Some initial findings regarding first language influence on playing brass instruments

Matthias Heyne[1,2], Donald Derrick[2,3]

[1]Department of Linguistics, University of Canterbury, New Zealand
[2]New Zealand Institute of Language, Brain, and Behaviour, University of Canterbury, New Zealand
[3]MARCS Institute, University of Western Sydney, Australia

`matthias.heyne@pg.canterbury.ac.nz, donald.derrick@canterbury.ac.nz`

Abstract

This paper presents some initial findings regarding the influence of First Language on playing brass instruments. Using ultrasound imaging of the tongue, vowel production and sustained trombone notes were compared for a New Zealand English and a Tongan speaker. It is suggested that, during trombone playing, the tongue shapes used by the Tongan participant pattern with the back vowel /o/ while those used by the New Zealand English player pattern with the centralized KIT vowel (/ə/) and schwa (/ə/). It is argued that these findings provide preliminary evidence of First Language influence on brass playing.

Index Terms: laboratory phonology, phonetics, phonetics of music, ultrasound imaging of the tongue (UTI)

1. Introduction

1.1. The acoustics of brass instruments

In general, tone production on brass instruments is regarded as happening via an outward-striking lip-reed mechanism (often referred to as the player's 'embouchure') that excites the air column within the instrument, producing a spectrum of standing waves which are controlled by the natural frequencies of the air column and which are emitted from the bell at varying volumes [1, 2]. Regarding this in analogy to the human voice, one can think of the embouchure as the brass player's larynx and the instrument as the vocal tract which serves merely as an amplifier. The number of close pitches that can be produced, however, is much smaller than for the human voice due to the much greater length of instrument as compared with the vocal tract; there is also no mechanism comparable to the human tongue which would enable brass players to produce such a wide range of timbres as we do in speech and singing.

The central issue regarding the proposed influence of First Language on brass playing is the impact of vocal tract influences on brass instrument sound. As early as 1987, Campbell and Greated considered the possibility of looking "on the players windway as effectively a second 'brass instrument', with the air flowing in the 'wrong' direction (that is, towards the lips rather than away from them)" [1, p. 324]. More recently, the documentation of vocal tract tuning to produce the notes in the highest register of the saxophone [3], as well as investigations of the playing technique of the didgeridoo [4] have shown that vocal tract resonances are indeed quite important for the production of wind instrument sound and changes in timbre. For the trombone, recent research [5] has documented the increasing relevance of vocal tract impedance as one ascends to the higher partials of the instrument, but not so much for the lower register of the instrument.

Based on this literature review, we conclude that different tongue positions are possible in playing brass instruments (at least at the lower end of the register) and that they can lead to perceivable differences in timbre independent of the produced pitch.

1.2. Prior research on the function of the tongue in playing brass instruments

As far as we know, no previous empirical research has investigated the influence of language on playing brass instruments, although Budde [6] included a substantial section on phonetics plus another eight pages specifically dealing with native language and its potential effects on brass playing and their pedagogical implications in his Ph.D. thesis.

There exist, however, a small number of theses undertaken in the years from 1954 to 1970 [7, 8, among others] which used x-ray imaging to observe the tongue during playing and two of the researchers even recorded tongue shapes for selected vowels [7, 8]. Overall, the findings of these studies revealed significant inter-individual differences for most of the dimensions that were measured and except for Hall [7], researchers let participants play on their individual instruments and mouthpieces, which would have introduced an additional confound. As for matching up the tongue shapes used in playing with those of the elicited vowels, similarly no consistency was found although the reason for this may have been the limited number of vowels (/a/, /i/, /u/) that were recorded. An interesting finding from Hiigels study was that "[N]o evidence was found to support the postulate that thinking a syllable during performance will tend to simulate the tongue position resulting from the enunciation of that syllable" [8, p. 108] which proved that the long documented history of recommendations passed on to students by brass teachers on using specific syllables during playing (cf. [9]) might help achieve their aim but did not reflect the physical reality. A single more recent study exists [10], which makes use of MRI but carries limited interpretability because no real trumpet could be used within the scanner and only a single subject was observed.

