INTERSPEECH 2016
September 8–12, 2016, San Francisco, USA

A preliminary ultrasound study of nasal and lateral coronals in Arrernte

Marija Tabain [1], *Richard Beare* [2, 3]

[1] La Trobe University, Melbourne, Australia
[2] Monash University, Melbourne, Australia
[3] Murdoch Children's Research Institute, Melbourne, Australia

m.tabain@latrobe.edu.au, richard.beare@monash.edu

Abstract

Ultrasound tongue image data are presented for two female speakers of the Central Australian language Arrernte, focusing on the nasal and lateral coronal consonants. These coronal places of articulation are: dental, alveolar, retroflex and palatal. It is shown that the tongue back is particularly far forward for the palatal consonant, and to a lesser extent also for the retroflex consonant. There is a general flattening of the tongue for the dental consonants. In addition, the back of the tongue is consistently further forward for the nasal consonants than for the laterals – this is true for all places of articulation. Finally, a double-pivot pattern for the retroflex articulations is observed for one speaker, but not for the other.

Index Terms: Australian languages, coronal consonants, nasals and laterals, ultrasound.

1. Introduction

Arrernte is a language of Central Australia, spoken by about 2000 people in and around the administrative township of Alice Springs (*Mparntwe*). Like most Australian languages, it has a very rich inventory of coronal consonants [1, 2, 3]. It has six places of articulation in the oral stop, nasal, and pre-stopped nasal series: bilabial, dental, alveolar, retroflex, alveo-palatal, and velar. There are also four lateral consonants in parallel with the other coronals: dental, alveolar, retroflex and alveo-palatal. The dental and palatal sounds are classified as laminal consonants, and the alveolar and retroflex sounds are classified as apical consonants [4, 5].

In this paper, we present preliminary ultrasound data for the nasal and lateral coronal consonants in this language. This work follows on from our previous articulatory work on Arrernte, using the techniques of electro-palatography and electro-magnetic articulography [6, 7, 8, 9, 10].

2. Method

2.1 Speakers and recordings

Seven female speakers of Arrernte were recorded to ultrasound using the Telemed Echo Blaster 128 CEXT-1Z, the Articulate Instruments stabilization helmet [11], the Articulate Instruments pulse-stretch unit, and the AAA software version 2.16.07 [12]. In addition we used an MBox2 Mini soundcard, a Sony lapel microphone (electret condenser ECM-44B), and an Articulate Instruments Medical Isolation Transformer. The ultrasound machine, sync pulse, sound card and a software dongle were connected via USB to a Dell Latitude E6420 laptop running Windows software. Typical frame rate was 87 f.p.s., using a 5-8 MHz convex probe set to 7 MHz, a depth of 70 mm and a field of view of 107.7 degrees (70%). An eighth potential speaker was not recorded because we were unable to see a clear outline of her tongue.

For each speaker, sample palate traces were taken in order to aid with subsequent tracking of the tongue contours. Bite plane was also measured by pressing the tongue up against a ruler held in place by the molars, and all data presented here have been rotated to this bite plane.

Recordings took place either in a hotel room in Alice Springs (five speakers) or in the staff-room of the Santa Teresa school, 85 km south-east of Alice Springs (two speakers).

Speakers read a list of 92 Arrernte words designed to present the four coronal places of articulation for the oral stop, nasal and lateral consonant series. Some of these words illustrated homorganic nasal+stop, stop+nasals or lateral+stop clusters. Wherever possible, surrounding vowels were the central vowels /a/ or /ə/, which are the most common vowels in Arrernte (Arrernte has three phonemic vowels, /a ə i/, and a fourth vowel [u] which occurs as a result of rounding on a consonant – rounded consonants were avoided in the list, though were not entirely absent). Where possible, the target consonants were illustrated both in stressed and in unstressed word position (note that schwa can be stressed in Arrernte).

The words were displayed on the laptop screen. Speakers were asked to say each word three times, or as often as possible within the 5-second recording window set by the Articulate Assistant software. Some speakers were able to produce four or five repetitions in each 5-second window. Each speaker read the list through at least once, and four speakers read through the list a second time. Some speakers chose not to produce a particular taboo word which was accidentally included on the list, and some speakers weren't sure of some words. Note also that the ultrasound machine does not begin recording until about 150 ms into the 5-second audio recording window – as a result, some repetitions were discarded because they were cut off by these limitations.

