# Preliminary Investigations into the Australian English Articulatory Vowel Space

*Louise Ratko[1], Michael Proctor[1], Felicity Cox[1], Sean Veld[2]*

[1]Department of Linguistics, Macquarie University, Australia
[2]Department of Cognitive Science, Macquarie University, Australia

{louise.ratko, michael.proctor, felicity.cox}@mq.edu.au, sean.veld@students.mq.edu.au

## Abstract

Articulation of vowels produced by a single speaker of Australian English in CVC contexts was examined using Electromagnetic Articulography. Dorsal articulatory activity for each vowel was compared by tracking the midsagittal trajectories of the tongue body. Articulatory targets were determined and a companion articulatory vowel space constructed. Comparison of dorsal trajectories of vowel pairs /ɐː-ɐ/, /eː-e/, /oː-ɔ/ confirms a close articulatory relationship between long-short pairs that has previously only been examined in the acoustic domain. Long vowels were characterised by greater excursion from the centre of the midsagittal articulatory space, compared to their short equivalents.

**Index Terms**: Vowel Production, Australian English, Articulatory Phonetics, Electromagnetic Articulography, Kinematics

## 1. Introduction

The relationship between articulatory and acoustic targets for vowels is complex and remains imperfectly understood [1]. It is still not clear to what extent vowel contrasts are defined in terms of acoustic [2,3] or articulatory [4,5] goals of production, or both [5,6]. A limiting factor in our understanding of these issues is the relative paucity of articulatory data available, particularly on Australian English (AusE) vowel production.

The explanation for this scarcity is twofold; the acoustic properties of vowels position them as ideal candidates for spectrographic analysis [1,2,4]. This coupled with the historical difficulty of obtaining readily analysable articulatory data has meant that progress in vowel articulation research has lagged behind work in the acoustic domain even as technology improves.

While ultrasound and x-ray have been used for vowel analysis with a focus on tongue shape [7,8], kinematic data from Electromagnetic Articulography (EMA) provides high resolution tracking of tongue midsagittal displacement across vertical, horizontal and lateral dimensions in relation to both the participant's occlusal plane and hard palate. EMA is particularly useful for tracking the velocity and acceleration of lingual regions [9]. These data are important for informing models of articulatory kinematics in relation to vowel production [10].

In this study we examine the production of AusE monophthongs in a CVC context, by a single speaker. As a preliminary investigation, we explore the viability of characterising AusE vocalic targets using articulatory kinematics and present some initial observations about patterns of dorsal articulation to complement existing acoustic findings on AusE vowels.

### 1.1. Articulatory vowel space

Traditionally in the acoustic analysis of monophthongs, a single point is chosen, usually midway through the vowel, to represent the vowel's acoustic target [1, 11]. The values of (at least) the first two formants may be extracted at the target to construct a vowel space diagram. The F1 (high-low) and F2 (front-back) dimensions, when plotted on the Y and X axes respectively, indicate each vowel's relative position in terms of phonetic height and fronting and are often sufficient to differentiate vowels in the inventories of many of the world's languages [1].

In AusE the F1 x F2 vowel space is sufficient to differentiate all but four of the 12 AusE stressed monophthongs [11, 12], with /ɐː, ɐ/ '*Bart'* vs *'but'* (and possibly /eː, e/ *'bared'* vs *'bed'*) differing only in duration [9]. The acoustic vowel space is a useful tool for comparing vowels in a range of applications such as the investigation of differences between languages, vowel change, differences between social groups, and phonetic or prosodic contexts [11, 12, 13, 14].

Watson et al. [13] have demonstrated the limitations of the formant relationships described by the acoustic vowel space in their kinematic and acoustic study of the New Zealand English (NZE) high front vowel in 'hid'. Acoustic analysis has indicated that the target of this vowel has both retracted and lowered over time. Yet this was not reflected in the articulatory position of the tongue dorsum during target production, with a high tongue position maintained by all five of their participants.

