Summer 2013

Brittany Bernal - Sensorimotor Adaptation of Vowel Production in Stop Consonant Contexts

Brittany A. Bernal
Marquette University, brittany.bernal87@yahoo.com

Recommended Citation

Bernal, Brittany A., "Brittany Bernal - Sensorimotor Adaptation of Vowel Production in Stop Consonant Contexts" (2013). *Ronald E. McNair Scholars Program 2013*. Book 6.
http://epublications.marquette.edu/mcnair_2013/6

Sensorimotor Adaptation of Vowel Production in Stop Consonant Contexts

Brittany Bernal

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

Marquette University

McNair Scholars Program

Abstract

The purpose of this research is to measure the compensatory and adaptive articulatory response to shifted formants in auditory feedback to compare the resulting amount of sensorimotor learning that takes place in speakers upon saying the words /pep/ and /tet/. These words were chosen in order to analyze the coarticulatory effects of voiceless consonants /p/ and /t/ on sensorimotor adaptation of the vowel /e/. The formant perturbations were done using the *Audapt* software, which takes an input speech sample and plays it back to the speaker in real-time via headphones. Formants are high-energy 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) can uniquely distinguish among the vowels of American English. For this experiment, *Audapt* shifted F1 down and F2 up, and those who adapt were expected to shift in the opposite direction of the perturbation. The formant patterns and vowel boundaries were analyzed using *TF32* and *S+* software, which led to conclusions about the adaptive responses. Manipulating auditory feedback by shifting formant values is hypothesized to elicit sensorimotor adaptation, a form of short-term motor learning. The amount of adaptation is expected to be greater for the word /pep/ rather than /tet/ because there is less competition for articulatory placement of the tongue during production of bilabial consonants. This methodology could be further developed to help those with motor speech disorders remedy their speech errors with much less conscious effort than traditional therapy techniques.

The goal of this study is to gain an understanding of how typically functioning speakers respond to acoustic perturbations to their speech, and to evaluate the amount of vowel sensorimotor adaptation that takes place upon exposure to these experimental conditions in different consonant contexts. Articulatory changes were assessed by measuring formant values, which are high-energy acoustic resonance patterns that reflect the articulatory working space during the production of speech. The effects of coarticulation were analyzed by requiring participants to say the words /pep/ and /tet/. The resulting articulatory changes that took place in participants based on the varying consonant contexts were compared. The methods used in this line of research can be further developed as a potential form of neurorehabilitation for disordered speakers.

Review of the Literature

I: Articulatory Properties of Vowels and Consonants

Vowels are typically produced with voicing by the larynx, due to the vibration of the vocal folds, but different vowels require distinct articulatory positions. Vowels are produced with an open vocal tract. On a spectrogram, a graphical display of acoustics characterizing time, frequency, and intensity, vowels have very well defined resonances or formant patterns (Shriberg & Kent, 2003). The center frequencies of the two lowest frequency formants (F1 and F2, measured in Hertz) uniquely distinguish among the various vowel sounds. Vowels are primarily lingual, meaning the tongue is the main articulator during vowel production (Hixon, Weismer, and Hoit, 2008). Thus, changes in the frequency of F1 and F2 reflect changes in lingual position. Roughly speaking, the frequency of F1 decreases as lingual height increases, and the frequency of F2 increases as lingual position advances forward in the oral cavity (Fant, 1970). The vowel analyzed in this study was /e/, which is produced with the tongue positioned in the middle and front of the vocal tract.

Stops are the types of consonants that were the focus of this study. Stops are classified as obstruents, which are sounds that have either a complete or narrow constriction within the vocal tract (Shriberg & Kent, 2003). Stops begin with occlusion of the airway at the place of articulation. Then, there is a buildup of oral air pressure behind the occlusion in the tracheal space. After the pressure builds up, the airway is abruptly opened and a burst of air is released. Typically accompanying this release is a burst of noise, which is prominent for voiceless stops because of the significant amount of intraoral air pressure. Voiceless stops usually are produced with aspiration that promptly follows the burst noise (Shriberg & Kent, 2003). Aspiration is noise that is generated as air rushes through the open vocal folds. Stop sounds that were analyzed in this study included /p/ and /t/. /p/ is a voiceless, bilabial stop. Bilabial signifies that the two lips are involved in the production of the sound. /t/ is a voiceless, alveolar stop. Alveolar sounds are those in which the tongue comes to contact the upper gum ridge inside the teeth (Hixon, Weismer, and Hoit, 2008).