In the present study, we are only presenting data from two speakers – these particular speakers were chosen for these initial analyses because their tongue images were particularly clear for a large set of the target consonants. Note that the two speakers presented in the current study were not literate in Arrernte, and were prompted by another (literate) speaker who was present in the room, and/or by the author MT providing an English gloss of the target word. One of these speakers read through the list once, and the other speaker read through the list twice.

2.2 Analyses

Acoustic data were labelled using the EMU speech software package [13] version 2.3, and tongue contours were tracked

http://dx.doi.org/10.21437/Interspeech.2016-568

using the AAA software. Tracked tongue data were exported to Simple Signal File Format (SSFF) for compatibility with EMU. Analyses were conducted using EMU/R version 4.4, interfaced with the R statistical package version 3.1.2 [14]. Smoothing spline ANOVAs were calculated using the *gss* package in R [15]. Figures presented were created using the *ggplot2* package [16].

2.3 The current study

The current study focuses on the laterals and the nasals, since these have clearly defined acoustic onsets and offsets. Table I gives the number of tokens analyzed in the present study. It can be seen that speaker Mia has more tokens than speaker Phyllis, because Mia read through the list twice.

All vocalic and prosodic contexts were included. In particular, it should be noted that no attempt was made to separate the apicals (alveolar or retroflex) according to prosodic context: the apical contrast is in principal neutralized in word-initial position [7, 17], and the retroflex is more prototypically retroflex when it is in unstressed position [18].

Table 1. *Number of tokens.*

Mia		dental	alveolar	retroflex	palatal	Total
	lateral	46	60	34	39	**179**
	nasal	106	95	101	62	**364**
Phyllis		**dental**	**alveolar**	**retroflex**	**palatal**	**Total**
	lateral	27	46	17	20	**110**
	nasal	61	54	70	40	**225**

Tongue spline data were sampled at the acoustic onset, midpoint, and offset of the consonant. In the initial plots shown below, the offset was chosen for speaker Mia because this is the point in time where formant transitions into the vowel begin. However, for speaker Phyllis, the midpoint was chosen, since there were problems tracking the front part of the tongue for the palatal consonant for this speaker, and the problems were worse as the consonant approached the following vowel.

3. Results

Figure 1 shows nasal and lateral tongue splines for speaker Mia. It can be seen that the tongue back is more forward, and the tongue body is higher, for the palatal (blue line) than for the other coronal consonants. It can also be seen that the retroflex (pink line) has a more forward tongue back position than do the alveolar (red line) and the dental (green line) – however, the tongue back for the retroflex is not as far forward as for the palatal. It can also be seen that the middle part of the tongue is much flatter for the dental than for the other places of articulation. In general, the tongue tip is not distinguished between the alveolar, retroflex and the dental, and for this speaker, the back part of the tongue is not distinguished between the alveolar and the dental.

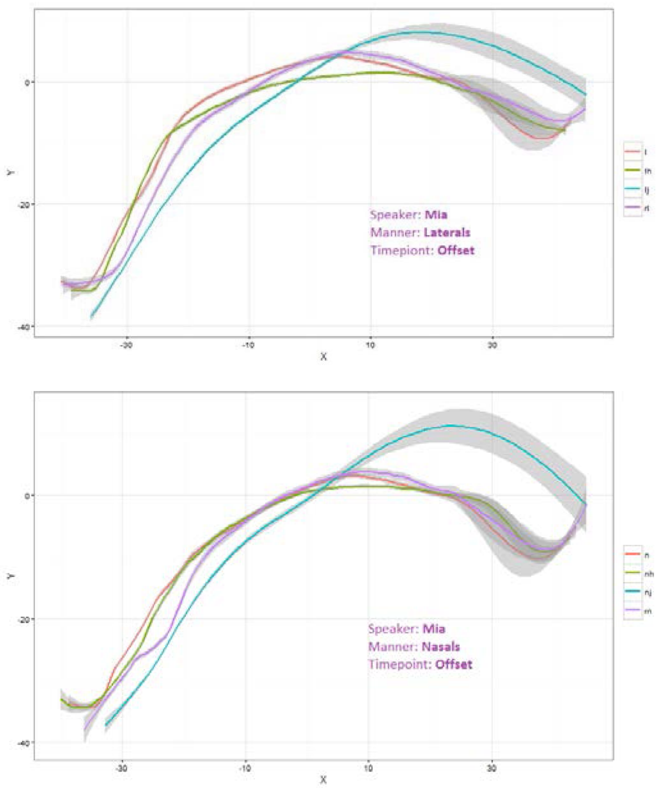

Figure 1: *Tongue contours for speaker Mia, sampled at the offset of the consonant. Top panel: laterals. Bottom panel: nasals. The red line denotes an alveolar consonant, the green line denotes a dental, the blue line denotes a palatal, and the purple line denotes the retroflex. The grey shadows surrounding each contour represent the confidence intervals generated by the SSANOVA.*

