

# THE INFLUENCE OF TONGUE POSITION ON TROMBONE SOUND: A LIKELY AREA OF LANGUAGE INFLUENCE

Matthias Heyne<sup>1,2</sup>, Donald Derrick<sup>2</sup>

<sup>1</sup>Department of Linguistics, University of Canterbury, New Zealand

<sup>2</sup>New Zealand Institute of Language, Brain and Behaviour, University of Canterbury, New Zealand  
matthias.heyne@pg.canterbury.ac.nz, donald.derrick@canterbury.ac.nz

## ABSTRACT

This paper builds on initial evidence of First Language influence on brass playing presented in Heyne and Derrick (2013) [13] by indicating how tongue positioning might affect trombone timbre. Ultrasound imaging of the tongue was used to compare vowel production and sustained trombone notes for three participants, one each of New Zealand English, Tongan and Japanese, whose musical production was also analyzed acoustically.

Comparison of the sound spectra produced by two semiprofessional players shows that the player using a higher, more retracted tongue position displays a larger component of high frequencies in the produced sound spectrum. We believe that this could explain why brass players can notice differences between players from different language backgrounds.

**Keywords:** laboratory phonology, phonetics, phonetics of music, ultrasound imaging of the tongue (UTI), acoustic analysis

## 1. INTRODUCTION

### 1.1. Vocal tract influences on brass instrument sound

We argue that the tongue position used while playing the trombone influences the timbre of the instrument, contra older models that downplay this notion [4, 8]. Recent research has shown that so-called ‘vocal tract tuning’ is necessary to produce notes in the highest register of the saxophone [6], as well as for producing different timbres on the didgeridoo [21]. For the trombone, the increasing relevance of vocal tract impedance as one ascends to the higher partials of the instrument has been shown for some players [9] and observations of an ‘artificial lip reed player’ have shown that large changes in tongue position affect the intonation and intensity of partials within the produced overtone spectrum; however, overall, matching vocal tract to instrument impedance seems to be less of a requirement for the lower register of the instrument.

Observations by Hall as part of his early x-ray investigation on “physical changes during trumpet playing” [12] provide empirical documentation of

the phenomenon, but the scanned quality of the x-ray negatives and spectrograms included in the study prevents the critical inspection of his claims.

### 1.2. Recent research on the influence of tongue position on brass instrument sound

In [13], we provided a summary of some older research using x-ray imaging to investigate the function of the tongue in brass instrument playing; here we would like to add more detailed information on an acoustical investigation employing MRI by Kaburagi et al. [14], who investigated the effect of vocal tract resonances on the trumpet sound produced by a single professional player. They found that although vocal tract alterations accompanying changes in pitch occur all along the vocal tract from the glottis to the lips, tongue positioning plays a big part in changing vocal tract impedance. Their Japanese player used a tongue position similar to /o/ for a low and mid-range note and positioned the tongue similarly, but “slightly posteriorly”, to /u/ (/uu/) for a high note.

### 1.3. Hypothesis

Different tongue positions assumed while playing sustained notes on the trombone lead to differences in timbre. More specifically, we speculate that a higher, more retracted tongue position should produce a larger component of high frequencies in the produced sound spectrum, based on some commonly shared assumptions in the brass teaching community. These tend to associate smaller cavities (e.g., narrow bore of an instrument, small mouthpiece cup size, a ‘tight’ throat) with the production of ‘brighter’ sounds.

## 2. METHODS

### 2.1. Ultrasound imaging of the tongue

Ultrasound imaging of the tongue is a noninvasive and relatively inexpensive method for imaging the tongue and has previously been used to collect midsagittal tongue contours during wind instrument playing [10]. We used a modified non-metallic head-mounted ultrasound probe holder designed to allow

trombone tubing to run along the left side of the player's neck without bumping the probe or head mount; this holder stabilizes the ultrasound probe against the jaw. Assessment of this system [7] shows that 95% confidence intervals of probe motion and rotation were well within acceptable parameters described in the HOCUS paper [19].

