THE ROLE OF SOMATOSENSORY FEEDBACK IN THE PRODUCTION OF THE ENGLISH VOWEL /i/ IN FEMALES

Noor Al-Zanoon[1], Angela Cullum[1], Jacqueline Cummine[1], Caroline Jeffery[1,2,3], Bill Hodgetts[1,3], and Daniel Aalto[1,3]
[1]Faculty of Rehabilitation Medicine, University of Alberta, [2]Faculty of Medicine and Dentistry, University of Alberta,[3] Institute for Reconstructive Sciences in Medicine (iRSM)
alzanoon@ualberta.ca, acasson@ualberta.ca, jcummine@ualberta.ca, cxu@ualberta.ca,
bill.hodgetts@ualberta.ca, aalto@ualberta.ca

ABSTRACT

Background. Auditory and somatosensory feedback modulate speech production. In particular, perturbing the somatosensory system has been shown to impact consonant production.
Objective. To investigate the role of somatosensory feedback in the production of the English vowel /i/.
Methods. Thirty-three female, native English speakers were randomly assigned to control and experimental groups. The experimental group received 15 ml of 2% lidocaine mouthwash, whereas the control group received a visually comparable solution without anaesthetic. Participants produced 10 repetitions of four words (/bit, bæt, but, bɑt/) in random order. Formant frequencies (F1, F2) of /i/ were extracted, and analysed separately using a two-way mixed ANOVA.
Results. In the vowel /i/, F1 decreased in the experimental condition compared to the control group.
Conclusions. The results suggest that lidocaine impacts the production of the vowel /i/ due to the reliance on somatosensory feedback.

1. INTRODUCTION

The speech production system is highly complex and involves the coordination of auditory, motor, and somatosensory subsystems [1-5].

Somatosensory feedback plays a crucial role in monitoring and correcting complex motor tasks like hand grip force [6], and gait [7,8]. Similar somatosensory feedback mechanisms are expected to play a role in the speech production system [3-5]. Specifically, the motor system adapts and corrects the somatosensory perturbation caused by the mechanical load, without affecting the auditory output of speech. Mechanical perturbation studies [3-5] provide an indirect measurement of the effect of somatosensory feedback on speech.

Early studies that directly perturb somatosensory feedback using nerve blocks have found observable effects in affricates, sibilants and liquids [9-11]. This observable effect has also been demonstrated in vowels. Niemi et al., [12-14] analysed Finnish vowel productions after anesthetizing participant lingual nerves. Changes in formant values varied across individuals with no specific group level perturbation pattern emerging.

In the present study, the role of oral somatosensory feedback in vowel /i/ is studied by experimentally manipulating the oral sensations, specifically, by topically anaesthetizing the oral cavity. For the vowel/i/, the anterior, lateral and posterior portions physically contact the hard palate [15, 16]. The abundance of biomechanical evidence on the tongue-to-hard palate contact in the vowel /i/ makes the case for an ideal preliminary study on somatosensory feedback. Based on the results from mechanical perturbation studies, it is expected that perturbing the oral sensory system with a numbing agent would lead to acoustic changes. If systematic changes are observed, the potential role of somatosensory feedback in vowel production will be further understood.

2. METHODS

The present study was conducted as part of a larger study on the relationship between reading and speaking. The present paper only reports the methods and results of the speech production task.

2.1. Participants

Thirty-five university students (F=33, M=2) were recruited (mean age= 26.5, SD =4.95) to participate in the study. The study took place at the University of Alberta (Edmonton, Alberta). All participants were native English speakers, reported normal hearing and speech, and no neurophysiological problems. Informed consent was obtained from every participant. The study protocol was approved by University of Alberta Health Research Ethics Board - Health Panel (Pro00068658).

2.2. Stimuli

Recordings of four words (/bit, bæt, but, bɑt/) were made by a 28-year-old female native English speaker with a Southern Ontario Canadian accent.