2. Methods

2.1. Ultrasound imaging of the tongue (UTI)

Ultrasound imaging of the tongue provides a noninvasive and relatively inexpensive method for imaging the tongue and has previously been shown to be suitable for recording tongue positioning while playing wind instruments [11].

In order to make it possible to record trombone players who

have tubing of their instrument running along the left side of the neck, we modified a non-metal jaw brace previously designed at the NZILBB and reduced its width so that for most players it now does not touch the tubing except for a thin strap running along the side of the player's face. The jaw brace stabilizes the ultrasound probe against the jaw, tying the probe motion to jaw motion and thereby reducing motion variance.

Assessment of the system is currently in progress [12]. On-going analysis of combined articulometry and ultrasound data collection of 10 speakers reading sentences for 40 minutes each suggests that the ultrasound probe translates around 2 mm, and rotates about 0.5 degrees (standard deviation) during tasks where the speaker is focused on a screen or page similar to this research. These motion ranges are well within the 3mm of motion of variance expected as a result of ultrasound data noisiness combined with the human and computer measurement variance expected during tongue surface measurements .

2.2. Participants

The findings presented here are based on a full dataset recorded with a male speaker of Tongan who grew up in Tonga and started learning English once he arrived in New Zealand as an adult twenty years ago. He reported himself as being an amateur player on the trombone although he had started playing brass instruments in High School in his home country. The second participant was a professional trombonist whose first and only language was New Zealand English (NZE); this participant's data was recorded as our pilot study and the size and reliability of the dataset is thus much reduced as compared to that of the other participant.

2.3. Instruments

The trombones used in recording both participants were identical plastic 'pBone' trombones made by Warwick Music in the UK. While the pilot participant brought his own mouthpiece, the Tongan participant used a standard 6 1/2 AL mouthpiece supplied by the researchers.

2.4. Recording

Both participants were asked to come to the NZILBB sound attenuated booth and given sufficient time to warm up and familiarize themselves with the 'pBone' as well as filling in a short questionnaire about their language proficiency and playing experience. They were then asked to put on the jaw brace with the ultrasound transducer and adjustments were made to allow for a comfortable fit.

The ultrasound machine used for the recordings was a GE Healthcare Logiq e (version 11), with a 8C-RS wide-band microconvex array 4.0-10 MHz transducer.

Ultrasound tongue images were captured on a late 2013 15", 2.6 GHz macbook pro with retina display using an epiphan VGA2USB pro frame grabber for the video, and a sound devices LLC USB Pre 2 microphone amplifier through ffmpeg. Video was encoded using X.264 and audio was encoded as uncompressed 44.1 kHz mono. Frame rates varied between 58 and 60 Hz and were encoded in a progressive scan uyvy422 pixel format at 1024x768.

2.5. Procedure

The first part of the experiment consisted of reading English and Tongan wordlists, respectively; while number of words was much smaller and printed on paper for the pilot participant, the

Tongan participant was asked to read Tongan words in blocks of 3-5 from a computer screen that automatically advanced to the next block after a set timeframe.

In the second part of the experiment, the participants played an almost identical set of eight musical passages which featured sustained notes at varying dynamics, in different registers, and different kinds of articulation including double-tonguing (for which a second place of articulation behind the alveolar ridge has to be used) and lip slurs (production of different pitches by changing the vibrating frequency of the lips, without moving the slide). Only two of these exercises required the use of the slide to alter the fundamental pitch of the instrument; these were original etudes written for trombone (with only some slight modifications).

2.6. Data analysis

To analyze the tongue shapes for vowels (speech) and sustained notes (during trombone playing), the relevant frames of the ultrasound video were identified and annotated using Praat (for audio only) [13] and verified using ELAN [14]. For speech, we used the midpoint of vowels based on the presence of a clear F2 and its relative steadiness while for sustained notes the selection of frames was based on visual inspection of tongue steady-state for the pilot data and set at a third of the sustained note duration for the Tongan participant's data. The tongue contours on the selected video frames were then extracted by clicking on the visible contour with a script programmed for MATLAB [15, 16] and average contours were calculated by translating the locations of the points on the tongue contours into polar coordinates with the position of the ultrasound transducer as point of origin and using R [17] to fit an ssanova [18] on this data.