To reconcile the discrepancy between acoustic and articulatory evidence, and unravel the complex relationship between vowel acoustics and vowel articulation, a methodological approach to the creation of an articulatory vowel space must be established. The present study aims to determine whether the articulatory parameters of tongue dorsum height and tongue dorsum fronting are useful in differentiating the monophthongs of AusE, and whether the relative dorsal positions of these vowels reflect their relative positions in the acoustic vowel space.

### 1.2. Dynamic articulatory trajectories

Although it is possible for AusE vowels to be differentiated by their acoustic target and duration [11], many intrinsic properties of vowels are dynamic in nature. The presence and length of acoustic onglides and offglides [11, 12] and formant transitions from the preceding and into the following consonant may improve vowel differentiation rates, even for monophthongal vowels [11].

This is not the case for all AusE monophthongs: the long-short vowel pairs /ɐː-ɐ/ and /eː-e/ have been shown to differ only in duration, with formant trajectories and acoustic target values near identical [12]. This is in contrast to other long-short vowel combinations in AusE such as /oː- ɔ/, which differ in spectral, as well as temporal properties.

Thus to further explore dynamic characteristics of these vowel pairs we will examine midsagittal tongue dorsum trajectories during articulation of /ɐː-ɐ/ (in *Bart* and *but*) and /eː-e/ (in *bared* and *bet*) and compare these to the dorsal trajectories of the spectrally distinct vowel pair /oː-ɔ/ (in *bought* and *pot*). It is predicted that within the long/short vowel pairs /ɐː-ɐ/ and /eː-e/ similar articulatory trajectory shapes and articulatory targets will be found consistent with prior acoustic descriptions of these vowel pairs [12]. We expect the items within the vowel pair /oː-ɔ/ to show distinct dorsal trajectories and distinct target locations as suggested by their acoustic descriptions [12].

## 2. Method

### 2.1. Participant and speech material

A native speaker (Sydney, male, 45 y.o.) produced the 18 stressed vowels of AusE across a variety of phonetic contexts including hVd, bVt and tVn. bVt tokens were chosen for preliminary analysis as they provided two clear articulatory landmarks; the labial /b/ gesture and the coronal /t/ gesture, between which articulatory trajectories could be compared across tokens. /eː/ was only elicited in bVd context. All vowels were elicited from common English words.

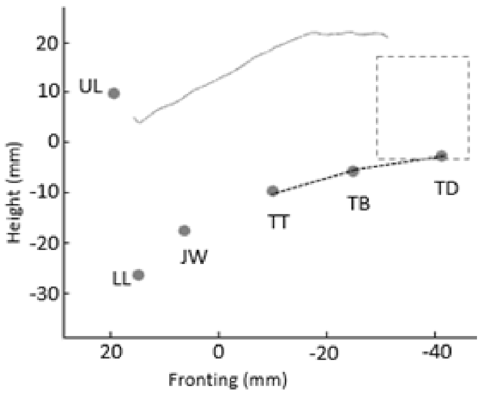


Figure 1: ***Configuration of articulators during /ɐː/ production****. Sensors attached to: Upper Lip (UL); Lower Lip (LL), Jaw (JW), Tongue Tip (TT), Tongue blade (TB), and Tongue dorsum (TD). Solid line indicates midsagittal palate trace. Vertical origin located at occlusal plane. Dashed line between TT-TB-TD sensors approximates midsagittal tongue line. Dashed box indicates limits of excursion of TD sensor values during vowel articulation.*

### 2.2. Data Acquisition

Articulatory data were acquired using an NDI Wave system sampling each sensor at a rate of 100 Hz. Nine sensors were attached to the participant. Three (nasion, left and right mastoid) were used to correct for head movement, two sensors (upper, lower lip) tracked labial aperture, and a sensor attached below the lower incisors tracked jaw movement. Three sensors were affixed to the midsagittal line of the tongue at the tongue tip (TT) 25 mm from participant's anatomical tongue tip, tongue blade (TB) 20 mm posterior to TT sensor and tongue dorsum (TD) 35 mm posterior to TT sensor (Fig.1). The occlusal plane was located with a bite trial, and the midline of the palate was traced with a custom 6D palate probe (Northern Digital Inc.).