## 2.2. Participants and instruments

We recorded one male speaker each of the following three languages: (1) A speaker of New Zealand English (NZE) who did not report speaking any other language; he is a semiprofessional player on the trombone, having played the instrument for eighteen years and is also active as a singer in a barbershop quartet. (2) A speaker of Tongan who grew up in Tonga and only acquired English upon arrival in New Zealand as an adult; he described himself as an amateur player although he had started playing various brass instruments as a secondary school student in his home country. And finally, (3) a speaker of Japanese who has lived in New Zealand for eleven years and only started learning English in his forties, a few years before relocating to New Zealand; he indicated his playing level as almost professional, having played the instrument for forty years although he does not work as a musician.

The trombone used for recording all participants was a plastic 'pBone' trombone by Warwick Music in the UK; the mouthpiece used was a standard 6 1/2 AL by Arnold's and Son's, Wiesbaden, Germany.

## 2.3. Recording procedure

All participants were asked to come to a small sound attenuated room on campus and given sufficient time to fill in a short questionnaire about their language proficiency and playing experience, as well as to familiarize themselves with the 'pBone'. They were then asked to put on the jaw brace with the ultrasound probe and a comfortable fit was assured by making some adjustments.

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 video was captured on a separate laptop using a USB frame grabber for the video, and for the audio, a USB audio interface connected to a Sennheiser MKH 416 shotgun microphone which was placed as close as possible to the participants' lips for the speech recordings and at about one bell size diameter's distance for the musical passages. Frame rates varied between 58 and 60 Hz and the video codec used was X.264 with uncompressed 44.1 kHz mono audio.

The first part of the experiment consisted of reading English, Tongan and Japanese wordlists, respectively; these lists were designed to elicit all vowels of each language in different phonetic contexts and also included consonants, which will be analyzed in future research.

In the second part of the experiment, the participants played an almost identical set of eight musical passages including sustained notes at varying dynamics, in different registers, and different kinds of articulation including double-tonguing and lip slurs (production of different pitches by changing the vibrating frequency of the lips, without moving the slide). The use of the slide to alter the fundamental pitch of the instrument was required in only two of these exercises, slightly modified original etudes written for trombone.

## 2.4. Analysis of ultrasound data

For the analysis of tongue shapes for vowels (speech) and sustained notes (during trombone playing), the relevant frames of the ultrasound video were identified and annotated using Praat [3] and verified in ELAN [20]. For speech, we used the midpoint of vowels based on manual annotation; for the trombone notes, the selection of frames was set at a third of sustained note duration. Selected tongue contours were extracted by clicking just below the visible contour using Get Contours [15, 18] and polar coordinates of the contours based on the transducer head as the vertex were used to calculate average curves by fitting an SSANOVA [11] using R [16].

## 2.5. Acoustic analysis

To extract the average harmonics of notes played by the participants we used a MATLAB function for harmonics analysis [17], employing settings that were kept constant for all analyzed notes. Each sound sample was 60ms long and taken from 30ms before to 30ms after the timestamp value of the ultrasound frame used for extracting the corresponding tongue contour; frequency resolution was improved by applying zero padding and notes whose fundamental frequency deviated more than an equal temperament quartertone from the mean of all tokens for that note were eliminated. Finally, average frequencies and magnitudes of each harmonic peak were calculated, and evaluated statistically using unpaired, one-tailed t-tests in MATLAB [15].

To address within and between subject recording quality variance, we averaged across a large number of tokens produced while playing identical musical passages. In addition, participants were instructed to

keep the same distance from the microphone, keep the slide locked for six out of eight musical passages, and we checked that the tuning slide was completely pushed in before starting recordings. Furthermore, no frequencies beyond the documented cutoff frequency for the trombone at approximately 1000 Hz [2] were included in the analysis.