Taking into consideration the differences in the place of articulation for the consonants /p/ and /t/, there is less competition for articulatory placement of the tongue when saying the word /pep/ compared to /tet/. This is because vowels are lingual productions, whereas /p/ is bilabial, meaning that the lips are the primary articulator. Thus, the tongue can move more fluidly from one place in the vocal tract to the next rather than maneuvering through production of two adjacent sounds using similar places of articulation. Conversely, for the word /tet/, both the vowel /e/ and the consonant /t/ are lingual, because the articulatory production of /t/ places the tongue at the alveolar ridge. There is less competition for placement of the tongue in the word /pep/ than the word /tet/, and the tongue has more freedom to move in /pep/. Based on the differing coarticulatory demands of the two target words in this study, it is reasonable to hypothesize that the amount of sensorimotor adaptation will be greater in the word /pep/.

II: Coarticulation

Coarticulation is the concept that a speech sound is not produced identically in different contexts, but rather it depends upon the preceding or following speech sounds. Although some cases of coarticulation can be observed perceptually (listeners hear differences in the same sound produced in different contexts), there are many cases in which the only evidence of coarticulation is measured quantitatively through programs such as *TF32,* which analyze the acoustic signal (Hardcastle & Hewlett, 1999).

There are two types of coarticulation primarily discussed in speech science. The first is forward coarticulation, which takes place when the articulatory features of an anticipated sound influence the features of the presently produced sound. Backward coarticulation occurs when a presently produced sound is influenced by the articulatory features of a previous sound. (Hixon, Weismer, and Hoit, 2008). For this study, we controlled for both forward and backward coarticulation by using symmetric contexts, meaning that the vowel is surrounded by the same consonant for both of the words /pep/ and /tet/.

Formant transitions are an acoustic effect of changing tongue position, which results in a shift in formant frequency from the beginning to the end of the speech sound. These transitions occur as the articulators move from the constricted position of the stop to the open position for the following sound (Ferrand, 2007). Formant transitions further support the idea of coarticulation and the importance of vowel and consonant contexts affecting formant values.

Acoustically, there are several features that define the various vowel and consonant phonemes. Bonneau and Laprie (2008) characterize acoustic cues that identify phonetic features of French, voiceless stops. A “distinctive region” is defined as a given acoustic space. They characterize selective cues for the place of articulation of voiceless stop consonants. Acoustic detectors for stops include the transient form of the burst as well as the acoustic parameters from the transitions of formants. To determine these parameters, F1 and F2 formant frequencies are evaluated at both the beginning and end of the stop consonant: the vowel onset for CV (consonant-vowel) words and the vowel offset for VC (vowel-consonant) words. These vowel boundaries can also be measured to determine the temporal midpoint of the vowel in Hz/ms, as was done in the current work. The conclusion of Bonneau and Laprie’s (2008) study signifies that the formants transitioning into and out of the vowel are distinct depending on the stop place of articulation (e.g. bilabial or alveolar). This is important because it implies that time-varying formant changes throughout the whole vowel are affected by the surrounding stop consonants. This study relates to our own because it evaluates the acoustic and articulatory features of consonants. We too, measured vowel boundaries to determine the formant values at the midpoint, which essentially define the acoustic characteristics of a phoneme. We then evaluated these formant values and analyzed how they changed over time with the acoustic perturbations. However, measuring the formant values at the midpoint may have been a limiting factor in the current analysis because it did not consider the time-varying characteristic of vowels.

III: Sensorimotor Adaptation

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). Rehabilitative applications of sensorimotor adaptation have been developed for motor control of gait and upper limb movements, suggesting the same principles can be used for rehabilitation purposes in those with disordered speech.