These words were designed to contain four vowels within a consonant-vowel-consonant sequence (CVC) realized as /i, æ, u/, and /ɑ/. The CVC context was used to keep the phonetic context consistent for the /i/ vowel. The purpose of including four words was to provide variation and reduce a potential repetition effect. Stimuli were recorded using a dynamic microphone (SHURE SM58), and an amplifier (Steinburg UR 22 mkII). The mouth-to-microphone distance was 10 cm. Audacity (2.1.2) was used for recordings, which were sampled at a rate of 44.1 kHz. Once recorded, each sound file was scaled in amplitude to ensure level (dB) was the same across all four words. The time before the onset of each stop consonant was also standardized to 0.05s for all words. To reduce natural variation, the recorded words were repeated ten times and randomized to create stimulus set of 40 productions. The inter-stimulus interval was set to length of each of word in seconds.The F1 and F2 of the stimulus /i/ were 247.63 Hz and 2944.67 Hz respectively (measured as described in section 2.4.1).

2.3. Procedure

2.3.1 Group Assignment

Participants were randomly assigned to two groups, using an online random number generator (https://www.randomizer.org/): control (N=16), and experimental (N=19). In the experimental group, participants were given 15 ml of 2% lidocaine solution. Lidocaine is a sodium-channel inhibitor that acts rapidly to induce anaesthesia lasting up to one hour when applied topically to mucous membranes, such as the oral cavity [17]. To isolate the effect of lidocaine, a control group was given a corn syrup solution similar in colour and consistency to the experimental group which did not contain lidocaine (15 ml in total: 10 ml water, 5 ml corn syrup, two to three drops of food colouring). Participants were instructed to: "*Swish this in your mouth for 60 seconds (I will time you). Then spit it into the sink. Do not swallow the mouthwash*."

2.3.2 Group Assignment

The same recording set up as for speech stimuli (microphone and amplifier) was used for participant recordings. Stimuli were presented binaurally through a set of headphones. The sound quality of recordings was checked using Audacity (2.1.2) at the beginning of each recording session. Participants were given the following instructions: "*Once you hear the word, you will be asked to repeat it right after*." Stimuli were randomized for each participant, presented, and recorded using a custom MATLAB (2018b) script [18]. Recordings took place before and immediately after the mouthwash manipulation.

2.3.3 Degree of Topical Anaesthesia Measurements

Immediately after the second recording block (post mouthwash), participants completed two measurements to determine the degree of topical anaesthesia (i.e. 'numbness'). The first, a visual analogue scale with two endpoints: "no numbness" to "completely numb". Participants were asked to indicate with a tick mark the degree of numbness felt. The second, an image of the oral cavity, where participants were asked to mark anatomical regions in which numbness was felt.

2.4. Data Analysis

2.4.1 Pre-processing

Vowel formants (F1, F2) were extracted using custom code written in MATLAB. The most stable portion of vowel was segmented in Praat (6043, 2018) [19]. A pre-emphasis filter was applied and then Burg's method was used to extract the spectral envelopes of each of the vowels (8192 FFT, order 40). The parameters were selected based on trial and error to broadly match the typical range of formant values for /i/ [20]. F1 frequencies were slightly lower than values in [20] and F2 frequencies were similar to values in [20].

2.4.2 Statistical Analysis

All statistical analyses were carried out using R statistical software (version 3.5.1) [22]. Two separate mixed ANOVAs were conducted for F1 and F2 of the vowel [i], with Time (Pre-and post - mouthwash) as the within subjects factor, and Group (Control, Experimental) as the between subjects factor. Model assumptions were checked using visual inspection and statistical tests of normality and homogeneity of variance. The data showed a positively skewed distribution and was log10 transformed. Standard deviations for each subject were calculated pre and post for F1 and F2. The standard deviations were comparable for both groups (paired t-test, ns). The two male participants were excluded from the analysis. Outliers (values >3 standard deviations) were removed by trial rather than by participant. Post hoc tests were performed with Bonferroni correction (alpha=0.017). Effect size was measured using Cohen's d.

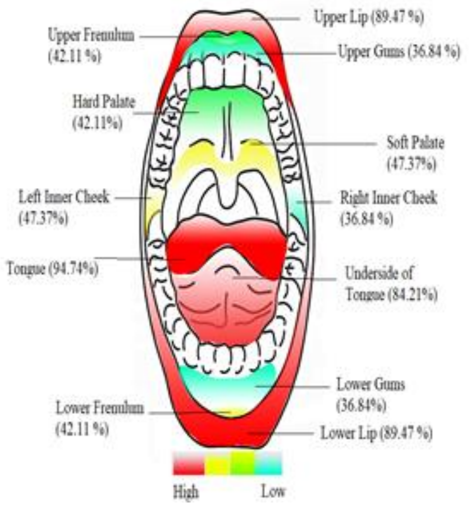
3. RESULTS

3.1. Numbness measurements

3.1.1 Self-perceived numbness scale

In the control group, the mean numbness rating was 0.51/10, with a range of 0-0.21. In the experimental group, the mean numbness rating was 6.25/10 with a range of 2.1-9.3. This difference was statistically significant (t-test, p<0.001). Figure 1 shows the percentage of oral cavity regions that were perceived as numb by participants in the experimental group. Highest rates of numbness were reported for the tongue and lips.

