Fall 2014

Brittany Bernal - Sensorimotor Adaptation of Speech Through a Virtually Shortened Vocal Tract

Brittany A. Bernal
Marquette University

Recommended Citation

Bernal, Brittany A., "Brittany Bernal - Sensorimotor Adaptation of Speech Through a Virtually Shortened Vocal Tract" (2014). *Ronald E. McNair Scholars Program 2014*. 10.
https://epublications.marquette.edu/mcnair_2014/10

Sensorimotor Adaptation of Speech Through a Virtually Shortened Vocal Tract

Brittany Bernal

PI: Jeffrey Berry, Ph.D., CCC-SLP

Marquette University

McNair Scholars Program

Abstract:

The broad objective of this line of research is to understand how auditory feedback manipulations may be used to elicit involuntary changes in speech articulation. We examine speech sensorimotor adaptation to supplement the development of speech rehabilitation applications that benefit from this learning phenomenon. By manipulating the acoustics of one's auditory feedback, it is possible to elicit involuntary changes in speech articulation. We seek to understand how virtually manipulating participants' perception of vowel space affects their speech movements by assessing acoustic variables such as formant frequency changes. Participants speak through a digital audio processing device that virtually alters the perceived size of their vocal tract. It is hypothesized that this modification to auditory feedback will facilitate adaptive changes in motor behavior as indicated by acoustic changes resulting from speech articulation. This study will determine how modifying the perception of vocal tract size affects articulatory behavior, indicated by changes in formant frequencies and changes in vowel space area. This work will also determine if and how the size of the virtual vowel space affects the magnitude and direction of sensorimotor adaptation for speech. The ultimate aim is to determine how important it is for the virtual vowel space to mimic the talker's real vowel space, and whether or not perturbing the size of the perceived vowel space may facilitate or impede involuntary adaptive learning for speech.

Introduction:

The purpose of this research is to evaluate the effectiveness of auditory feedback manipulations used to elicit involuntary changes in speech articulation. Previous research in other fields indicates that adaptation occurs when there is a change in movement based on a perceived sensory error (Bastian, 2008). When individuals make involuntary changes in speech articulation due to perceived speech errors, it is referred to as sensorimotor adaptation. Sensorimotor adaptation of speech is the learning phenomenon that is evaluated in the current study. By manipulating the acoustics of one’s auditory feedback in experimental conditions, it is possible to elicit involuntary changes in speech articulation (Houde & Jordan, 1998, 2002). The altered speech patterns can be made to persist for some time, even after the signal modifications end, which suggests that speech can be re-learned (Perkell, 2012). The experimental methodology used in this research study may be further developed to help rehabilitate individuals with motor speech disorders, such as dysarthria, who do not benefit from traditional, voluntary therapy techniques. This is hypothesized because currently there are no effective treatments for those with severe motor speech disorders; which is partly due to the therapy techniques requiring voluntary modifications to articulation, which may be ineffective. Additionally, effective rehabilitation strategies for other sensorimotor skills such as walking have utilized adaptation paradigms (c.f., Bastian, 2008).

The goal is to understand how virtually manipulating subjects’ vowel space affects their speech movements by assessing formant frequency changes. The vowel space is the acoustic space defined by the first and second resonance frequencies, called formants (F1 and F2) of the four vowels that define the extreme points of tongue articulation (/i/, /æ/, /u/, and /ɑ/) (Kent & Read, 1992, 2002). Every other vowel in English is within this space.

The acoustic vowel quadrilateral can be used to describe the vowel working space, which is a key component of the current study. The size of a vowel space can help quantify speech intelligibility in various disorders. It has been established that the vowel space is reduced in several adult speech disorders (Turner, Tjaden, & Weismer, 1995; Ziegler & von Cramon, 1983). The reason for reduced vowel space in some speech disorders is that a smaller acoustic space indicates a constricted articulatory space. If one's articulatory space is constricted, this implies less range of movement for key articulators such as the tongue, lips, and jaw (Kent & Read, 2002). Reduced range of movement results in reduced perceptual contrast between different speech sounds, affecting the intelligibility of speech.