### 3. FINDINGS

#### 3.1. Ultrasound data

Figure 1 shows the average tongue contours for the NZE, Tongan and Japanese participants, respectively. For the vowels, ‘prominent’ tokens according to the prominence patterns of each language (syllable-based with ‘stress’ for NZE, mora-based with ‘accent’ for Tongan and Japanese; NZE schwa, of course, is unstressed), are plotted in color and these are based on at least twenty-one tokens each. For the note contours, represented by different line styles, the pitches are labeled as per the US standard system for specifying pitch and these are based on a minimum of eight tokens each.

Although no direct mapping of sustained note contours onto those for the vowels is possible, visual analysis shows that for the two participants speaking languages with only five vowels, these tongue curves are closest to the tongue position for the vowel /o/, whereas for the NZE participant, the contours fall in between the tongue positions for schwa (/ə/) and /e/. Furthermore, the two participants with at least semiprofessional playing ability (S5 NZE and S7 Japanese) raise the back of their tongues for higher notes, while the Tongan amateur player shows the opposite pattern.

It is impossible to directly compare absolute place of articulation across players and languages, but the location within the respective vowel systems provides some clues regarding the relative position of vowels. In terms of our hypothesis, we observe that the NZE player uses a playing tongue position that is typically associated to be quite central, while the other players use a typically higher and more retracted tongue position during playing, which we assume would lead to a smaller oral cavity. We thus predict that the NZ player may produce a less ‘bright’ sound, with lower magnitudes of higher harmonics, as compared to the other participants.

#### 3.2. Acoustic data

For our acoustic analysis, due to limited quality of data from our Tongan player, we focused on comparing the NZE and Japanese players only.

Results show that the Japanese player had significantly higher magnitude values for all of the

harmonics of all of the notes included in the analysis except for the fundamentals of F3, Bb3, and D4, and the second harmonic of D4, none of which were significantly different. In other words, the magnitude of the higher harmonics is larger, but the magnitude of the fundamental is not – corresponding to a ‘brighter’ sound.

Furthermore, these significant differences met or exceeded the ‘Just Noticeable Difference’ reported in Carral (2011) [5] for non-musicians to distinguish trombone sounds with different spectral centroids at 75% accuracy (1.28dB; the difference for the second harmonic of was 1.26dB).

There are also small differences in intonation for the different harmonics that seem to pattern with the magnitude differences in such a way that louder harmonics display slightly flatter intonation [cf. 21].

### 4. DISCUSSION

We believe that the patterns for the ultrasound data described in 3.1. provide further evidence for our previous hypothesis of language influence on brass playing, namely, that players whose First Languages include centralized vowels seem to use that position (or close to it) while players who do not have such a vowel use back vowels such as /o/ in Tongan and Japanese (cf. [13]). The patterns for changes in the height of back of the tongue for different pitches displayed by our three participants seem to indicate that changes in the vocal tract to facilitate the production of higher notes can be affected not only by altering the tongue position but also by making changes in the pharyngeal cavity; S5 NZE, whose tongue position changes only minimally, reported that he can feel his pharyngeal cavity widening while ascending throughout the register. Both observations are in agreement with articulatory data for the production of extreme vowels [1] and might provide a direct link between the observed tongue shapes and acoustical output.