Figure 1: Percentage of self-reported numbness by oral cavity region. Figure is adapted from: OpenStax college, obtained from Anatomy & Physiology, Connexions, http://cnx.org/content/col11496/1.6/.

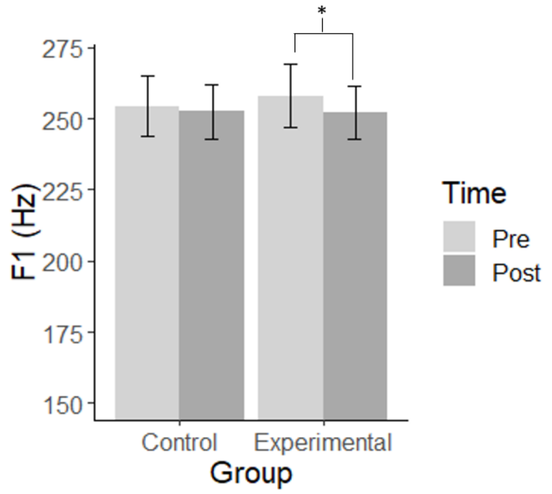


3.1.2

Formant Results

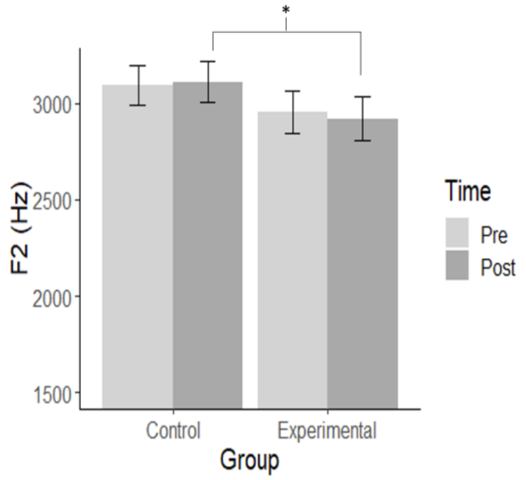
ANOVA analysis for F1 showed no significant main effects. However, there was a significant interaction for Group x Time (F (1, 32) =5.49, *p*=0.03, Cohen's *d*=0.336). Post-hoc tests showed no difference in control group (paired t-test, *p*=0.46, Cohen's *d*=0.005) whereas in the experimental group, F1 decreased after lidocaine exposure (paired t-test, *p*=0.01, Cohen's *d*=0.338). The results are depicted in Figure 2 in raw units although the comparisons were made in log-transformed units.

Figure 2: Mean F1 divided by group and time.



F2 showed only a significant main effect of Group (F= (1,32) = 6.37, *p*= 0.02, Cohen's *d*=0.196). Post hoc tests showed no difference between control and experimental groups before mouthwash (*p*=0.055, d=0.632). Finally, a significant difference was found between the control and experimental groups after mouthwash (p=0.014, Cohen's d= 0.802). The results are visualized in Figure 3.

Figure 3: Mean F2 divided by group and time.



3.1.3 Additional Analyses

To further check data quality, F1 and F2 of the stimuli were compared to participant formants. In the recordings before the mouthwash, the participants trended towards a higher F1 compared to the stimulus target (unpaired t-test, p=0.06). The pre-mouthwash F2 values of both groups were substantially higher than the stimulus target (unpaired t-test, $p<10^{-15}$).

4. DISCUSSION

The results of this study show that perturbation of the intraoral somatosensory system modulates F1 of the English vowel /i/. However, the way in which this perturbation modulates both F1 and F2 is subject to further investigation.

Congruent with earlier findings [12-14], we observed changes in formants after numbing the oral cavity. As opposed to the findings in [12-14], our findings indicate systematic changes in F1 at the group level. Since earlier studies have largely used a case series design, the present results may reflect an increase in statistical power showing the relatively weaker group level effects. In addition, the present study used topical anaesthetic as opposed to nerve blocks used in earlier studies potentially inducing more congruent compensatory behaviours [10-14].