Figure 2 shows nasal and lateral tongue spline for speaker Phyllis. It will be recalled that there were problems tracking the front part of the tongue for the palatal for this speaker. However, the image of the back part of the tongue was good for this consonant, and it can be seen that as for speaker Mia, the back part of the tongue is more forward for the palatal consonant for speaker Phyllis as well (blue line). It can also be seen that the dental (green line) is flatter in the mid part of the tongue (though due to great variability in the palatal, this is not significantly different in the case of the lateral). It also seems that the mid-back part of the tongue is higher for the alveolar (red line) than for the other consonants, though this is clearer in the case of the laterals than in the case of the nasals. A similar effect of a high tongue back for the alveolar can be seen in Figure 1 for speaker Mia's laterals. By contrast, for speaker Phyllis, the retroflex (pink line) does not seem to pattern consistently for the two manner classes: in the case of the nasal, the tongue back is more forward than for either the alveolar or the dental, as was the case for speaker Mia. But for the lateral, the retroflex is more back than the dental and alveolar, at least for the most posterior portions of the tongue.

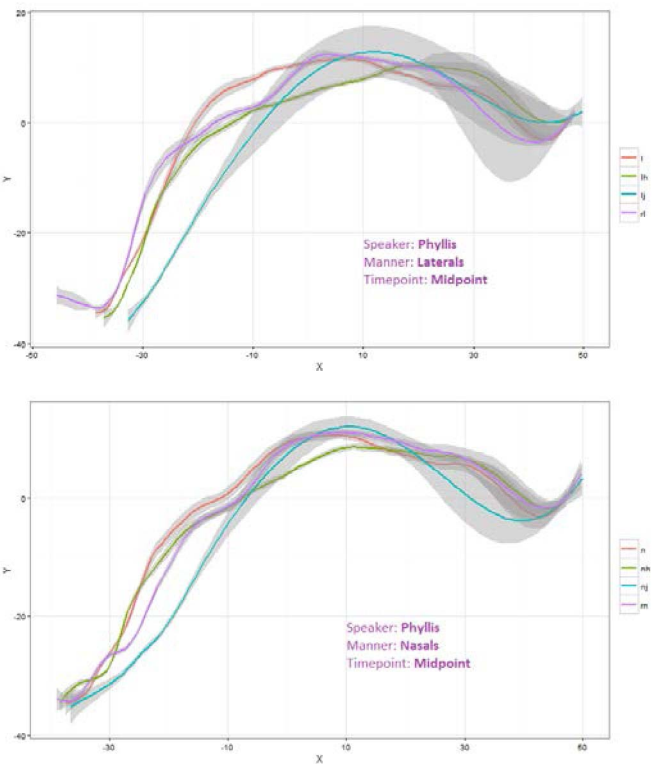

Figure 2: *Tongue contours for speaker Phyllis, sampled at the temporal midpoint of the consonant. Top panel: laterals. Bottom panel: nasals. The red line denotes an alveolar consonant, the green line denotes a dental, the blue line denotes a palatal, and the purple line denotes the retroflex. The grey shadows surrounding each contour represent the confidence intervals generated by the SSANOVA.*

Figures 3 (speaker Mia) and 4 (speaker Phyllis) show nasal and lateral tongue splines on the same plot for the different places of articulation.[1] (For the sake of direct comparison, both sets of data in this instance are sampled at the temporal midpoint of the consonant.) It can be seen that in all cases, the back of the tongue is more forward for the laterals than it is for the nasals.

Figure 5 shows the tongue contours for speaker Mia sampled at three points in time – acoustic onset, midpoint and offset of the consonant – and Figure 6 shows the same data for speaker Phyllis. It can be seen that the back of the tongue is relatively stable throughout the consonant for the dental, alveolar and retroflex, with the front part of the tongue rising from the start to the midpoint of the consonant and remaining there until the end of the consonant. By contrast, for the palatal, the back part of the tongue moves forward during the consonant at the same time as the front part of the tongue rises quite rapidly – this forms a clear pivot point in the mid part of the tongue. In addition, in the case of the palatal nasal, the front part of the tongue continues to rise from the midpoint to the end of the consonant.