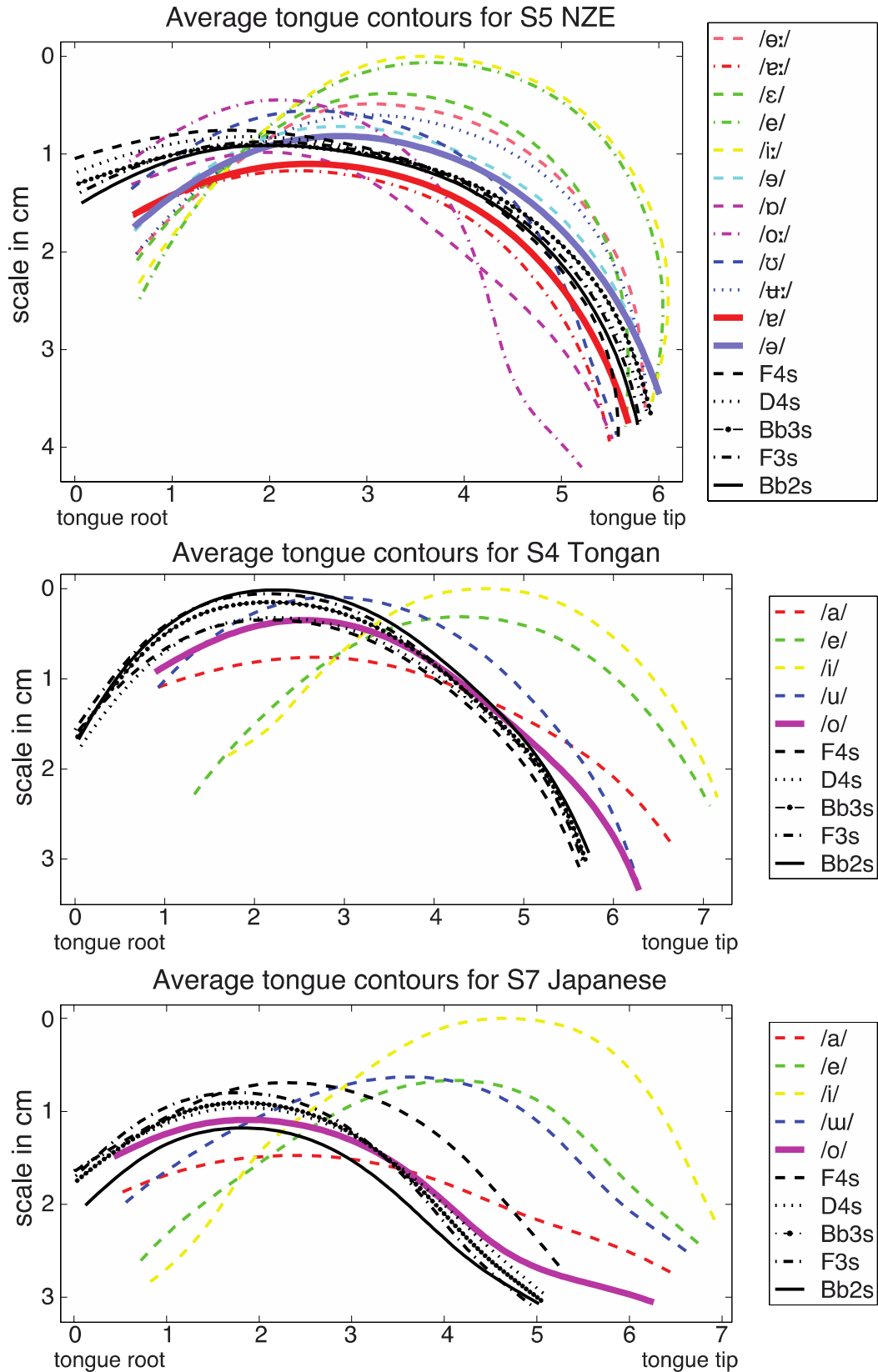
Our acoustic findings support our current hypothesis that variable tongue position (attributed to native language influence) may lead to perceptible acoustic differences as the Japanese player with a higher, more retracted tongue position produced a brighter sound. If we compare average formant values for centralized and medium high back vowels we can see that they differ to a small extent in their F1 and more in their F2 values. In future work, we thus plan to extract average formant data from our speech recordings for the vowels closest to the tongue positions used by the different players to determine whether these show up directly in the sound spectrum produced while playing the trombone.