### 2.3. Data analysis

#### *2.3.1. Acoustic/ and articulatory vowel space*

Acoustic vowel targets were extracted from spectrogram using Praat [15] and plotted in a traditional F1 x F2 vowel space (Fig. 3) based on criteria outlined in [12].

Articulatory data were analysed in MView [16]. For each token, an automatic gesture labelling procedure was used to locate the dorsal articulatory target of the nuclear vowel, from the tangential velocity profile of the TD sensor in each vocalic interval [17]. The articulatory target was the point at which the tongue dorsum sensor reached a point of minimum velocity. Tongue dorsum height was calculated in relation to the participant's occlusal plane; tongue dorsum fronting was determined in relation to the point of maxillary occlusion in line with [6, 7, 14].

#### *2.3.2. Dorsal Trajectory*

A comparable interval for vocalic analysis was identified in each token by locating articulatory landmarks in the marginal consonants. The vocalic analysis interval was bounded by the labial onset release and the closure of the coronal coda (Fig. 2).

The midsagittal location of the dorsal sensor was tracked through the interval defined between the onset and coda consonants in each token (Fig.2). For each vowel, this produced a 2-dimensional signal indicating the location of the top of the tongue dorsum in the midsagittal plane at 10 ms intervals which was used to plot dynamic midsagittal dorsal articulatory trajectories.

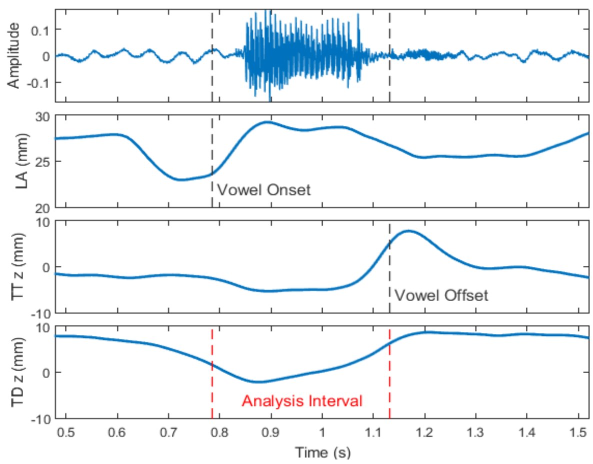


Figure 2: ***Method of location of interval of analysis****. Top: speech waveform for utterance 'bert' /bɜːt/; 2nd row: time aligned labial aperture (UL-LL mm); 3rd row: vertical location of TT sensor (mm); 4th row: vertical location of TD sensor (mm). Start of vocalic analysis interval: release of /b/ closure gesture; End of interval: onset of /t/ closure.*

## 3. Results

### 3.1. Acoustic vowel space

F2 and F1 for the participant's vowels are plotted in Fig.3. Comparison with vowel data reported in the Sydney AusTalk corpus [18] shows that the majority of the participant's tokens were within two standard deviations of reported means, with slightly lower F1 values reported for /æ, ɐ, ɐː/, and slightly lower F2 values for /ɔ/.

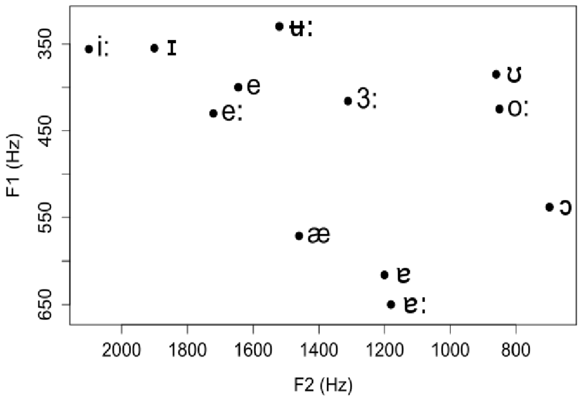


Figure 3: ***Acoustic Vowel Space.*** *First and second formants measured at point of maximum acoustic stability in each vocalic interval.*

### 3.2. Articulatory Vowel Space

Dorsal targets in the midsagittal plane for each vowel in the corpus are compared in Fig. 4. The relative configuration of articulatory targets is in general agreement with the acoustic spacing of the vowels produced by this speaker, with some notable exceptions. /ʉː/ is much more fronted in the articulatory space compared to its acoustic target, /oː/ is acoustically closer to /ʊ/ yet articulatorily closer to /ɔ/, and /ɜː/ has a low front dorsal target, yet a high mid acoustic target.