[1] We would like to thank Alexei Kochetov for suggesting this analysis.

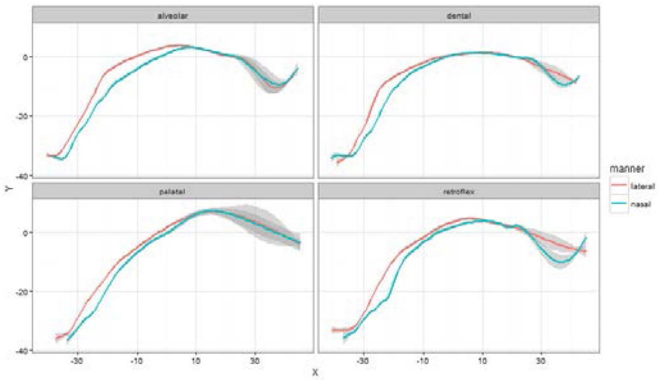

Figure 3: *Tongue contours for speaker Mia, showing nasal (blue) and lateral (red) articulations on the same panel. Each place of articulation is shown in a separate panel: alveolar (top left), dental (top right), palatal (bottom left), and retroflex (bottom right).*

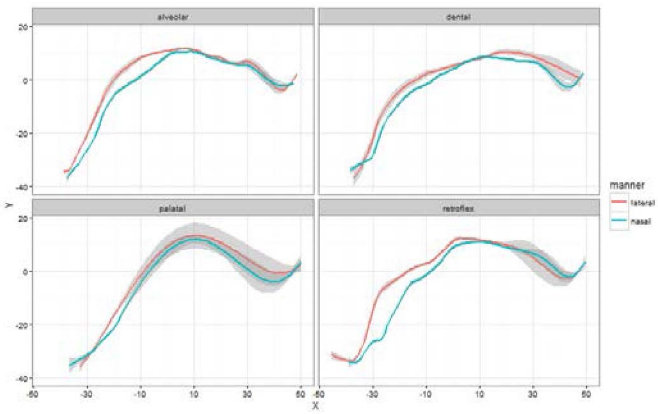

Figure 4: *Tongue contours for speaker Phyllis, showing nasal (blue) and lateral (red) articulations on the same panel. Each place of articulation is shown in a separate panel: alveolar (top left), dental (top right), palatal (bottom left), and retroflex (bottom right).*

Unfortunately the palatal data for speaker Phyllis are not very reliable. However, the alveolar consonants show very little change over the course of the consonant, with only the tongue tip rising from start to midpoint (and in the case of /l/, returning to the starting point at the end of the consonant). By contrast, the dental consonants show a distinct flattening of the mid part of the tongue from start to midpoint, together with a slight backing of the tongue for the lateral 'lh' (of course, there is also tongue tip movement for the laterals). Similarly, the retroflex may show a slight backing of the tongue from the start to the midpoint of the consonant, together with a possible forward movement in the mid-back region. This results in a sort of double pivot pattern over time, leading to the particular retroflex tongue shape evident for this speaker.