Having subjects speak through a virtually shortened vocal tract size (as compared to their actual vocal tract size) may be facilitative to adaptive motor behavior indicated by acoustic changes in speech articulation. One consideration for having participants speak through a virtually shortened vocal tract would be that a smaller vocal tract correlates with a larger acoustic space, and manipulating the acoustic vowel space would likely create changes in vowel production as indicated by changes in formant values. Some studies indicate that subjects with greater impairments tend to make greater progress with virtual reality intervention as compared to typically functioning individuals (Fluet & Deutsch, 2013). In order to virtually manipulate participants' vocal tract sizes, the TC Helicon VoiceWorks Plus© hardware was used. The TC Helicon is a speech signal manipulation hardware that virtually shifts the size of the subjects' vocal tracts by scaling formant frequencies based on a formant parameter setting. In the current study, each parameter setting, ranging from values of 0 to 50, served as a completely novel vocal tract through which the participants spoke. The intention was to determine which parameter values on the TC Helicon elicit the greatest amount of adaptive change in speakers, which is indicated either by changes in the talker's true formant values or changes in vowel space area. Furthermore, the degree to which the size of the new vowel space affects the magnitude and direction of compensatory and adaptive speech behaviors was evaluated. Ultimately, the objective was to understand if this type of speech adaptation study can generalize to disordered speakers' normal everyday speech.

The current hypothesis is that when participants' vowel space area is perturbed to be bigger than their actual vowel space area (which is the acoustic correlate of a shortened vocal tract), they will respond by lowering their formant frequency values and reducing their vowel space area. This hypothesis is due to the correlation between bigger vowel space area and higher formant frequency values; thus compensatory learning behaviors would be indicated by participants changing their own formant values and vowel space area in the opposite direction of the perceived vocal tract perturbation.

Review of the Literature:

Sensorimotor adaptation is the foundation of this study due to its link with potential rehabilitative applications for those with motor speech disorders. Sensorimotor adaptation is a form of involuntary, short-term sensorimotor learning. Adaption for all types of motor learning consists of a nervous system response in which a change occurs in movement based on sensory feedback errors. These sensory prediction errors are discrepancies between the brain’s expected outcome of a movement and the actual sensory consequences of that movement (Bastian, 2008).

Houde & Jordan (1998) developed a method for measuring and altering formant patterns in real time. Formant patterns are determined by vocal tract shape during speech, and varying shapes creates different vocal tract resonances (Kent & Read, 1992). Houde and Jordan’s experimental method of altering formant patterns generates a synthetic version of speech, which allows a speaker to hear the manipulated feedback in real time. Houde and Jordan’s work formed a foundation for the current study, which generated synthetic speech, creating auditory feedback perturbations for participants.

The current study focused on the acoustic consequences of large-scale auditory feedback manipulations. Acoustic changes that participants exhibited were evaluated throughout the experiment by measuring their formant frequencies across varying TC Helicon parameter values. Formants are acoustic resonance patterns measured in hertz that reflect positions of articulators during the production of speech sounds. The two lowest frequency formants (F1 and F2) are the main quantitative measures in this study. The first two formants are sufficient to measure because they have the greatest impact on the acoustics and perception of vowels (Hixon, Weismer, & Hoit, 2008). In a previous study in which participants were asked to recognize another’s voice when exposed to recorded vocal stimuli, results indicated that both the formants of F1 and F2 heavily contribute to voice recognition (Xu, et al., 2013). Thus, tracking the resulting acoustics of these manipulations provides key insights into articulatory behaviors of individuals in the current study.

In experimental conditions, the formant patterns can be shifted to make speakers think that they are producing the wrong vowel or consonant. Manipulating a speaker's formants in real time typically elicits compensatory changes in the positions of the articulators such as the tongue, lips, and jaw (Houde & Jordan, 1998). Results from Houde and Jordan (2002) indicate that some participants in these sensorimotor adaptation experiments do indeed adapt their speech to change articulatory positions, and that these changes can continue, suggesting short-term, involuntary motor learning. Houde and Jordan's work supports the idea of sensorimotor adaptation of speech as a relationship between articulatory movement patterns and auditory feedback. The current work will expand upon Houde and Jordan's work not by manipulating individual formant frequency values, but rather by manipulating the entire acoustic working space as a whole through each virtual vocal tract parameter. In perceptual terms, acoustic manipulations in the current work make a talker's voice sound like another person.