Houde (1998) developed a method for measuring and altering formant patterns in real time. This method generates a synthetic version of speech, which allows the manipulation of formant patterns as they occur so that a speaker hears the manipulated feedback in real time. The formant patterns can be shifted to make speakers think that they are producing the wrong vowel or consonant phonemes. Shifting a speaker's formants in real time typically triggers compensatory changes in the positions of the speaker's tongue, lips, and jaw (Houde and Jordan, 1998). For example, if the speaker is asked to produce the /e/ vowel, but the formants are perturbed so that the speaker hears a sound closer to the /i/ vowel (a perceived auditory-sensory error), the speaker responds by lowering the tongue during /e/ to compensate for the perceived error in vowel height. In other words, the speaker lowers his tongue position because the auditory feedback he is hearing makes it sound like the tongue is too high. Results from Houde (2002) indicate that some participants in these sensorimotor adaptation experiments do indeed adapt their speech to change articulatory positions for vowels and that these changes can persist, evidencing short-term motor learning. Studies such as this one support the conclusion that sensorimotor learning for speech is in fact an interaction between movement patterns and auditory feedback. In order for the phenomenon of sensorimotor adaptation to be used as an effective means of rehabilitation, it is best to gradually present auditory perturbations because this will more likely lead to a broader generalization for learning the new motor behavior

(Bastian, 2008). Houde's research forms a framework for the experimental phases in our own study.

Sensorimotor learning is highly dependent upon different neural functions within the brain. Shadmehr, Smith, & Krakauer (2010) address the effects that adaptation has on motor control as a whole. They define a forward model as the brain predicting the consequences of a motor command. Based on the sensory feedback that the brain receives, this model lets predictions be stored about formulating future actions. Thus, all motor commands rely upon a combination of both sensory feedback and forward control models. If the brain had to depend exclusively on feedback, then there would be a time delay between the planning of motor commands and the actual execution of the commands. There is a benefit of making sensory predictions: the brain does not have to wait for the results of the feedback before acting (Shadmehr, Smith, & Krakauer, 2010). In practical terms this means that motor behaviors, such as the movements of the tongue, lips, and jaw for different speech sounds, can be executed very quickly, but will also be able to adapt in the event of errors. Shadmehr, Smith, and Krakauer (2010) highlight the importance that feedback plays when performing motor behaviors such as speech. Thus, when participants in our study heard their own auditory feedback as altered, their brain perceived that as an error in their speech, and then proceeded to make compensatory articulator movements to correct the perceived error.

Perkell (2012) addresses many variables that affect the motor control of speakers, and discusses adaptation more specifically within the domain of speech. One hypothesis in Perkell's review paper is that variables for speech movements are time-varying patterns of auditory and somatosensory sensation. Thus, speakers rely not only on what they hear, but also on what they feel. Research in his lab uses a neural network model that has a foundation in neurophysiology.

The model is called DIVA (Directions into Velocities of Articulators). This model computes relations among motor goals formulated by the brain and the resulting motor output of speech, as well as auditory and somatosensory consequences of articulatory movements (Perkell 2012). Error maps are used to take note of discrepancies between the expected and the actual sensory consequences of speech motor patterns. These error maps allow for the recognition of speech production errors. The correction of these errors is a primary means for short-term motor relearning (sensorimotor adaptation). This model can characterize relationships between the underlying neural representations of speech sounds and the resulting auditory-acoustic and somatosensory manifestations of speech sounds. Methods reviewed by Perkell (2012) provide a foundation for our own study. These methods can trigger sensorimotor adaptation, causing speakers to unconsciously correct perceived errors in speech while simultaneously increasing accuracy of articulatory targets. Perkell (2012) does not address the long-term clinical goals of our work. Specifically, our intent is to apply sensorimotor adaptation to the neurorehabilitation of motor speech disorders. Currently, much basic, pre-clinical research must be completed before clinical research is viable. In particular, it is critical to understand how different consonant contexts affect sensorimotor adaptation of vowel production. Knowledge of these coarticulatory influences is essential to advancing potential clinical applications of sensorimotor adaptation because speech therapy is highly dependent upon identifying linguistic contexts that support generalization and maintenance of newly learned behaviors. Because there are so many possible combinations of different speech sound sequences, therapists cannot feasibly work on all possible combinations. Thus, we need to identify speech sound contexts that best facilitate sensorimotor adaptation in order to maximize potential learning effects.