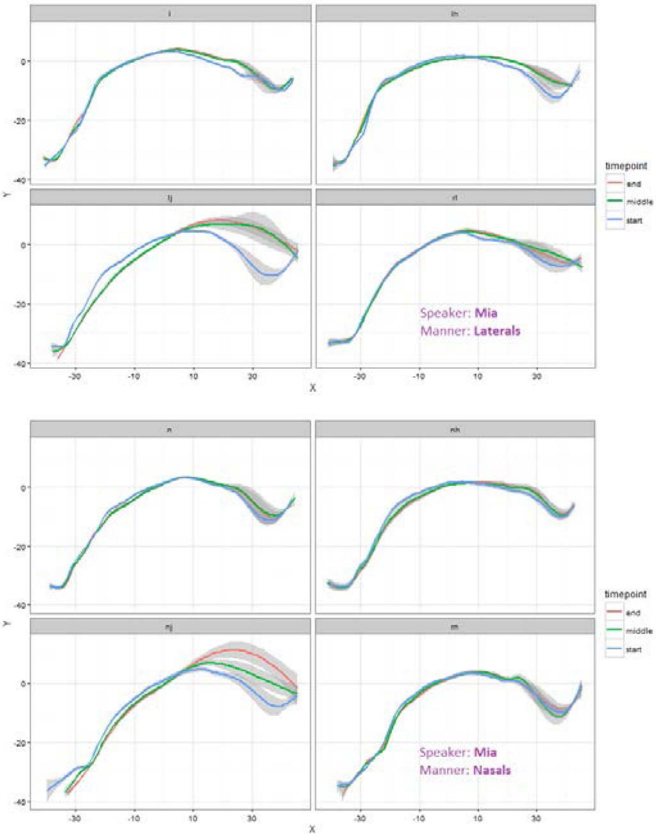

Figure 5: *Tongue contours for speaker Mia, sampled at three points in time. Top panels: laterals. Bottom panels: nasals. Each consonant is shown in a separate panel. Note that 'nh' and 'lh' denote dentals; 'rn' and 'rl' denote retroflexes; and 'nj' and 'lj' denote palatals. The blue line is the start of the consonant; green is the midpoint; and red is the end of the consonant.*

4. Conclusions

We have seen that the tongue back is particularly far forward for the palatal consonant, and to a lesser extent for the retroflex consonant. We have also seen a general flattening of the tongue for the dental consonants. These data are based on two speakers whose tongue contour images were relatively clear, and data from further speakers will confirm whether or not these observations apply across all speakers.

The tongue back is clearly further forward for the nasal than for the lateral consonants, for all places of articulation. This may be due to the need to avoid contact between the tongue back and the velum, which has been lowered in order to allow airflow through the nasal cavity. As pointed out by Fant [19], it is possible for the back of the tongue to contact the centre of the uvula during nasal consonant production, with air flowing along the sides of this connecting point—this may occur, for example, with a high tongue position and a fully lowered velum. Such a configuration may set up a different set of oral and nasal resonances, which may not be desirable in the case of a place-rich consonant system [20]. By contrast, the hydrostatic nature of the tongue may mean that as the tongue sides are lowered for lateral production, the back of the tongue is pushed further back to compensate for this. Work in progress on oral stop production in Arrernte will clarify the question of why the stops and nasals pattern differently with respect to tongue body position. It may be noted that recent work on the Indian language Kannada [21] suggests that the laterals pattern with the stops in having a tongue body position that is further back than for the nasals.

Further work which teases out prosodic contexts will likewise clarify to what extent the alveolars and retroflexes can be reliably differentiated. The double-pivot pattern seen for the retroflex for one speaker was not evident for the other speaker, and it is not yet clear to what extent there is interspeaker variation in the production of these sounds in Arrernte. It is likely that extending the analysis in time (i.e. looking at tongue shape during the previous vowel, and during the following vowel) will elucidate the articulatory strategies for the retroflex sounds.