The goal of the current research is to assess speech motor control in healthy speakers in order to better understand adaptation as a tool for rehabilitation for those with motor speech disorders. Motor speech disorders such as dysarthria, which is the prospective clinical focus of this study, are characterized by paralysis, incoordination, or reduced range of motion in the muscular control necessary to produce intelligible speech. This is manifested in symptoms such as slowness or incoordination of the speech mechanism caused by nervous system damage. Different sub-types of dysarthria are due to damage to specific parts of either the central or peripheral nervous system (Darley Aronson, & Brown, 1969).

There are various etiologies associated with dysarthria, the primary ones being traumatic brain injury (TBI), stroke, dystonia, and degenerative diseases such as multiple sclerosis (MS), amyotrophic lateral sclerosis (ALS), and Parkinson’s Disease (Dysarthria, 2014). Because there is so much variability among different types and severities of dysarthria, it is a clinically challenging population to address and treatment methods are still being developed.

Currently, there is a large gap in the literature regarding successful treatment methods for dysarthria. The current study sought to gain a better understanding of speech adaptation for healthy speakers. This goal was achieved by comparing information about how participants used their articulators to produce speech under normal conditions and how they produced an acoustic signal when speaking through a virtually shortened vocal tract. Assessing these factors let us have a better understanding of how feasible it is to create changes in articulation when speaking through a virtual vocal tract. Future work along this line of research will be utilizing the same experimental methods for those with dysarthria. Such research may be a stepping-stone for developing successful treatments for this population. Thus, the current work will evaluate the importance of the perceived vocal tract size in eliciting adaptation. This knowledge will then be applied to the RASS system (Rehabilitative Articulatory Speech Synthesizer) (Berry, North, Meyers, & Johnson, 2013). The RASS system places sensors on the articulators, has talkers produce speech movements without requiring an acoustic signal, and subsequently manipulates the acoustic output of the virtual vocal tract, which has potential to elicit speech adaptation. The RASS system previously could not match the size of a talker’s vowel space. Thus, the current work will address the implications of this technological limitation.

Speech is considered to have both auditory and somatosensory targets, which are predictions formulated by the brain about the outcome of speech sounds. Auditory targets are defined by formant frequencies, whereas somatosensory targets are defined by articulator positions (Tourville & Guenther, 2011). The forward model is a significant component of the current study. The forward model is the brain's estimation of the sensory consequences of a motor command (Christoffels, et al., 2011). This means that before speech is even produced, the brain creates sensory targets, which are expectations of how the sound will be perceived acoustically as well as how it will feel to physically produce that sound in one's vocal tract (Tourville & Guenther, 2011). Thus, manipulating auditory feedback in this study causes subjects to recognize an error in their speech production and may lead to a remapping of their motor plan for forward control of speech output.

One hypothesis states that neural structures such as the basal ganglia and the cerebellum (which may be damaged in various types of dysarthria), contribute to a talker's ability to process sensory information and affect the ability to execute precise and intelligible speech (Kent, Kent, Weismer, & Duffy, 2000). This may imply that those with dysarthria have "target regions" for speech sound production that are larger than the target regions of healthy speakers, which signifies that their brain no longer efficiently registers errors in articulation. This error recognition deficit indicates that when experiencing auditory feedback perturbations, those with dysarthria have less competition with their old target regions, meaning that they are more likely to re-map new target regions for speech sounds. This concept serves as the justification for potentially using auditory feedback perturbations as a form of therapy for those with dysarthria.

As of now, there are no treatment approaches that use speech adaptation as a rehabilitative approach for dysarthria. One plan of assessing speech behaviors that was implemented was measuring participants' vowel spaces and mapping out their targets for various speech sounds. Typically, those with dysarthria, specifically associated with ALS, are perceived to have imprecise vowel productions. This is likely due to the characteristics of incoordination and motor planning as well as motor programming deficits (Kent, et al., 2000). These deficits are likely due to physiological damage to the orofacial muscles required for speech production or limitations in the range and speed of articulator movements (Turner, Tjaden, & Weismer, 1995). Thus, gaining a greater understanding of how vowel space influences auditory feedback driven articulatory learning in healthy speakers may serve as a stepping-stone towards better understanding treatment options for those with dysarthria.