## 6. ACKNOWLEDGEMENTS

We would like to thank our three participants, the New Zealand Institute of Language, Brain and Behaviour at the University of Canterbury for

supplying the equipment to do ultrasound research, Warwick Music UK for providing a free 'pBone', the University of Canterbury for providing a Doctoral Scholarship for the first author, and Jennifer Hay for providing valuable feedback.

**Figure 1:** Average tongue contours for the S5 NZE, S4 Tongan, and S7 Japanese trombone players



## 7. REFERENCES

- [1] Baer, T., Gore, J. C., Gracco, L. C., Nye, P. W. 1991. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels *J. Acoust. Soc. Am.* 90, 799-828.
- [2] Beauchamp, J. W. 2012. Trombone transfer functions: Comparison between frequency-swept sine wave and human performer input *Archives of Acoustics* 37, 447-454.
- [3] Boersma, P., Weenink, D. *Praat: Doing phonetics by computer* [Computer program]. Version 5.3.52, retrieved 18 May 2014 from <http://www.praat.org/>
- [4] Campbell, M., Greated, C. A. 1987. *The musician's guide to acoustics* New York: Schirmer.
- [5] Carral, S. 2011. Determining the just noticeable difference in timbre through spectral morphing: A trombone example." *Acta Acustica united with Acustica* 97, 466-476.
- [6] Chen, J.-M., Smith, J., Wolfe, J. 2011. Saxophonists tune vocal tract resonances in advanced performance techniques *J. Acoust. Soc. Am.* 129, 415-426.
- [7] Derrick, D., Best, C. T., Fiasson, R. (In Press). Non-metallic ultrasound probe holder for co-collection and co-registration with EMA. *Proc. 18th International Congress of Phonetic Sciences* Glasgow.
- [8] Fletcher, N., Rossing, T. D. 1991. *The physics of musical instruments* New York: Springer.
- [9] Fréour, V. 2013. *Acoustic and respiratory pressure control in brass instrument performance* Ph.D. thesis, McGill Univ., Montreal.
- [10] Gardner, J.T. 2010. *Ultrasonographic investigation of clarinet multiple articulation* D.M.A. thesis, Arizona State Univ., Tempe.
- [11] Gu, C. 2014. *gss: General smoothing splines* <http://CRAN.R-project.org/package=gss>
- [12] Hall, J. C. 1954. *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. thesis, Indiana Univ., Bloomington.
- [13] Heyne, M., Derrick, D. 2014. Some initial findings regarding first language influence on playing brass instruments *Proc. 15th Australasian International Conference on Speech Science and Technology* Christchurch. 180-183.
- [14] Kaburagi, T., Yamada, N., Fukui, T., Minamiya, E. 2011. A methodological and preliminary study on the acoustic effect of a trumpet player's vocal tract *J. Acoust. Soc. Am.* 130, 536-545.
- [15] MathWorks, Inc. 2013. *MATLAB Release 2013b*
- [16] R Core Team 2013. *R: A language and environment for statistical computing* <http://www.R-project.org/>.
- [17] Sinex, D. *Harmonics* [Computer program]. Sent 22 January 2015.
- [18] Tiede, M. *Get Contours* [Computer program]. Sent 1 May 2014.
- [19] Whalen, D. H., Iskarous, K., Tiede, M. K. 2005. The haskins optically corrected ultrasound system (hocus). *Journal of Speech, Language, and Hearing Research* 48, 543-553.
- [20] Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., Sloetjes, H. 2006. ELAN: a professional framework for multimodality research *Proc. 5th International Conference on Language Resources and Evaluation* Genoa. 1556-1559.
- [21] Wolfe, J., Tarnopolsky, A.Z., Fletcher, N. H., Hollenberg, L. C. L., Smith, J. 2003. Some effects of the player's vocal tract and tongue on wind instrument sound *Proc. Stockholm Music Acoustics Conference* Stockholm, 307-310.