Measuring formants was the primary method of analysis in our own study because they allowed us to assess how subjects adapted their articulation. Cai (2010) analyzes native Mandarin speakers' responses to auditory perturbations of the first formant (F1) of their feedback upon saying the triphthong /iau/. A triphthong is vowel combination in which there is a quick transition between three vowel qualities, as opposed to a monophthong, which is a pure "single" vowel sound. The results of the study indicate that on average, participants show partial compensation upon hearing these auditory perturbations. These findings imply that auditory feedback control of speech movements does in fact apply to time-varying gestures, such as the triphthong /iau/. Assessing time-varying gestures was an important consideration in our study because these gestures characterize formant transitions, which depend on the consonant context. In the past, it was believed that auditory feedback was restricted to the quasi-static gestures in monophthongs (e.g., the F1 and F2 frequencies at a single moment during the time course of a speech sound). In addition to this consideration, Cai's conclusions indicate that a weak pattern of generalization was observed in the participants. Generalization is the carryover of an adaptive behavior outside of the trained context. An example of generalization would be changes in the articulation of the vowel /o/ in a study that manipulated auditory feedback for the vowel /e/.

Adaptation experiments require different experimental phases with varying acoustic perturbations. Cai (2010) included start, ramp, stay, and end phases. Start is the baseline phase, in which the speaker hears her own speech without auditory perturbations. Ramp is the training phase, in which the perturbations are gradually presented to the speaker. Stay is the full perturbation phase, in which the F1 and F2 perturbations are at their highest magnitudes. This creates a large discrepancy between what speakers expect to hear themselves say and what they actually hear. Taken together, the ramp and stay phases comprise the primary period of sensorimotor relearning. End is when the perturbations cease, in which subjects are expected to return to their baseline values, or show signs of adaptive behavior. The results show that the changes made are in the opposite direction of the perturbation. Most of the compensatory changes are specific to F1. It was found that the generalization of the adaptation (to a new, untrained context) is less significant than the adaptation to the training context. Cai (2010) concludes that even though the triphthong /iau/ is a time-varying gesture, there is still a significant function of auditory feedback involved. Overall, Cai’s study forms a foundation for the methodology of our own research by the design of the different phases of the perturbations and the *Audapt* software used to manipulate auditory feedback. In the current work, a masking phase was added, during which participants heard only masking noise and no speech. This phase was used to determine if compensatory articulatory changes persist in lieu of auditory feedback. Differences between the start (baseline) phase and the masking phase were used as the primary test of sensorimotor adaptation.

Cai (2012) performed the first sensorimotor adaptation study of disordered speech, which is consistent with the long-term goals of our own research. Cai’s (2012) study analyzes whether or not the speech motor systems of persons who stutter (PWS) use auditory feedback abnormally during speech production. Prior research studies suggest that PWS might have deficits in auditory processing ability, and this may inhibit them from perceiving errors as effectively as persons with fluent speech (PFS). Another aim of Cai’s study is to distinguish what deficits PWS have in their integration of auditory and motor control systems. An auditory perturbation to F1 is employed during the production of the monophthong /ɛ/. Results indicate that PWS do in fact use compensations that are similar to the control group, but their compensations are significantly smaller than the control participants. There are three different hypotheses in this study. The first hypothesis is that PWS have an auditory perception deficit, which in turn negatively affects their compensatory abilities for auditory perturbations. The second hypothesis is that PWS have an abnormal gain in their auditory feedback control systems for speech, and this is based on the evidence of smaller or larger than average responses to auditory feedback perturbations. The third and final hypothesis is that PWS have an abnormal variability in their motor responses to auditory feedback errors, which is evidenced by their greater variability than PFS in their responses to perturbations. For the experimental procedure, *Audapt* is used to track and shift the formant frequencies, which are played back in real time to the participants. Participants are required to say the words “head” and “pet”, each word containing the monophthong /ɛ/. The software requires both consistency and intensity in speaking rate. There are also trials in which the participants receive no auditory feedback (Cai, Beal, Ghosh, Tiede, Guenther & Perkell, 2012). This experimental design is similar to our own study because we used *MATLAB* alongside *Audapt* to shift F1 down and F2 up, and saw compensatory changes in the participants. Also, we attempted to keep participants consistent in their volume by using a sound pressure level meter, which let participants monitor their own volume and keep it within a certain range.