To enhance the clarity of the plots presented on the following page, we collapsed the initially separately analyzed tokens for stressed and unstressed, regular and long vowels in the Tongan dataset and calculated joint average curves. The same was done for the curves labelled low and high notes for both participants; low notes include the pitches Bb2 and F3, middle notes only Bb3, and high notes D4 and F4 as per the US standard system for specifying pitch.

3. Findings

Figures 1 and 2 show the average tongue contours for the Tongan and NZE participants, respectively; note that the different number of curves represented reflects the much smaller number of vowels in Tongan as compared to NZE. The average curves for the Tongan participant are based on a minimum of 80 tokens for each vowel and 33 for the sustained notes. For the pilot data recorded with the NZE participant the numbers are 5 and 36, respectively, except for /ɒ, ɵː, ʊ/ with only three or four tokens each; we nevertheless decided to include these vowels in the plot for reference.

While no one-on-one mapping of the sustained note contours onto those of the vowels is possible, it can be seen that the average curves for the high notes of the Tongan participant and those for the low notes of the NZE participant pattern with certain vowels in the languages respective vowel systems.

In the case of the Tongan participant the vowel most closely approximated by the tongue shape used while playing high notes on the trombone is /o/, although the positions of the tongue tip and dorsum are further back and as the tongue raises to produce lower notes the closest vowel becomes /u/. In the case of the NZE participant, the tongue shape for low notes lies
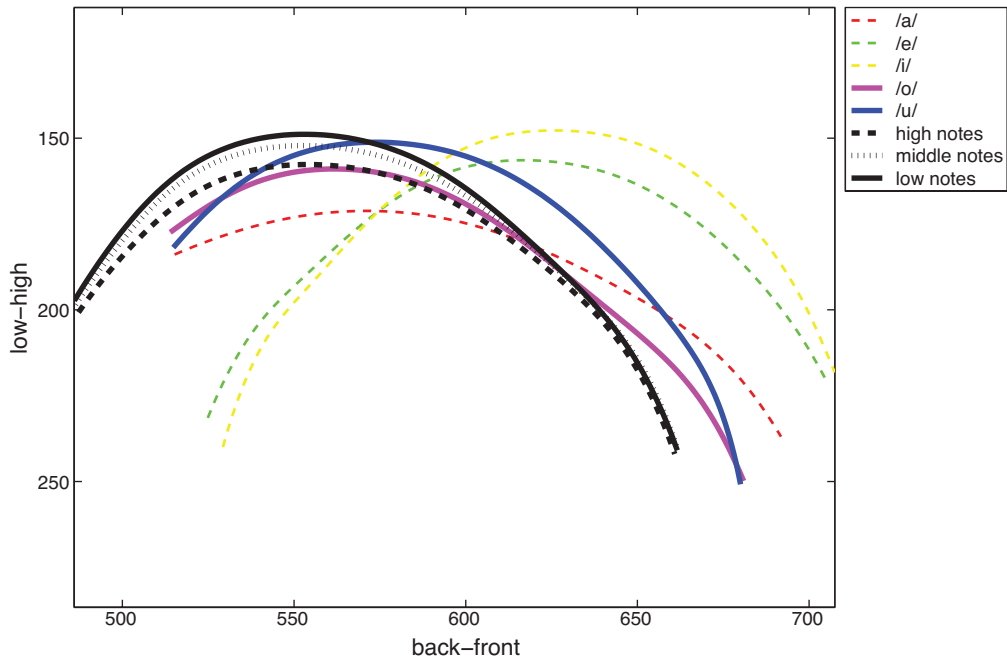
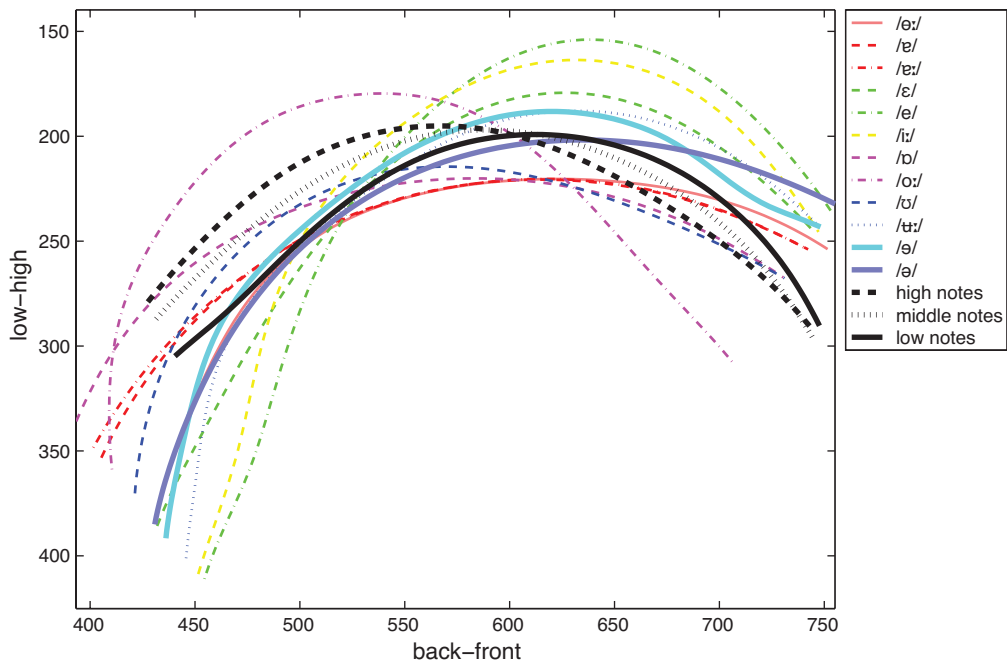Figure 1: *Average tongue contours for vowels and sustained notes for our Tongan participant.*