While articulatory changes underlie the observed changes in formants, they were not directly measured. Formant extraction is an error-prone, indirect measurement of the biomechanical configurations of the vocal tract [22]. Extant experimental evidence does not allow for a direct prediction of acoustic related changes based on biomechanics [23, 24]. Therefore, direct kinematic measurements would be required to understand changes underlying tongue position. Based on perturbation analyses of vocal tract shapes and corresponding computed formant frequencies, the observed difference in formant patterns during /i/ may be explained by the tongue being bunched slightly further along hard palate towards the soft palate; thereby lengthening the narrow air column between tongue and palate in the experimental group compared to the control group [25-30]. In the experimental group, the tongue could be pushed pushed harder against the palate to restore oral somatosensory feedback.

For F2, statistical analyses showed only a significant main effect for Group in F2. Post hoc tests only reached significance when comparing post F2 values between the control and experimental groups. Surprisingly, the control group seemed to have increased F2 values after swishing with the mouthwash while the experimental group seemed to have a decreased F2. However, a systematic lowering or raising pattern was not found for F2. This may be explained by natural variation in the way the experimental and control groups imitated the stimulus target. Statistical analyses showed that both groups were statistically different in their F2 productions before the lidocaine manipulation. Future studies using greater sample sizes may reveal a more systematic effect of lidocaine. Another possible explanation could be a confound effect by using auditory stimulus target. As opposed to F1, the F2 in the target stimulus was lower than the average participant F2 productions. Between-subject variability in vowel formants is greater in read words than when imitating an acoustic target [31]. Hypothetically, the participant-specific tongue-palate target for /i/ and the acoustical target may have been in conflict. This could have led to relative changes in the somatosensory and auditory weights in monitoring the production. Future studies are required to dissociate the interplay of auditory and somatosensory feedback in vowel production.

Finally, in this study, topical anaesthesia of the oral cavity was achieved using a lidocaine mouthwash solution. While the questionnaires show congruent patterns within each participant group (control, experimental) in terms of degree of anaesthesia, the areas (fig 1) were based on self-reported region markings. These may not include all the areas that were affected by numbing agent. Specific regions can only be approximated, and we can only assume numbing of the anterior part of the oral cavity. In future studies, the impact of the numbing agent could be verified with stereognostic or two-point discrimination tests. Furthermore, on a physiological basis, the tongue consists of superficial and deep sensory receptors that are used to detect changes in touch, temperature, and taste [32]. Topical lidocaine may only inhibit the superficial receptors.

5. CONCLUSION

Perturbation of somatosensory feedback modulates the production of /i/ as observed through the lowest two formants. F1 was lowered in the experimental group after exposure to lidocaine. A systematic change was not found for F2. Three possible mechanisms could underlie the results: 1) natural variation within the sample; 2) stronger reliance on auditory feedback; and/or 3) the tongue being pushed harder towards the palate to compensate for reduced somatosensory feedback.

6. ACKNOWLEDGEMENTS

The authors would like to thank Cassie Tam, Chris Parbery, Aimee-Jeanne Garthus, and Tyson Sereda for helping with participant recruitment and data collection.

7. REFERENCES

[1] MacDonald, E., Purcell, D., Munhall, K. 2011. Probing the independence of formant control using altered auditory feedback. *JASA*, 129, 955-965.

[2] Ghosh, S., Matthies, M.., Maas, E., Hanson, A., Tiede, M., Ménard, L., ... Perkell,S. 2010. An investigation of the relation between sibilant production and somatosensory and auditory acuity. *JASA*, 128, 3079-3087.

[3] Tremblay, S., Shiller, D., Ostry, D. 2003. Somatosensory basis of speech production. *Nature*, *423*, 866.

[4]McAuliffe,J., Robb, M., Murdoch, B. 2007. Acoustic and perceptual analysis of speech adaptation to an artificial palate. *Clin Linguist Phon*, 21, 885-894.

[5] Baum,S. McFarland, D .2000. Individual differences in speech adaptation to an artificial palate. *JASA*, 107, 3572-3575.

[6] Nowak, D., Glasauer, S., Hermsdörfer, J. 2004. How predictive is grip force control in the complete absence of somatosensory feedback?. *Brain*, 127, 182-192.