Overall, our research study aimed to assess how individuals acoustically perceive their own speech, and focused on using involuntary changes in perturbed acoustics as a means of eliciting sensorimotor learning that caused changes in vowel articulation. We hope to gain a better understanding of how typically functioning speakers react to perturbed auditory feedback to determine methods that can further be developed and applied to those with disordered speech. Previous literature indicates that there is an interaction between articulatory movement patterns and auditory feedback. The current study aims to understand how coarticulatory influences (consonant context) can support maintenance of newly learned changes in vowel articulation. This adaptation phenomenon is not well understood in terms of using it as a long-term form of rehabilitation, especially in those speakers with disordered speech. Thus, there is much more to be learned in terms of sensorimotor adaptation of speech before it can benefit those with motor speech disorders.

Methods

Overview of Methods

This study consisted of five experimental phases in which participants said words with different consonant and vowel contexts. Participants were wearing a headset with a microphone and insert earphones. As they spoke, participants heard perturbations to their auditory feedback. These perturbations were designed to elicit involuntary changes in the articulator movements of the speaker. A sound pressure level meter was used to help participants monitor their speech volume and keep it constant. The vowel being assessed was /e/ rather than /ɛ/, which is less commonly used in studies of adaptation. The vowel /e/ was chosen because it is a mid vowel; there was room for subjects to compensate and adapt by shifting their formants to sound more like a high vowel, such as /i/, or shifting their formants to sound more like a low vowel, such as /ɑ/. The articulatory contexts of words in this study included voiceless stop consonants surrounding the vowel /e/. The focus of this study was on the words /pep/ and /tet/. Voiceless consonants were chosen in order to more clearly distinguish the boundaries of the target vowel. This project tested the effects of sensorimotor adaptation in typically functioning individuals who were American English speakers with no history of speech, language, or hearing pathology. Subjects were within the age range of 20-22 years. There was 1 male and 3 female subjects. Speech of the participants was recorded with a microphone and processed into a computer program called *Audapt* that measured the formant values in hertz. A speech synthesis tool in this program shifted their F1 values down and F2 values up. The software played the manipulated speech back to the participants in real-time. This created a synthetic version of the participants' speech that they perceived as their own production.

Specific Methods

The focus of the study was on how the bilabial stop context in /pep/ and the alveolar stop context in /tet/ differentially affected sensorimotor adaptation of the vowel /e/. Using the *Audapt* software, these words were perturbed to sound like /pip/ and /tit/. The acoustic shift caused participants to perceive an error in articulation as an increased vowel height. To check if generalization occurred for other vowels, the speakers also said the words /pop/ and /tot/ but did not hear acoustic perturbations of these words. As a neutral context for all trials, experimental runs with the word /he/ were also completed. This word was perturbed to sound like /hi/, and generalization was assessed with the word /ho/. The reason for choosing this word is because /h/ is considered to have little or no coarticulatory influences on the production of adjacent vowels. Thus, this "neutral" context provided a reference for evaluating the coarticulatory effects of the two primary target words.

For this experiment, F1 was shifted down, and F2 was shifted up by the *Audapt* software. The magnitude of the shift varied person by person. The average parameters that F1 was shifted ranged from 175-200 Hz, and the average parameters for F2 was 200-400 Hz. Due to the effects of adaptation, speakers in this study were expected to compensate for the perceived errors in their speech by shifting their F1 value up and their F2 value down, which would be compensating for the perceived error in the opposite direction of the shift.

The various phases of the experiment are shown in Table 1. First, a baseline was established, in which participants heard their speech as they said the words /pep/ and /pop/ or /tet/ and /tot/ and it was played back to them without the formant values being altered. This was used to create a control variable to establish normal data which could then be contrasted with the speech produced with perturbed auditory feedback.

Following baseline was the ramp phase, in which formants were gradually shifted to be more and more variant from participants' actual speech production. The participants said /pep/ but they heard something progressively closer to /pip/, or they said /tet/ but they heard something progressively closer to /tit/. The ramp phase gradually incorporated the changes in auditory feedback, which eventually elicited the compensatory articulatory response.

The next phase was full perturbation, in which the greatest amount of auditory perturbations took place because formants were completely shifted to their highest or lowest magnitudes. In this phase, it was anticipated that the greatest compensation would occur due to the large discrepancy between the expected and the actual speech production of the participant. Taken together, the ramp and full perturbation phases comprised the primary periods of sensorimotor relearning.

Following full perturbation was a masking phase, in which loud noise was played back into the earphones so the participants could not hear their own speech. This phase helped establish whether or not sensorimotor learning had actually occurred because participants were unable to hear their own auditory feedback. This phase provided a test of adaptation because subjects were no longer hearing any perturbations, but if adaptation had occurred, they would continue to demonstrate changed articulation.