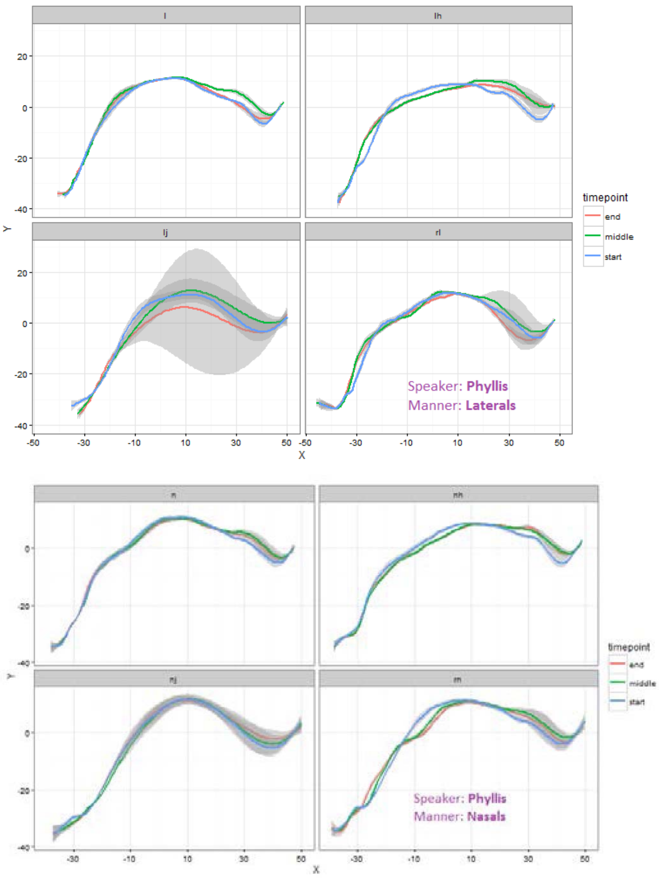

Figure 6: *Tongue contours for speaker Phyllis, sampled at three points in time. Top panels: laterals. Bottom panels: nasals. Each consonant is shown in a separate panel. Note that 'nh' and 'lh' denote dentals; 'rn' and 'rl' denote retroflexes; and 'nj' and 'lj' denote palatals. The blue line is the start of the consonant; green is the midpoint; and red is the end of the consonant.*

5. Acknowledgements

We would like to thank the Australian Research Council for funding this research via a Future Fellowship to the first author. We would also like to thank Gavan Breen for his support, and Casey Tait for labelling the acoustic data. Finally, we would like to thank our speakers for their commitment to speech research. *Kele.*

6. References

[1] R. Dixon, *The Languages of Australia*. Cambridge: Cambridge University Press, 1980.

[2] R. Dixon, *Australian Languages: Their Nature and Development.* Cambridge: Cambridge University Press, 2002.

[3] N. Evans, "Current issues in the phonology of Australian languages," In John Goldsmith (Ed.), *The Handbook of Phonological Theory* pp. 723–761). Cambridge, MA, Oxford: Blackwell, 1995.

[4] G. Breen and V. Dobson. "Central Arrernte," *Journal of the International Phonetic Association,* vol. 35, pages 249-254, 2005.

[5] J. Henderson, *Topics in Eastern and Central Arrernte grammar*. Lincom Europa: Germany, 2013.

[6] M. Tabain, "An EPG study of the alveolar vs. retroflex apical contrast in Central Arrernte," *Journal of Phonetics,* vol. 37, pages 486-501, 2009.

[7] M. Tabain, "A preliminary study of jaw movement in Arrernte consonant production," *Journal of the International Phonetic Association*, vol. 39, pages 33-51, 2009.

[8] M. Tabain, "EPG data from Central Arrernte: a comparison of the new Articulate palate with the standard Reading palate," *Journal of the International Phonetic Association*, vol. 41, pages 343-367, 2011.

[9] M. Tabain, J. Fletcher and A. Butcher, "An EPG study of palatal consonants in two Australian languages," *Language and Speech*, vol. 54, pages 265-282, 2011.

[10] M. Tabain, "Jaw movement and coronal stop spectra in Central Arrernte," *Journal of Phonetics,* vol. 40, pages 551-567, 2012.

[11] Articulate Instruments Ltd., *Ultrasound Stabilisation Headset Users' Manual: Revision 1.4*. Edinburgh, UK: Articulate Instruments Ltd., 2008.

[12] Articulate Instruments Ltd., *Articulate Assistant Advanced User Guide: Version 2.14*. Edinburgh, UK: Articulate Instruments Ltd., 2012.

[13] Jonathan Harrington, *The Phonetic Analysis of Speech Corpora*. Blackwell, 2010.

[14] R Core Team "R: A language and environment for statistical computing. R Foundation for Statistical Computing," Vienna, Austria. URL http://www.R-project.org/, 2014.

[15] Chong Gu, "Smoothing Spline ANOVA Models: R Package gss," *Journal of Statistical Software*, vol. 58(5), pages 1-25. URL http://www.jstatsoft.org/v58/i05/, 2014.

[16] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York, 2009.

[17] M. Tabain, G. Breen, A. Butcher, A. Jukes and R. Beare, "Stress effects on stop bursts in five languages," *Laboratory Phonology,* 2016, in press.

[18] M. Tabain and R. Beare, "An EPG and EMA study of apicals in stressed and unstressed position in Arrernte," *18th International Congress of the Phonetic Sciences.* Glasgow: Scotland, 2015.

[19] G. Fant, *Acoustic Theory of Speech Production,* 2nd ed. (Mouton, The Hague), 1970.

[20] M. Tabain, A. Butcher, G. Breen and R. Beare, "An acoustic study of nasal consonants in three Central Australian languages," *Journal of the Acoustical Society of America*, vol. 139, pages 890-903, 2016.

[21] A. Kochetov and N. Sreedevi, "Manner-specific tongue shape differences in the production of Kannada coronal consonants," paper presented at the Spring 2016 meeting of the Acoustical Society of America, 2016.