A major problem that persists in those with dysarthria is that their target region for vowel production is too large. Too large of a target region means those individuals perceive errors in articulation as acceptable, due to an inability to produce fine phonemic contrasts. The target space is the acceptable articulatory posture for a particular speech sound, whereas the vowel space is the acoustic working space for all vowels. Typically, those with ALS as well as those with closed head trauma or cerebellar lesions are known to have a reduction in their vowel space (Turner, Tjaden, & Weismer, 1995). Smaller vowel spaces are associated with less intelligibility, meaning that these types of patients are harder to understand when speaking. Thus, artificially changing the perceived size of one's vowel space may have potential for rehabilitative purposes in this population. When those with dysarthria speak through a virtual vocal tract, there is likelihood for them to re-map their own vowel spaces in order to form more precise articulatory targets.

Currently, there are various methods used when seeking to treat dysarthria. A compensatory treatment approach that has been implemented is having patients slow down the rate of their speech. Another compensatory approach for those with dysarthria is having patients increase their loudness during speech production (Tjaden & Wilding, 2010). However, these methods, along with others, may help facilitate communication but may be contrary to rehabilitation, since they do not exploit sensorimotor learning principles. The hope for the current work is that utilizing sensorimotor adaptation as a rehabilitation tool will create effective changes in articulation for those with dysarthria with less conscious effort than traditional therapy techniques.

Methodology:

The current work focuses on the issue of interactions between articulation and voice. Participants in this study experienced a virtual voice in order to define the relationship between adapting to a virtual voice and changes in articulatory speech behaviors.

The TC Helicon VoiceWorks Plus$^{©}$ speech signal manipulation hardware allowed the manipulation of the acoustics of a talker's speech in real time. The talker heard this modified speech acoustic signal via headphones. The TC Helicon virtually manipulated the size of the subjects' vocal tracts across different formant parameter values. The acoustic manipulations caused subjects to perceive themselves to be speaking through a vocal tract of decreased size. Each parameter on the TC Helicon creates a global recalibration of the speech sensorimotor control system. Vocal tract size corresponds with the size of the speaker; thus, small children have much smaller vocal tract sizes than fully-grown adults. The size of the vocal tract is inversely correlated with one's vowel space. Thus, those with smaller vocal tract sizes have larger vowel spaces, and vice-versa (Turner, Tjaden, &Weismer, 1995). The size of one's vocal tract determines the formant working space of the subject. To artificially perturb a speaker's vocal tract size, there was a manipulation of the formant parameter setting on the TC Helicon that ranged from -50 to +50. The 0 dial setting is a subject's speech output without any apparent perturbation to his or her speech signal. According to previous research with virtual environments, the more convincing the virtual environment, the greater the motor response (Wright, 2014). By increasing the formant parameter value in the negative direction to the -50 setting, this virtually shifts a subject's vocal tract to be smaller (making him or her sound more like a child). By increasing the formant parameter value in the positive direction, this shifts the subject's perceived vocal tract size to be larger. In creating synthetic acoustic manipulations in the speech, this method causes the talker to perceive a completely altered acoustic working space. The talkers thus had to learn an unfamiliar mapping between their articulatory movements and the resulting auditory feedback that they heard themselves saying. This method served as a global recalibration because it had effects on multiple acoustic cues for all speech sounds produced.

In order to set up the experiment, participants' vowel spaces at each TC Helicon parameter were recorded first. This phase was labeled "No TC Helicon Feedback." During this phase, participants were neither wearing earphones nor a bone conduction vibrator. Participants were asked to say the words /hid/, /hæd/, /hud/, and /hɑd/; and the phrase “I owe you a yo-yo”, repeated a total of five times at each parameter setting. The reason these words were chosen was because the /h/ phoneme has essentially no acoustic significance, because it is a glottal sound with little impact on the acoustics of surrounding speech sounds. The goal was to be able to easily track the changes in vowel formants without any coarticulatory influence of surrounding consonant sounds. The phrase “I owe you a yo-yo” was used because it is a sentence that consists entirely of vowels. In this experiment the acoustic changes were not tracked in the phrase; however, the plan is to use it for a continuation of this study in the future. Participants’ vowel spaces were measured from each parameter setting on the TC Helicon which they heard their own auditory feedback, and their speech was recorded from the 0 setting all the way to negative 50, in intervals of 5. The purpose in doing this was to determine a reference point for the virtual vocal tract without any perturbed sensory input.