[7] Pearson, K. G. (2004). Generating the walking gait: role of sensory feedback. In *Progress in brain research*. 143, 123-129.

[8] Nielsen, J., Sinkjær, T. 2002. Afferent feedback in the control of human gait. *Journal of electromyography and kinesiology*, 12, 213-217.

[9] Schliesser, H. F., & Coleman, R. O. 1968. Effectiveness of Certain Procedures for Alteration of Auditory and Oral Tactile Sensation for Speech. *Perceptual and Motor Skills*, 26, 275–281.

[10] Borden, G. 1976. The effect of mandibular nerve block upon the speech of four-year-old boys. *Language and speech*, 19, 173-178.

[11] Borden, G, Harris, K., Oliver, W. 1973. Oral feedback, part I: Variability of the effect of nerve-block anesthesia upon speech. Haskins Laboratories Status Report on Speech Research SR-34 (this issue).

[12] Niemi, M., Laaksonen, J., Aaltonen, O., Happonen, R. 2004. Effects of transitory lingual nerve impairment on speech: an acoustic study of diphthong sounds. *J. Oral Maxillofac. Surg*, 62, 44-51.

[13] Niemi, M., Laaksonen, J., Vähätalo, K., Tuomainen, J., Aaltonen, O., Happonen, R. 2002. Effects of transitory lingual nerve impairment on speech: an acoustic study of vowel sounds. *J. Oral Maxillofac. Surg*, 60, 647-652.

[14] Niemi, M., Laaksonen, J., Forssell, H., Jääskeläinen, S., Aaltonen, O.,Happonen, R. 2009. Acoustic and neurophysiologic observations related to lingual nerve impairment. *Int J Oral Max Surg*, 38, 758-765.

[15] Gick, B., Allen, B., Roewer-Després, F., Stavness, I. 2017. Speaking tongues are actively braced. *JSLHR*, 60, 494-506.

[16] Stone, M., Lundberg, A. (1996). Three- dimensional tongue surface shapes of English consonants and vowels. *JASA*, 99, 3728-3737.

[17] Catterall, W., Mackie, K. 2011. Local anesthetics. In: *Goodman & Gilman's the pharmacological basis of therapeutics.* McGraw-Hill: New York (NY), 565-582.

[18] MATLAB and Statistics Toolbox Release 2018b, The MathWorks, Inc., Natick, Massachusetts, United States.

[19] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer [Computer program]. Version 6.0.43, retrieved 8 September 2018 from http://www.praat.org/

[20] Hillenbrand, J., Getty, L. A., Clark, M. J., Wheeler, K. 1995. Acoustic characteristics of American English vowels. *JASA*, 97, 3099-3111.

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

[22] Shadle, Christine H., Hosung, N., Whalen, D. 2016. Comparing measurement errors for formants in synthetic and natural vowels. *JASA*, 139, 713-727.

[23] Lindblom, B., Sundberg, J.(1971). Acoustical consequences of lip, tongue, jaw, and larynx movement. *JASA*, 50, 1166-1179.

[24] Lee, J., Shaiman, S., Weismer, G. (2016). Relationship between tongue positions and formant frequencies in female speakers. *JASA*, *139*(1), 426-440.

[25] Mermelstein, P. 1967. Determination of the vocal-tract shape from measured formant frequencies. *JASA*, 41, 1283-1294.

[26] Heinz, J.1967. Perturbation functions for the determination of vocal-tract area functions from vocal-tract eigenvalues. *STL-QPSR*, 8, 001-014.

[27] Story, B. 2007. A comparison of vocal tract perturbation patterns based on statistical and acoustic considerations. *JASA*, 122, 107-114.

[28] Fant, G. 1960. *Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations(No. 1).* Walter de Gruyter.

[29] Fant, G. 1970. *Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations(No. 2)*. Walter de Gruyter.

[30] Fant, G. 1975. Non-uniform vowel normalization. *STL-QPSR*, 1, 1-19.

[31] Broad, D. 1976. Toward defining acoustic phonetic equivalence for vowels. *Phonetica*, 33(6), 401-424.

[32] Fuller, D.R., Pimentel, J.Peregoy, B. 2012. *Applied Anatomy & Physiology for Speech-language Pathology & Audiology.* Wolters Kluwer: Baltimore