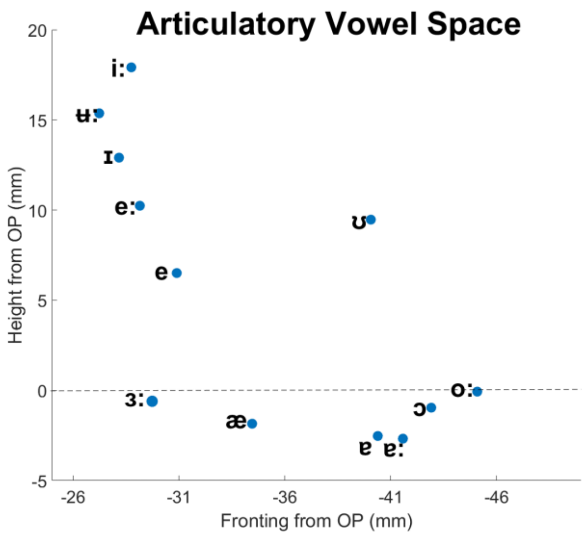


Figure 4: ***x-y coordinates of TD sensor at articulatory target of nuclear gesture in bVt/d tokens.*** *Dorsal height (mm) indicated with respect to Occlusal Plane (OP)*

### 3.3. Dynamic articulatory trajectories

Midsagittal TD trajectories for long-short vowel pairs /ɐː-ɐ/, /eː-e/ and /oː-ɔ/ produced in bVt/d contexts are illustrated in Figs. 5-7. The short vowels in the /ɐː-ɐ/ and /eː-e/ pairs were produced with similar trajectories and direction paths compared to their long counterparts. Short vowels /ɐ/ and /e/ exhibit a more centralised trajectory than their long equivalents, which show more peripheral excursion. Similar dorsal trajectories, with the same orientation are observed for back vowels /oː-ɔ/ (Fig. 7), which share similar spatial targets in the midsagittal plane, despite their considerable differences in F1 and F2 frequencies (Fig. 3).

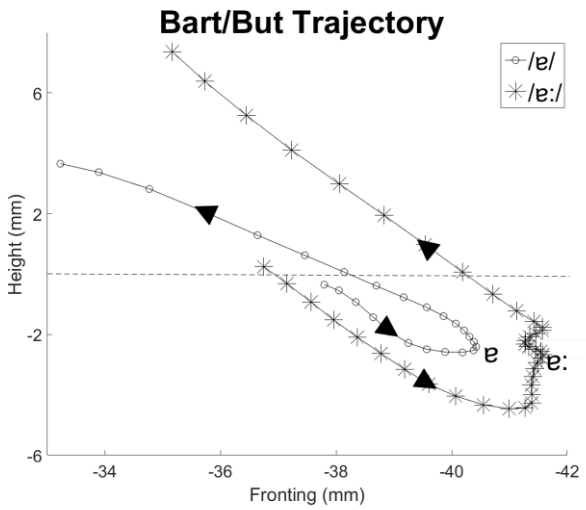


Figure 5: ***Midsagittal dorsal trajectories for long-short low vowels /ɐː-ɐ/.*** *IPA annotations indicate vowel target location. Sensor locations sampled at 10 ms intervals.*