After the phase with no auditory feedback from the TC Helicon, a loudness test was performed for each participant. During this phase, participants were wearing a headset with a microphone that was placed at the corner of their mouth, as they did in the previous phase. However, for the loudness test, participants were then asked to wear earphones. In addition, subjects wore a bone conduction vibrator. The purpose in doing this was to prevent subjects from hearing their own bone conduction, in order to create a false bone conduction for the voice that they would be hearing through the TC Helicon. However, there are believed to be some effects to the output signal under the 0 setting due to the fact that subjects cannot hear their own voice through bone conduction, as they would typically. Again, participants were asked to say the words /hid/, /hæd/, /hud/, and /hɑd/; and the phrase “I owe you a yo-yo”, repeated a total of one time. They were then asked to read The Bamboo Passage (Green, Beukelman, & Ball, 2004), which is one of the standard passages that are read in speech studies. The Bamboo Passage was read during the loudness test in order to ensure that participants would not hear themselves while producing connected speech. While subjects were speaking, they indicated to the experimenters whether they reached a loudness in which they could no longer hear their own auditory feedback at the loudest comfortable setting.

Following the loudness test, subjects' vowel spaces were recorded at Baseline, which was the 0 setting on the TC Helicon. This was a condition in which the TC Helicon did not create any acoustic manipulations. Formant frequency values for the vowel repetitions from this phase were measured acoustically. During Baseline, subjects said the words /hid/, /hæd/, /hud/, and /hɑd/; and the phrase "I owe you a yo-yo", repeated a total of five times. They were also prompted to say /hɔ͞ɪ/, /hɑ͞ʊ/, and /he/. These three words were added to this phase in order to test generalization. Generalization was noted if subjects changed their articulation when producing these vowel sounds (/ɔ͞ɪ/, / ɑ͞ʊ/, and /e/) that they never said under the experimental condition of a virtually shifted vocal tract. Testing generalization is important because it serves as an indication that speech adaptation may carryover to different phonetic contexts that were not trained during the perturbed auditory feedback. When generalization occurs, it suggests that adaptation as a form of speech rehabilitation is possible. If generalization does not occur, this means that essentially all linguistic contexts would have to be trained individually. Thus, generalization is useful because it makes rehabilitation more efficient; however, rehabilitation is still possible even if generalization does not occur. The purpose of the Baseline phase was to provide a reference point that is nearly equivalent to participants’ pure voices in order to measure the learning process through each repetition.

The next phase, called “Two Channel TC Helicon Feedback” consisted of participants saying the words /hid/, /hæd/, /hud/, and /hɑd/; and the phrase “I owe you a yo-yo”, repeated a total of five times. After each repetition, subjects were asked to read a short paragraph of either the “Stella” (Kunath & Weinberger, 2010), “Caterpillar”, or “Bamboo” (Green, Beukelman, & Ball, 2004) passages. These are passages that were created for speech studies such as this one; the “Caterpillar” in particular was designed to assess motor speech disorders (Patel, et al., 2013). Each passage is considered to have a phonetically balanced context. The purpose in having participants read these passages during each setting of the Two Channel Helicon Feedback phase was to “adapt” them to the new virtual vocal tract that they were speaking through. These passages were edited slightly so as to not contain any of the generalization words in the current study. The subjects’ actual voice without the TC Helicon perturbation was also recorded in a separate channel. The purpose of this phase was to quantify the learning process of the perturbation in both the real and virtual domains.

Following the Two Channel Helicon Feedback phase was the Masking Phase. During the Masking Phase, participants said the words /hid/, /hæd/, /hud/, and /hɑd/; and the phrase “I owe you a yo-yo”, repeated a total of five times. They were also prompted to say the generalization words /hɔ̄ɪ/, /hɑ̄ʊ̄/, and /he/. During this phase, participants heard white noise played into their earphones. The white noise was tuned to the frequency range of human speech. White noise was played in order to prevent subjects from hearing their own auditory feedback. The overall purpose of the Masking Phase was to determine if adaptation and generalization occur while the sensory input of the perturbation is absent.