Figure 2: *Average tongue contours for vowels and sustained notes for our NZE participant.*

in between those of the centralized /ɔ/ and /ə/ vowels except for the tongue tip and dorsum which behave similar to those of the Tongan participant, albeit deviating more significantly from the vowel contours.

Another observation is that although the back of the tongue assumes a higher position for the lower notes in the case of the Tongan participant, this pattern is reversed for the NZE player.

4. Discussion

The findings presented in the previous section seem to suggest that brass players' First Languages constrain their availability of tongue shapes to be used in producing sustained notes.

It has been shown using MRI [19] that for extreme vowels, namely high front and low back vowels, the airway in the oral and pharyngeal cavities, respectively, is heavily constricted and would significantly reduce airflow, which might explain why

brass players use centralized vowel tongue positions instead. We expect most players whose First Languages have such a centralized vowel to use it in playing, while players whose First Languages do not offer such a tongue position are likely to use a different one which closely resembles one of their other vowels; these tongue shapes should be readily available to brass players when they first start learning to play even at a young age. Furthermore, we do not expect frequency effects to play a role in the selection of the vowel tongue position to be used when playing brass instruments; rather, acoustical considerations seem to be more important, otherwise the Tongan player should be using /a/ instead of /o/.

If no central vowel is available in a brass player's First Language, mid to high back vowels seem to be next best choice as these do not produce a significant constriction in the pharynx and leave enough space at the tongue tip (compare below). This is exemplified by our Tongan subject and also data from a professional Japanese trumpet player recorded using MRI [10]. Even though many differences exist between our research and the MRI study, including the choice of instrument (trombone versus a plastic trumpet without valves, designed for use inside the MRI scanner), some interesting parallels should be noted: Japanese also has a five-vowel system similar to Tongan and the Japanese player likewise used tongue shapes approximating /o/ and /ɯ/, although the pattern for high and low notes was reversed (parallel to the NZE speaker in our study).

Differences in the height of the back of the tongue for the different registers may be related to brass players' skill levels; the pattern used by professionals (our NZE participant and the Japanese trumpet player) seems to keep the constriction inside the oral cavity as small as possible and only allows the tongue position to rise for high notes which on average require less airflow. In general, we expect professional players of brass instruments to display less language influence than amateurs as these individuals spend countless hours practicing their instruments to improve their sound and articulation which should lead to the gradual unlearning of tongue movement patterns acquired with their First Language.