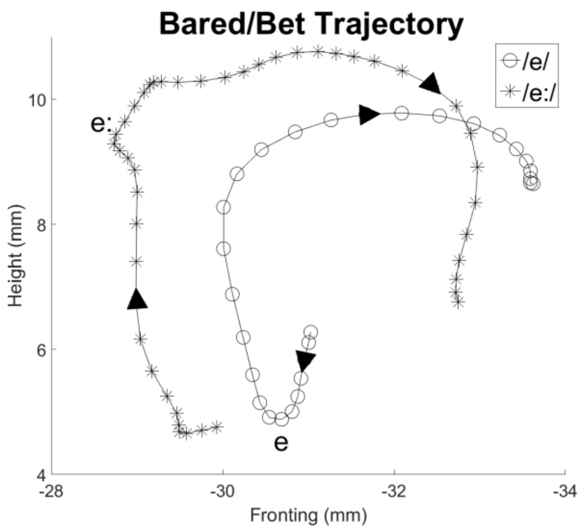


Figure 6: ***Midsagittal dorsal trajectories for long-short front vowels /eː-e/.*** *IPA annotations indicate vowel target locations. Sensor locations sampled at 10 ms intervals.*

## 4. Discussion

The vowel targets produced by this speaker were found to be broadly consistent with previous articulatory descriptions of AusE vowels [13]. Analysis of the relative distributions of

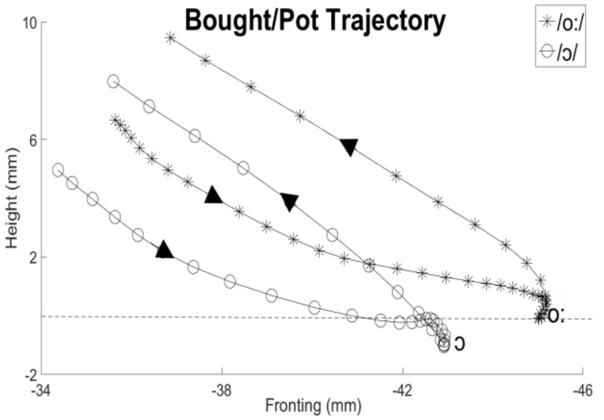

Figure 7: ***Midsagittal dorsal trajectories for long-short back vowels /oː-ɔ/.*** *IPA annotations indicate vowel target locations. Sensor locations sampled at 10 ms intervals.*

vowels in the midsagittal plane suggests that the articulatory dimensions of tongue dorsum height and dorsal fronting are sufficient to differentiate the majority of AusE monophthongs.

For /ɜː/ and /ʉː/, a discrepancy between the acoustic dimension of vowel fronting and the articulatory dimension of tongue dorsum fronting was observed, with a greater degree of dorsal fronting relative to the other vowels than the acoustic analysis of this speaker's second formant frequency suggests. A lower than expected dorsal height was also observed for back vowels /ʊ/ and /oː/, compared to the acoustic analysis of first formant targets for these vowels. This placed /oː/ in a similar dorsal articulatory space to /ɔ/ despite the acoustic similarity of /oː/ and /ʊ/.

These data suggest that a comprehensive characterisation of the AusE vowel space requires more articulatory data, as differences in dorsal posture are not always predictable from the relative acoustic properties of vowels. We acknowledge that analyses restricted to the tongue dorsum alone may fail to capture important parasagittal differences of articulation.

When we examine the dorsal target locations of the back-rounded vowels, it is clear that there is a discrepancy with the acoustic vowel space. Rounded vowels appear to be in relatively different positions in the acoustic space compared to the articulatory space. This is particularly the case for /oː/, /ɜː/ and /ʉː/. It is possible that differing degrees of lip-rounding could impact our perception of vowel height and mask quite different underlying dorsal articulations, in line with observations found in NZE [13].

Comparisons of /eː-e/, /ɐː-ɐ/ and /oː-ɔ/ reveal that the dorsal trajectories of the vowels within each pair show a close articulatory relationship. Articulatory targets of /e/ and /ɐ/ were found to be more centralised, consistent with the characterisation of these short vowels as undershot realisations of their more peripheral, longer equivalents [12]. This dorsal centralisation was not observed in production of /ɔ/, suggesting that /oː/ and /ɔ/ have distinct articulatory and acoustic targets. The similarity in dorsal trajectories observed for /oː-ɔ/ is unexpected, given that these vowels are typically realised with distinct acoustic characteristics in AusE [12].

