclassified as GOOSE (< 5%), indicating that the vowel remains distinct from other categories. These results also show that misclassifications are not necessarily symmetrical.

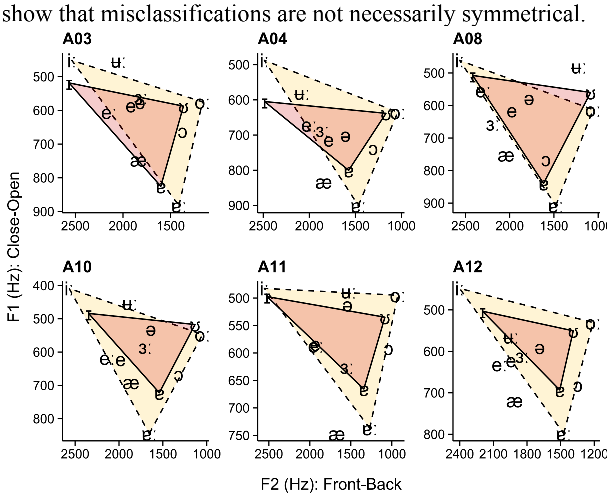

Figure 3. *Individual CAAE vowel plots*

3.2. CAAE vowel space

Figure 3 presents the acoustic vowel space of each CAAE speaker. Again, we visualised the triangles defined by the centroids of peripheral vowels, both tense and lax, (FLEECE-BATH-THOUGH, and KIT-STRUT-FOOT). All speakers produced all phonemic vowels in CAAE, but individual differences also existed in terms of how each vowel differed spectrally from those neighbouring categories. Here, we are primarily interested in the acoustic differences in critical (acoustically and articulatorily adjacent) pairs. The *p*-values reported below are the results of *F*-tests based on LMMs (mixed effects: *speaker* and *lexical item*).

For the FLEECE-KIT contrast, our analysis indicated that FLEECE vowels had longer durations ($p < .0001$), lower F1s ($p < .0001$), and higher F2s ($p = .0001$) than KIT vowels. The difference in F2 was more prominent in speakers A08, A10, and A12, as compared to other participants. With respect to the BATH-STRUT contrast, our analysis showed that BATH vowels had longer durations ($p < .0001$) and lower F1s ($p = .0164$) than STRUT vowels, but the BATH and STRUT vowels had similar F2s ($p = .2348$). The F1 difference was present in all six speakers. For the final tense-lax pair—SQUARE-DRESS—the analyses showed that SQUARE vowels had longer durations ($p < .0001$) and higher F2s ($p = .0172$) than DRESS vowels, but the difference in F1 was not significant ($p = .8175$). The difference in F2 was prominent in A03, A04, A08, and to a lesser extent A10 and A12, but potentially not A11. Finally, we compared between TRAP and DRESS, and the analysis showed that TRAP vowels had longer durations ($p = .0040$), and higher F1s ($p < .0001$) than DRESS vowels, but the two vowels had similar F2 values ($p = .6337$).

Our acoustic analysis also indicated that the THOUGHT and FOOT vowels were high back vowels in CAAE, as in SAE, and that the two vowels had similar F1 values ($p = .1062$) as well as F2 values ($p = .3220$), but THOUGHT vowels were longer than FOOT vowels ($p = .0002$), as expected. Lastly, the results indicated that that GOOSE vowels were produced with substantial fronting, as a central vowel, by all CAAE speakers, except for A08, who produced the vowel as a high back vowel (even higher than THOUGHT). Additionally, TRAP vowels ere produced with a similar vowel height to STRUT vowels, but substantial individual differences were also observed, e.g., sometimes the TRAP vowel was higher than the STRUT vowel (as in A03, A08), and sometimes it was lower than the STRUT vowel (as in A04, A11).

3.3. Contextual allophones

In what follows, we compare the acoustic quality of the contextual allophones with their canonical, or 'elsewhere' allophones. The relative position of these allophones is displayed in Figure 4, for each CAAE speaker. Note that some speakers did not produce all allophones, as a consequence of the naturalistic and conversational data collection.

We first consider the TRAP vowel and its pre-nasal allophone, the BAN vowel. Here our tests showed that BAN vowels had higher F2 values than TRAP vowels (p = .0011), and BAN vowels were also potentially longer than TRAP vowels (p = .0807), but these two had similar F1 values overall (p = .2900). In summary, BAN vowels tended to be longer, and more fronted than the default allophone, the TRAP vowel.