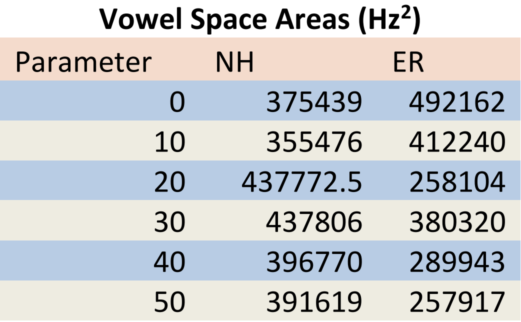
The final phase of the experiment was De-adaptation. During this phase, participants were asked to say the words /hid/, /hæd/, /hud/, and /hɑd/; and the phrase “I owe you a yo-yo”, repeated a total of five times. They were also prompted to say /hɔ̅ɪ/, /hɑ̅ʊ̅/, and /he/. This phase was to measure a “second baseline” following each vocal tract shift. Thus, this phase was recorded at the 0 setting on the TC Helicon, meaning there was no apparent perturbation to the participants’ auditory feedback. The purpose of this phase was to quantify the amount of time it takes the subject to return to his or her baseline vowel space area and assure that the next experimental cycle would not be affected by any adaptation that occurred during the current one.

Each of the phases (Baseline, Two Channel TC Helicon Feedback, Masking, and De-adaptation) was repeated at each parameter setting of the TC Helicon. The reason in doing so was to measure how subjects adapted to each individual vocal tract shift.

Various questions were addressed with this experimental design. Results were compared between subjects and within subjects. The amount of acoustic articulatory change that took place with the magnitude of the shift parameter of the TC Helicon was evaluated.

Results:

In order to measure changes in articulation, tables were created indicating changes in vowel space area values across different parameters on the TC Helicon, and this quadrilateral circumscribed by the corner vowels was calculated using a convex hull method. This was done in order to assess how the subjects’ own formant values changed across different parameters. The goal was to identify which settings on the TC Helicon were most effective in creating adaptive articulatory behaviors. Figure 1, found below, indicates vowel space areas across parameters ranging from 0 to 50 for two subjects, NH and ER.



Vowel Space Areas (Hz^2)

Parameter	NH	ER
0	375439	492162
10	355476	412240
20	437772.5	258104
30	437806	380320
40	396770	289943
50	391619	257917

Figure 1: Vowel Space Areas, participants NH and ER.

Vowel space areas serve as an overall measure of the acoustic working space. Participant NH has a tendency to start with a small vowel space area and then increase her vowel space area as the perceived vocal tract size shortens. This indicates a following, or mimicking response, which is contrary to the expectation. Participant ER, however, tends to start with a large vowel space area and decreases the size of her own vowel space area, indicating a compensatory response, which is consistent with the expectation.

Another method that was used to analyze this data was creating vowel spaces indicating the change in F1 and F2 space. Simply put, F1 corresponds inversely with tongue height and F2 corresponds directly with tongue forwardness. Thus, an increase in F1 (located on the x-axis) indicates a lowering of the tongue, and an increase in F2 (located on the y-axis) indicates moving the tongue farther forward in the vocal tract. These two variables were plotted on an X-Y coordinate system. Coordinates indicate where subjects typically produce their formant values for specific vowel sounds at each TC Helicon formant parameter value, indicated in the legend at the bottom right.

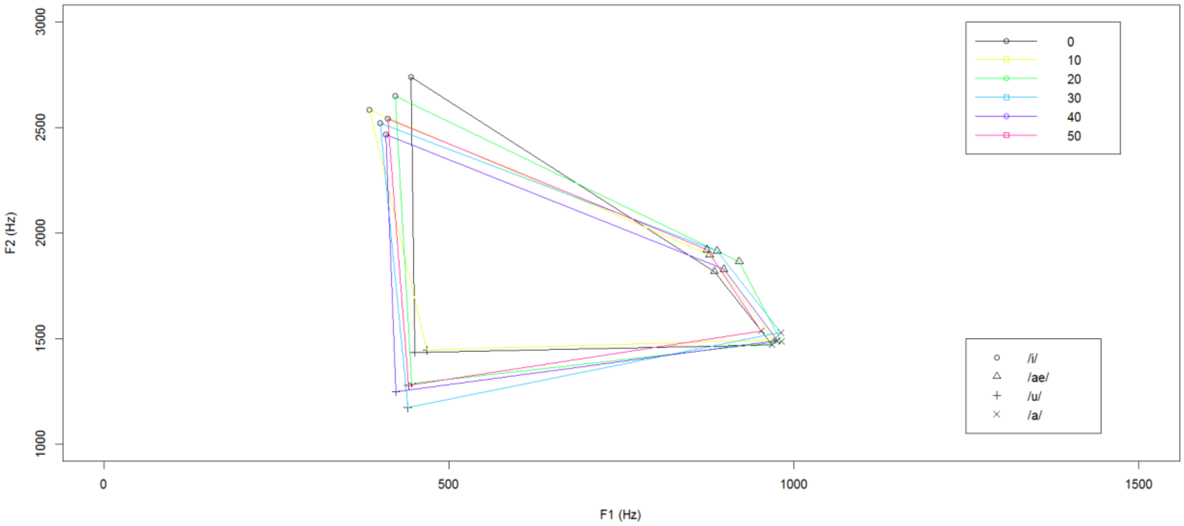


Figure 2: Vowel Space by Parameter for Participant NH.