The final phase was the return phase, in which participants heard their own speech without the formant perturbations. This helped establish whether there was fading or weakening of the adaptive response. This also provided a check to see if the last few productions were back to where they were during baseline, or if subjects continued to maintain adapted articulations.

Phase	**Learning Behavior**	**Auditory Feedback Condition**
Baseline	Baseline	Unperturbed
Ramp	Training	Gradual F1 Shift Down and F2 Shift Up
Full Pert	Compensation	Constant Maximum F1 Shift Down and F2 Shift Up
Masking	Adaptation	Noise: No Auditory Feedback
Return	Baseline	Unperturbed

Table 1: Adaptation experiment phases and conditions.

To analyze the patterns of articulatory change throughout the different experimental phases, *TF32* software was used to mark the boundaries of the vowels on a spectrogram, a process known as indexing. The indices were then labeled, differentiating between the spoken (input) files and the resynthesized, or perturbed files. After indexing was complete, the data were processed through a program called *S+*, which used the index labels to calculate the temporal midpoint of the vowel and find the corresponding F1 and F2 values in Hertz at the midpoint. These F1 and F2 values were then used for analysis to determine if compensation and adaptation occurred.

Results

F1 Shift	-72
F2 Shift	246
F1 Compensation	44
F2 Compensation	-157
F1 Adaptation	21
F2 Adaptation	-131
F1 Generalization	0
F2 Generalization	-20

Table 2: Formant value changes determined for subject 4 for the word /tet/.

Acoustic data obtained from an example (subject 4) are shown in Table 2 and Figure 1. In Table 2, rows labeled F1 Shift and F2 Shift indicate the magnitude that *Audapt* perturbed the subject's formant values to modify his auditory feedback. Subject 4 compensated by changing his F1 value up by 44 Hz, which was expected because it is in the opposite direction of the perturbation. The subject compensated by changing his F2 value down by 157 Hz, which was also expected because he shifted in the opposite direction of the perturbation. The F1 and F2 values for adaptation are also consistent with the hypothesized direction of the shift.

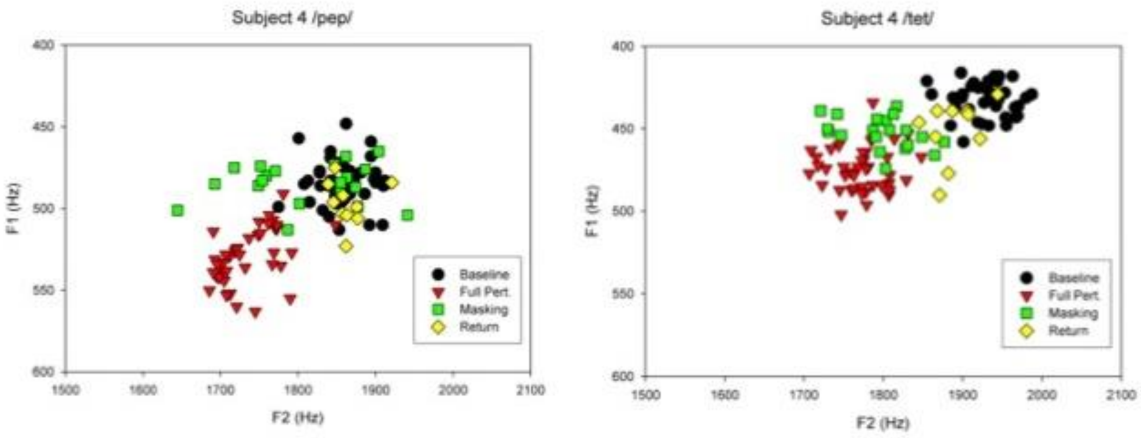

Figure 1: Acoustic data for production of /pep/ and /tet/ of subject 4.