As for why the tongue tip is further back in playing than for most vowels observed in speech we can only speculate that this is related to the amount of airflow used in brass playing and the requirements of embouchure formation.

5. Conclusion and future work

The differences in the vowel systems of the First Languages of two trombonists in our study seem to constrain the availability of tongue shapes to be used in playing their instruments. This finding represents the first evidence that playing differences among players from different countries are not simply cultural or a matter of style but may also be rooted in the phonetics and phonology of different languages. We also hypothesize that this can be shown when looking at coronal and velar consonants used for articulating notes on brass instruments and have recorded data for five additional participants which we are currently analyzing.

In the future, we plan to quantify our findings by recording and analyzing more data and investigate the aural perceptibility of playing differences using expert and non-expert listeners.

6. Acknowledgements

7. References

[1] Murray Campbell and Clive A. Greated, *The musician's guide to acoustics*, New York, Schirmer, 1987.

[2] Arthur H. Benade, *Fundamentals of music acoustics*, New York, Oxford Univ. Press, 1976.

[3] Jer-Ming Chen, John Smith, and Joe Wolfe, "Saxophonists tune vocal tract resonances in advanced performance techniques," *J. of the Acoustical Soc. of America*, vol. 129, no. 1, pp. 415-426, Jan. 2011.

[4] Joe Wolfe *et al.*, "Some effects of the player's vocal tract and tongue on wind instrument sound," in *Proceedings of the Stockholm Music Acoustics Conference*, Stockholm, Sweden, 2003, pp. 307-310.

[5] Vincent Frour, "Acoustic and respiratory pressure control in brass instrument performance," Ph.D. thesis, Schulich School of Music, McGill Univ., Montreal, Canada, 2013.

[6] Paul Joseph Budde, "An analysis of methods for teaching middle school band students to articulate," Ph.D. thesis, Grad. School, Univ. of Minnesota, Minneapolis and St. Paul, MN, 2011.

[7] Joseph C. Hall, "A radiographic, spectrographic, and photographic study of the non-labial physical changes which occur in the transition from middle to low and middle to high registers during trumpet performance," Ph.D. dissertation, Grad. School, Indiana Univ., Bloomington, IN, 1954.

[8] Lewis Edward Hiigel, "The relationship of syllables to pitch and tonguing in brass instrument playing," Ph.D. thesis, Ed. Dept.,Univ. of California, Los Angeles, CA, 1967.

[9] Luis Loubriel, *Lasting change for trumpeters*. Minneapolis, MN: Scholar Publications, 2011.

[10] Tokihiko Kaburagi *et al.*, "A methodological and preliminary study on the acoustic effect of a trumpet players vocal tract." *J. of the Acoustical Soc. of America*, vol. 130, no. 1, pp. 536-545, July 2011.

[11] Joshua T. Gardner, "Ultrasonographic investigation of clarinet multiple articulation," D.M.A. thesis, Graduate College, Arizona State Univ., Tempe, AZ, 2010.

[12] Donald Derrick, Romain Fiasson, and Catherine T. Best. In Prep. "Co-collection and co-registration of ultrasound and articulometry data for speech research."

[13] Paul Boersma and David Weenink, "Praat: Doing phonetics by computer [Computer program]." Version 5.3.52, retrieved 18 May 2014 from http://www.praat.org/

[14] P. Wittenburg *et. al.*, "ELAN: a professional framework for multimodality research," in *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, Genoa, Italy, 2006.

[15] Mark Tiede, "Get Contours [Computer program]." Sent 1 May 2014.

[16] MATLAB Release 2013b, The MathWorks, Inc., Natick, Massachusetts, United States. 2013.

[17] R Core Team, R: *A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna*, Austria, 2013 URL http://www.R-project.org/.

[18] Chong Gu, gss: General Smoothing Splines. 2014 URL http://CRAN.R-project.org/package=gss

[19] T. Baer *et al.*, "Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels," *J. of the Acoustical Soc. of America*, vol. 90, no. 2, pt. 1, pp. 799-828, Aug. 1991.