Figure 2 shows the acoustic vowel spaces for baseline (0) and masking conditions at different parameter values (10-50). Each parameter serves as a different vocal tract size. This comparison assesses involuntary (adaptive) learning. NH demonstrates vowel specific changes, primarily /i/ and /u/, that correspond with movement of the tongue up and backward as the perceived vocal tract shortens. This trend is not consistent, since extreme values appear to move back towards baseline, suggesting that when the perceived vocal tract is extremely different from the talker’s, adaptive changes may be reduced.

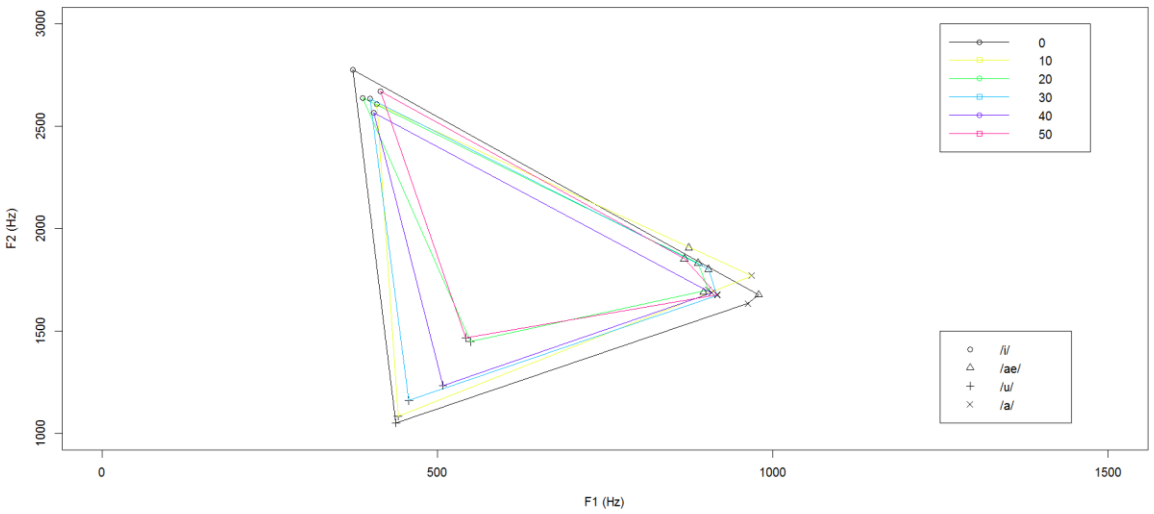


Figure 3: Vowel Space by Parameter for Participant ER.

Figure 3 above suggests that ER demonstrates vowel specific changes, primarily for /u/, that correspond with movement of the tongue down and forward as the perceived vocal tract shortens. This trend is quite strict, although the effect is almost exclusively related to the vowel /u/.

Conclusions:

The acoustic results indicate idiosyncratic responses in both the magnitude and direction of the articulatory adaptation effects for both vowel space area and vowel space by parameter measures. Participant NH followed, or mimicked, the perceived changes in vowel space area by moving the tongue up and backward as the perceived vocal tract shortened. On the other hand, participant ER compensated for the perceived changes in vowel space area by moving the tongue down and forward as the perceived vocal tract shortened.

In terms of changes in formant frequency values, NH’s movement of the tongue up and backward suggests that both F1 and F2 decreased, which indicates a compensatory response in regards to formant values. Conversely, ER’s movement of the tongue down and forward suggests that both F1 and F2 increased, which is contrary to the hypothesis.