Data from different phases of the experiment are shown by differently colored and shaped symbols. F2, which is on the x-axis, represents tongue forwardness, whereas F1, which is on the y-axis, represents tongue height. These two axes define an approximation of the articulatory space. Subject 4 serves as an example for the hypothesized compensatory and adaptive responses in this experiment. For the word /pep/, the four clouds of data points appear to differ between phases, signifying that subject 4 used different articulatory spaces for his productions throughout the experimental phases. He increased his F1 values and decreased his F2 values from baseline to full perturbation, indicating compensation. Subject 4 also increased his F1 values and lowered his F2 values from baseline to masking, indicating adaptation. The acoustic differences between experimental phases suggest that subject 4 made large changes in articulation from both baseline to full perturbation and baseline to masking, consistent with sensorimotor adaptation of the vowel in /pep/. Subject 4 also displayed sensorimotor adaptation for /tet/ based on F1 and F2 changes between phases, but with smaller overall magnitude compared to /pep/ indicated by the more condensed data points.

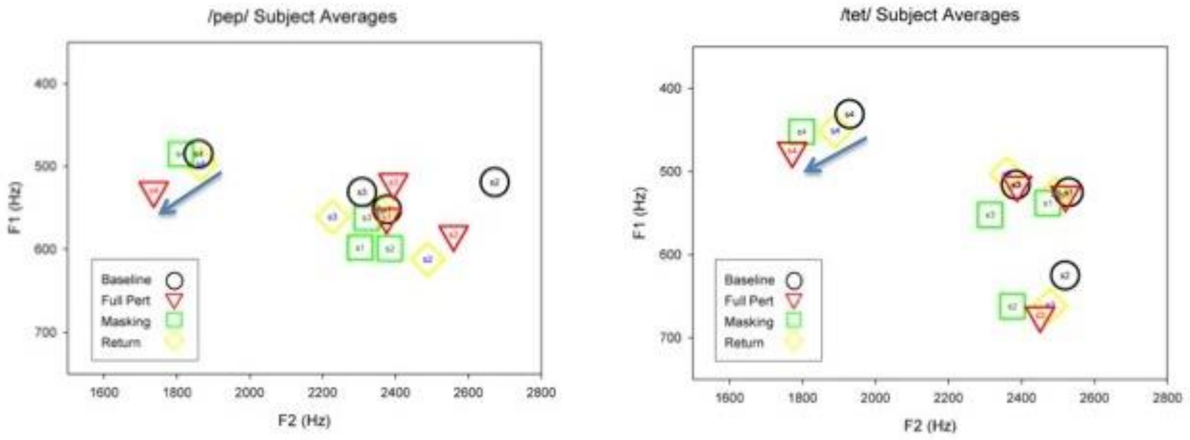

Figure 2: Subject means for the words /pep/ and /tet/.

Average data by subject and experimental phase are shown in Figure 2. The vectors for subject 4 indicated by the blue arrows in these figures follow the expected direction and magnitude for compensation and adaptation. Compensation and adaptation are evidenced by an increase in F1 from baseline and a decrease in F2 from baseline.

For both the words /pep/ and /tet/, subject 2 also compensated and adapted in the predicted direction, and the magnitude of the overall articulatory changes (measured acoustically) was larger in /pep/ compared to /tet/. Subjects 1 and 3 displayed more variable results. Subject 1 evidenced weak compensation, but her articulatory changes for adaptation were larger than those for compensation. Thus, subject 1 was making more changes in tongue position when her auditory feedback was eliminated than when she initially heard the perturbed auditory feedback, signifying a carryover of the learned changes in vowel articulation. Subject 3 followed the direction of the shift for her F1 compensation value, signifying that instead of moving her tongue down, which would be the hypothesized response, this subject moved her tongue up despite perceiving an increased vowel height. Subject 3's acoustic results were also unpredictable in the sense that her compensation for F2 was much greater in the word pep (signifying a backwards movement of the tongue, which was expected), but her F2 adaptation was much greater for the word /tet/ rather than /pep/, which was contrary to the hypothesis because /tet/ was anticipated to be the less facilitative consonant context for adaptation.