With respect to the NURSE vowel and its pre-lateral allophone, the GIRL vowel, the tests showed that the two vowels had similar durations (p = .8652), similar F1 values (p = .5096), but that the GIRL vowel potentially had lower F2 values (p = .0836). The F2 difference was more prominent in A04, A11, and A12, but not in other speakers. Finally, we consider the GOOSE vowel and its post-coronal allophone, TOO, as well as its pre-lateral allophone, POOL. Here, the tests showed that the three allophones did not differ statistically in duration (p = .3421), or F1 (p = .2799), though a significant difference was found in F2 values (p < .0001). A series of *post hoc* tests revealed that the POOL vowel had lower F2 values than both the GOOSE and TOO vowels (p < .0001 for two tests), but the difference between the GOOSE and TOO vowels was not significant (p = .2122). Therefore, our dataset indicates that the /ʉː/ vowel is more influenced by the lateral context than the coronal segments preceding it in CAAE than in SAE/MAE.

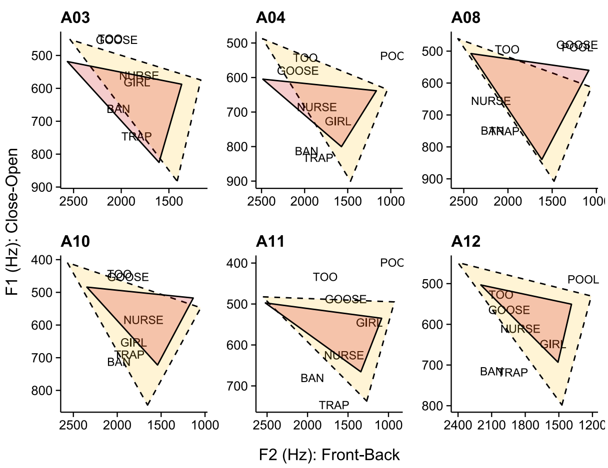

Figure 4. *Individual CAAE allophones*

4. General Discussion

The present study focused on characterising the structure of the monophthong inventory of CAAE. The results clearly indicate that the phonological structure of CAAE monophthongs is like the system in SAE, but also reveal phonetic differences in the realisation of CAAE vowels relative to SAE.

Our RM classification analysis shows that most vowels achieve an accuracy of higher than 50% in a twelve-alternative classification task with a chance-level of only 8.3%. The highest accuracy measures (>70%) were observed for peripheral vowels, presumably because these vowel categories do not share acoustic spaces with many other categories; conversely, accuracy measures were lower (>50%) in most non-peripheral vowels, potentially because they share acoustic spaces with multiple neighbouring categories. At the same time, lax vowels could be perceived/categorised as undershoots of tense vowels. Finally, the GOOSE vowel was often misclassified into other vowels, but other vowels were seldom misclassified as the GOOSE vowel, displaying an asymmetrical pattern. This phenomenon is potentially due to a large within-category variability of the GOOSE vowel over other neighbouring vowels, and more research is required to further investigate the role of relative dispersion level in vowel misclassification. In general, our RM classification suggests that all vowels were phonemically different in L1 CAAE, consistent with the previous report of L2 speakers based on a /hVd/ word-reading task [9].

Further acoustic comparisons between neighbouring pairs (Figure 2) also confirmed the finding. For example, the tense-lax pairs (e.g., BATH-STRUT, FLEECE-KIT, SQAURE-DRESS) all showed a clear difference in vowel duration, but spectral qualities also played an important role in these pairs, e.g., the BATH vowel was often lower than the STRUT vowel, and the FLEECE vowel was often more front and closer/higher than the KIT vowel. Additionally, the SQAURE vowel was sometimes more front than the DRESS vowel. These spectral features potentially differ from SAE documented in the literature, where the spectral differences between these three pairs were minimal [16]. In general, in the tense-lax pairs of CAAE, the tense vowels tend to be more peripheral than the lax vowels. See also that the red triangles were often enclosed in the yellow triangles (Figure 2). Additionally, we confirmed that a contrast between the TRAP and DRESS vowels was maintained by duration and F1, and a contrast between the THOUGHT and FOOT vowels was maintained by duration.

Finally, this study presents a first analysis of vowel allophony in CAAE, and tested whether the common allophones in MAE/SAE [12] also existed in this Aboriginal variety. Indeed, we observed that /æ/ was fronted and potentially lengthened in a pre-nasal context (i.e., the BAN vowel) as compared to its default allophone TRAP. For the /ɜː/ vowel our analyses showed that a pre-lateral context potentially led to lower F2 values, but the difference we found was nuanced and it did not reach the significance level. Finally, the /ʉː/ vowel is by default a central vowel in CAAE, as in SAE, but it can become a back vowel in a pre-lateral context (i.e., POOL), while the influence of a preceding coronal consonant was unclear, as we did not observe significant differences between TOO and GOOSE. These effects could be further investigated using a more controlled experiment-based design, potentially also with SAE control groups.