Overall, the results indicate that participant NH follows the perturbation in terms of vowel space area, but compensates in terms of formant frequencies. ER demonstrates a completely different behavioral response, because ER compensates in terms of vowel space area but follows in terms of formant frequency values. These results are complex but suggest that it is possible to elicit involuntary changes in articulation when speaking through a virtually shortened vocal tract, indicating potential rehabilitation applications. In conclusion, the purpose is to further develop these methods to determine whether having participants speak through a virtual vocal tract can be an effective means of eliciting involuntary changes in speech articulation for those with motor speech disorders.

Acknowledgements:

This work was supported by Marquette University’s McNair Scholars Program, the Biomedical Sciences Summer Research Program, and the American Speech-Language- Hearing Foundation.

A very special thanks to Jeffrey Berry, Michael Johnson, Rob Gries, Emily Runnoe, Nora Heiderscheidt Evelyn Werking, and Andrew Kolb.

References:

Bastian, A. J. (2008). Understanding sensorimotor adaptation and learning for rehabilitation. Current opinion in neurology, 21(6), 628-633.

Berry, J. J., North, C., Meyers, B., & Johnson, M. T. (2013, June). Speech sensorimotor learning through a virtual vocal tract. In *Proceedings of Meetings on Acoustics* (Vol. 19, No. 1, p. 060099). Acoustical Society of America.

Christoffels, I. K., van de Ven, V., Waldorp, L. J., Formisano, E., & Schiller, N. O. (2011). The sensory consequences of speaking: parametric neural cancellation during speech in auditory cortex. *PloS one*, *6*(5), e18307.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech, Language, and Hearing Research,12*(2), 246-269.

Dysarthria. (n.d.). . Retrieved June 5, 2014, from http://www.asha.org/public/speech/disorders/dysarthria/

Fluet, G. G., & Deutsch, J. E. (2013). Virtual reality for sensorimotor rehabilitation post-stroke: the promise and current state of the field. *Current physical medicine and rehabilitation reports*, *1*(1), 9-20.

Green, J. R., Beukelman, D. R., & Ball, L. J. (2004). Algorithmic Estimation of Pauses in Extended Speech Samples of Dysarthric and Typical Speech.*Journal of medical speech-language pathology*, *12*(4), 149-154.

Hixon, T. J., Weismer, G., & Hoit, J. (2008). Preclinical speech science.

Houde, J.F. and Jordan, M.I. (1998). “Sensorimotor adaptation in speech production,” Science, 279. 1213-1216.

Houde, J. F., & Jordan, M. I. (2002). Sensorimotor adaptation of speech I: Compensation and adaptation. Journal of Speech, Language, and Hearing Research, 45(2), 295.

Kent, R. D., & C. Read (2002). *Acoustic analysis of speech*. Singular Publishing Group.

Kent, R. D., & C. Read. "The acoustic analysis of speech, 1992." *Singular, San Diego*.

Kent, R. D., Kent, J. F., Weismer, G., & Duffy, J. R. (2000). What dysarthrias can tell us about the neural control of speech. *Journal of Phonetics*, *28*(3), 273-302.

Kunath, S. A., & Weinberger, S. H. (2010, June). The wisdom of the crowd's ear: speech accent rating and annotation with Amazon Mechanical Turk. In*Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (pp. 168-171). Association for Computational Linguistics.

Patel, R., Connaghan, K., Franco, D., Edsall, E., Forgit, D., Olsen, L., ... & Russell, S. (2013). “The Caterpillar”: A Novel Reading Passage for Assessment of Motor Speech Disorders. *American Journal of Speech-Language Pathology*,*22*(1), 1-9.

Perkell, J. S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. Journal of Neurolinguistics, 25(5), 382-407.

Tjaden, K., & Wilding, G. (2010). The impact of rate reduction and increased loudness on fundamental frequency characteristics in dysarthria. Folia Phoniatrica et Logopaedica, 63(4), 178-186.

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. Language and cognitive processes, 26(7), 952-981.

Turner, G. S., Tjaden, K., & Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*, *38*(5), 1001-1013.

Wright, W. G. (2014). Using virtual reality to augment perception, enhance sensorimotor adaptation, and change our minds. *Frontiers in systems neuroscience*, *8*.

Xu, M., Homae, F., Hashimoto, R. I., & Hagiwara, H. (2013). Acoustic cues for the recognition of self-voice and other-voice. *Frontiers in psychology*, *4*.

Ziegler, W., & Von Cramon, D. (1983). Vowel distortion in traumatic dysarthria: A formant study. *Phonetica*, *40*(1), 63-78.