Discussion

For the word /pep/, there appeared to be a greater degree of compensation and adaptation that occurred across subjects. For both compensation and adaptation of the words /pep/ and /tet/, 3/4 subjects displayed articulatory changes that were consistent with the hypothesis. 1/4 of the subjects (subject 3) shifted in the same direction of the perturbation, which was contrary to the hypothesis. A possible reason for unexpected responses such as these could be due to errors in measurement or simply inconsistent subject behavior. Also, it is established that some talkers are less sensitive to auditory feedback manipulations and do not compensate or adapt. The magnitude of the shift was more significant in /pep/ across subjects. This was evident because the data points were more spread out between experimental phases for the word /pep/, signifying a greater range of articulatory space that was used throughout the experiment. Speaker-specific differences were noted. For example, subject 4 appeared to display a greater degree of adaptation for F2 than for F1. Because F2 represents tongue forwardness, and subject 4 displayed a significant decrease in his F2 value, this suggests that in the bilabial context, there is a large degree of tongue movement posteriorly. This backwards movement of the tongue is much more evident than the vertical movements according to the less significant impact on F1, which represents tongue height. Thus, consonant context does have an effect on sensorimotor adaptation of vowels, because there was a greater degree of both compensation and adaptation in the bilabial rather than the alveolar context, as predicted by the hypothesis.

The long-term goal of this line of research is to determine whether or not sensorimotor adaptation of speech has a potential to be used as a form of neurorehabilitation for those with motor speech disorders. Currently, traditional digital signal processing techniques for manipulating auditory feedback are limited in how well they can be applied to disordered speakers because clear vocal quality is required. Unfortunately, those with disordered speech are unable to produce the robust vocal quality necessary for effective formant manipulations. Thus, further technical improvements will be required before sensorimotor adaptation can be studied as a potentially novel method of speech rehabilitation. The preliminary outcomes of the current research study demonstrate that consonant context influences sensorimotor adaptation of vowels.

Continuing analysis of the current data will assess the statistical significance of these effects as well as examine more complete formant history measures characterizing the articulatory changes across the total vowel durations. Our current findings indicate that there is in fact a greater degree of articulatory adaptation of vowel production at the temporal midpoint of the vowel in the bilabial compared to the alveolar context, and this is primarily characterized by backwards movement of the tongue. Using this information creates a basis of knowledge about sensorimotor adaptation of speech that will let us extend that knowledge to clinical populations in the future.

Future efforts related to this line of research will focus on further exploration of the potential learning applications of using sensorimotor adaptation. One of the main challenges for this line of work is to better evaluate the maintenance of changes in articulatory behavior elicited using auditory feedback perturbations. While short-term, compensation and adaptation often occur, but the persistence of these learning effects remains uncertain.

Acknowledgements

This work was supported by the American Speech-Language-Hearing Foundation New Century Scholar Research Grant and grant IIS-1141826 from the National Science Foundation.

A very special thanks to Jeffrey Berry, Michael Johnson, John Jaeger, Rachael Hefel, Julia Wesley, Melissa Strobel, Scott Palahniuk, Rachel Myatt, Sarah Zach, Kayla Shaw, and Trish Lyons.

References

Bastian, A. J. (2008). “Understanding sensorimotor adaptation and learning for rehabilitation,” *Current Opinion in Neurology*. 21 (6),628-633.

Bonneau, A., & Laprie, Y. (2008). Selective acoustic cues for French voiceless stop consonants. *The Journal of the Acoustical Society of America, 123,* 4482.

Cai, S., Beal, D. S., Ghosh, S. S., Tiede, M. K., Guenther, F. H., & Perkell, J. S. (2012). Weak responses to auditory feedback perturbation during articulation in persons who stutter: Evidence for abnormal auditory-motor transformation. *PLos ONE*, *7*(7), 1-13.

Cai, S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2010). Adaptive auditory feedback control of the production of formant trajectories in the Mandarin triphthong/iau/and its pattern of generalization. *The Journal of the Acoustical Society of America*, *128*, 2033.

Fant, G. (1970). *Acoustic theory of speech production*. The Hague: Mouton.

Ferrand, C. T. (2007). *Speech science: An integrated approach to theory and clinical practice*. (2nd ed.). Boston: Pearson Education, Inc.

Hardcastle, W. J., & Hewlett, N. (1999). *Coarticulation*. Cambridge: Cambridge University Press.

Hixon, T. J., Weismer, G., & Hoit, J. D. (2008). Pharyngeal-oral function and speech production. In T. J. Hixon (Ed.), *Preclinical Speech Science: Anatomy Physiology Acoustics Perception* (2 ed.).

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.

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

Shadmehr, R., Smith, M., & Krakauer, J. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience*, 90-108.

Shriberg, L. D., & Kent, R. D. (2003). *Clinical phonetics*. (3rd ed.). Boston: Allyn and Bacon.