In conclusion, the study contributes to the small set of instrumental investigations into the phonology and phonetics of Australian Aboriginal varieties of English, and thus contributes to our understanding of the linguistic variations across communities in Australia. The study also highlights the benefits of using spontaneous conversational speech data as such data allows for tapping into the natural variation of any given language variety in ways that word-list or nonse-word elicitation, such as the /hVd/-list implemented in [9] might not.

5. Acknowledgements

We thank the participants in the study, Tangentyere Council Town Camps Community Centres, Charles Darwin University, Central Australian Aboriginal Congress and Batchelor Institute.

6. References

[1] ABS, "Language statistics for Aboriginal and Torres Strait Islander peoples," 2021. [Online]. Available: https://www.abs.gov.au/statistics/people/aboriginal-and-torres-strait-islander-peoples/language-statistics-aboriginal-and-torres-strait-islander-peoples/latest-release

[2] M. Sharpe, "Alice Springs Aboriginal Children's English," in *Australian linguistic studies*, S. A. Wurm, Ed. Canberra, Australia, Australia: Pacific Linguistics, The Australian National University, 1979, pp. 733–748. doi: 10.15144/PL-C54.733.

[3] I. G. Malcolm, "Australian creoles and Aboriginal English: phonetics and phonology," in *Varieties of English 3: The Pacific and Australasia*, K. Burridge and K. Bernd, Eds. Berlin & New York: Mouton de Gruyter, 2008, pp. 124–141.

[4] A. Butcher, "Linguistic aspects of Australian Aboriginal English," *Clin. Linguist. Phonetics*, vol. 22, no. 8, pp. 625–642, 2008, doi: 10.1080/02699200802223535.

[5] J. Harkins, *Bridging two worlds: Aboriginal English and crosscultural understanding*. Brisbane, Australia: University of Queensland Press, 1994.

[6] J. C. Wells, *Accents of English: Volume 1*. Cambridge: Cambridge University Press, 1982.

[7] H. Koch, "Central Australian Aboriginal English: In Comparison with the morphosyntactic categories of Kaytetye," *Aisan Englishes*, vol. 3, no. 2, pp. 32–58, 2000, doi: 10.1080/13488678.2000.10801054.

[8] H. Koch, "The influence of Arandic languages on Central Australian Aboriginal English," in *Creoles, their substrates, and language typology*, C. Lefebvre, Ed. John Benjamins Publishing Company, 2011, pp. 437–460.

[9] A. Butcher and V. Anderson, "The vowels of Australian Aboriginal English," in *Proceedings of INTERSPEECH 2008*, 2008, pp. 347–350. doi: 10.21437/Interspeech.2008-145.

[10] J. Harrington, F. Cox, and Z. Evans, "An acoustic phonetic study of broad, general, and cultivated Australian English vowels," *Aust. J. Linguist.*, vol. 17, no. 2, pp. 155–184, 1997, doi: 10.1080/07268609708599550.

[11] F. Cox and S. Palethorpe, "Australian English," *J. Int. Phon. Assoc.*, vol. 37, no. 3, pp. 341–350, 2007, doi: 10.1017/S0025100307003192.

[12] B. Purser, J. Grama, and C. E. Travis, "Australian English vower time: Using sociolinguistic analysis to inform dialect coaching," *Voice Speech Rev.*, vol. 14, no. 3, pp. 269–291, 2020, doi: 10.1080/23268263.2020.1750791.

[13] F. Cox, "Vowel change in Australian English," *Phonetica*, vol. 56, no. 1–2, pp. 1–27, 1999, doi: 10.1159/000028438.

[14] P. Boersma and D. Weenink, "Praat, a system for doing phonetics by computer [Computer program]," *Version 6.3*, 2023, [Online]. Available: https://www.praat.org

[15] J. M. Hillenbrand, "Vowel classification based on fundamental frequency and formant frequencies," *J. Speech, Lang. Hear. Res.*, vol. 36, no. 4, pp. 694–700, 1993, doi: 10.1044/jshr.3604.694.

[16] F. Cox and J. Fletcher, *Australian English: Pronunciation and transcription*. Melbourne, Australia: Cambridge University Press, 2017.