As the realisation of /eː-e/, especially their length, may have been influenced by the difference in voicing of the following consonant (t/d), further investigation is required to examine the relationships between short and long vowels.

Extension of this research will involve the recruitment of more speakers, tokens and phonetic contexts. The investigation of articulatory parameters such as lip rounding and jaw height on AusE vowel acoustics is also warranted.

The creation of a methodological approach to kinematic vowel data analysis will provide a mechanism for fully understanding the complex relationship between vowel articulation and vowel acoustics, and offers an exciting opportunity for future researchers.

## 5. Conclusion

This study provides a preliminary characterization of the kinematics of dorsal articulation in Australian English vowel production. These data demonstrate that articulatory and acoustic vowel targets share a complex but close relationship, yet independent of acoustic representations, the articulatory dimensions of tongue dorsum height and tongue dorsum fronting are sufficient in differentiating AusE monophthongs in this speaker. Midsagittal dorsal trajectories revealed that long-short vowel pairs /ɐː-ɐ/, /eː-e/, /oː-ɔ/, previously identified as acoustically similar in AusE, also share a close articulatory relationship for this speaker. More data will be required to better understand the temporal and spatial mapping between acoustic and articulatory targets in Australian English vowel production.

## 6. References


[1] Ladefoged, P. & Johnson, K. “A Course in Phonetics”. 7th ed., Boston: Cengage, 2015.

[2] Nearey, T. M. “Phonetic features for vowels”. Bloomington: Indiana University Linguistics Club, Indiana, 1978.

[3] Perkell, J.S., Matthies, M.L., et al. “Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models”. Speech Comm, 22:227–250, 1997.

[4] Abercrombie, D. “Elements of General Phonetics”. Edinburgh: Edinburgh University Press; 1967.

[5] Lindblom, B. & Sundberg, J. “Acoustical consequences of lip, tongue, jaw and larynx movement” J. Acoust. Soc. Am, 50: 1166-1179, 1971.

[6] Noiray, A., Iskarous, K. & Whalen, D. “Variability in English vowels is comparable in articulation and acoustics” LabPhon, 5: 271-288, 2014.

[7] Hashi M., Westbury, J. R & Honda, K. “Vowel posture normalization” J. Acoust. Soc. Am., 104: 2426-2437, 1998

[8] Stone, M., Shawker, T. H., Talbot, T. L., & Rich, A. H. “Cross-sectional tongue shape during the production of vowels”. J. Acoust. Soc. Am, 83(4): 1586-1596, 1988

[9] Perkell, J., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I. & Jackson, M. “Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements”. J. Acoust. Soc. Am., 92: 3078-3096, 1992.

[10] Saltzman. E. & Munhall, K. “A dynamical approach to gestural patterning in speech production”, Ecological Psychology, 1: 333-382, 1989.

[11] Watson, C. & Harrington, J. “Acoustic evidence for dynamic formant trajectories in Australian English vowels”. J. Acoust. Soc. Am., 106: 459-468, 1999.

[12] Cox, F. “The Acoustic Characteristics of /hVd/ Vowels in the Speech of Some Australian Teenagers”. AJL, 26: 147-179, 2006.

[13] Watson, C., Harrington, J & Palethorpe, S. “A kinematic analysis of New Zealand and Australian English vowel spaces”. In ICSLP, 1998.

[14] Tabain, M., Perrier, P. & Savariaux, C. “A kinematic study of prosodic boundary effects on /i:/ articulation in French” in ICPhS, 2003.

[15] P. Boersma, and D. Weenink, *Praat: Doing phonetics by computer* (Version 5.4) [Computer program], 2009.

[16] Tiede, M. “MVIEW: software for visualization and analysis of concurrently recorded movement data, Haskins Laboratory, 2005

[17] Gafos, A., Kirov, C. & Shaw, J. “Guidelines for using Mview”, 2010

[18] Cox, F. “Formant Frequencies and Durations for /hVd/ vowels. Department of Linguistics, Macquarie University, Sydney, NSW, 2015.