2015-12-01

Changes in Acoustic and Kinematic Articulatory Working Space Across Three Intensity Levels

Panika Ellis Palmer
Brigham Young University - Provo

Changes in Acoustic and Kinematic Articulatory Working Space

Across Three Intensity Levels

Panika Ellis Palmer

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Christopher Dromey, Chair
Ron Channell
Kristine Tanner

Department of Communication Disorders

Brigham Young University

December 2015

ABSTRACT

Changes in Acoustic and Kinematic Articulatory Working Space
Across Three Intensity Levels

Panika Ellis Palmer
Department of Communication Disorders, BYU
Master of Science

The purpose of this study was to compare changes in acoustic and kinematic measures of articulation across soft, comfortable, and loud speech conditions. There were 19 participants, 9 male and 10 female, with age ranging from 20 to 34 with a median age of 25. Each participant had electromagnetic sensors glued to their tongue, jaw, and lips. It was anticipated that the acoustic measures would accurately reflect the kinematic measures of speech as articulation changed across the intensity levels. Vowel space area (VSA) and vowel articulation index (VAI) were computed from the three corner vowels, /ɑ, i, u/. Articulatory-acoustic vowel space (AAVS), a sentence-level acoustic measure, was computed from the continuous formant histories for all voiced segments in a sentence. Kinematic-vowel space area (KVSA), kinematic-vowel articulation index (KVAI), and articulatory-kinematic vowel space (AKVS) were the kinematic equivalents of the acoustic measures, and were newly developed for the present study. Stroke metrics based on the speed history of the lingual movements were also used to reveal average kinematic features of the articulatory gestures in each participant's speech. The data revealed that the isolated acoustic and kinematic measures that used corner vowels (VSA, VAI. KVSA, KVAI) did not change significantly with intensity. The sentence-level continuous measures of articulatory working space (AAVS and AKVS) increased as speech intensity increased. The other sentence-level kinematic metrics also changed significantly with speech intensity, including increases in hull volume, onset speed, peak speed, mean speed, and distance. Stroke duration decreased as speech intensity increased. These findings suggest that measures based on isolated corner vowels are not as reflective as continuous measures of changes in articulatory movement in speech.

Keywords: vowel space area, vowel articulation index, articulatory acoustic vowel space, speech kinematics, acoustic speech measures

ACKNOWLEDGMENTS

I would like to thank Dr. Christopher Dromey for all of the feedback and encouragement he has given me while writing my thesis. I would also like to thank Dr. Ron Channell and Dr. Kristine Tanner for their willingness to be on my thesis committee and provide me with valuable feedback. Lastly, I would like to thank my friends and family for all of their love, encouragement, and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

DESCRIPTION OF THESIS STRUCTURE

This thesis is part of a larger study involving four thesis projects that may result in publications for which I may be listed as a co-author. This thesis was written as a manuscript that is suitable for submission to a peer-reviewed journal in speech language pathology. This thesis is in the format of a journal article with an annotated bibliography in the appendix.

Introduction

The primary purpose of this study was to compare changes in acoustic and kinematic measures of articulation across soft, comfortable, and loud speech conditions. A further purpose was to better understand the parallels between acoustic and kinematic indexes of speech, while also comparing metrics based on isolated vowel samples with measures from complete sentences.

Speech Acoustics

Perceptual evaluation is useful in a clinical setting because speech-language pathologists (SLPs) are able to describe the characteristics of their client's speech and then choose therapy techniques that address the identified deficits. SLPs frequently assess progress based on what they hear. However, perceptual evaluation has its limitations because it is a subjective process that lacks the precision and reliability that is needed when doing speech analysis in research. Acoustic measures can help overcome these limitations by providing objective measures of speech performance.

Two major advantages of speech acoustic measures are that they are inexpensive and noninvasive. Acoustic analysis software such as Praat can quantify several acoustic characteristics, including fundamental frequency, intensity, and formants (Boersma & Weenink, 2015). These objective measures can be useful in research. The main instrumentation that is required for speech acoustic analysis consists of a recording device and software. Once the participant's speech sample has been recorded, it can be imported into the software and analyzed for the specific speech features that the researchers are studying (Tjaden & Wilding, 2004). Speech acoustics are useful in research because they can indirectly reflect the activity of the vocal tract.

The configuration of the vocal tract is what determines the sounds that are produced during speech. For example, when a person rounds their lips, they lengthen their vocal tract which, along with the high back position of the tongue, results in the vowel /u/; this is in contrast to the result of lip retraction and tongue advancement, which we hear as an /i/ vowel. Changes in the shape of the vocal tract will determine which frequencies are most prominently resonated. Formants are resonant peaks, and they are determined by the placement of the tongue, jaw, and lips during speech.

Formants. Formant frequencies are used to distinguish vowels (Sandoval, Berisha, Utianski, Liss, & Spanias, 2013). A basic explanation of the first two vowel formants is that the first formant (F1) is strongly influenced by the height of the tongue and the second formant (F2) increases with tongue advancement. In addition, other factors can influence formant frequencies, such as larynx height or lip rounding (Dromey, Jang, & Hollis, 2013). Changes in F1 and F2 can be fairly predictable based on the movement of the articulators and the three-dimensional formation of the vocal tract (Sapir, Ramig, Spielman, & Fox, 2010). Sapir et al. (2010) explained that the frequencies for F1 and F2 often have an inverse relationship. For example, the frequency of F1 decreases when the tongue moves forward while F2 increases. Conversely, F2 decrease when the tongue moves backward. The F1 and F2 are both lowered when lips are rounded and increase when the lips are retracted. Formants have been used in speech acoustic research for many years.

Traditional formant-based metrics. Vowel Space Area (VSA) is an acoustic measure that is used to reflect vowel separation. VSA is considered a measure of acoustic working space, because the frequencies of F1 and F2 are related to the size and shape of the cavities created by the jaw opening and tongue position (Sandoval et al., 2013). Researchers use the four corner vowels, /i/, /æ/, /u/ and /ɑ/ to determine the VSA of the vowel quadrilateral (Berisha, Sandoval, Utianski, Liss, & Andreas, 2013). Three corner vowels, such as /i/, /u/ and /ɑ/, can also be used to determine the triangular VSA. F1 and F2 vowel midpoint frequencies are plotted on an X-Y graph, where the corner vowels mark the corners of the quadrilateral or triangle. The area of the VSA can be computed in Hz^2.

VSA can be used to reflect the differences between disordered and typical speech. For example, speakers with hypokinetic dysarthria often have a reduced VSA compared to typical speakers because they under-articulate their vowels. When vowels are under-articulated in dysarthric speech, they become more centralized, which causes the production of a vowel to be similar to a schwa. This process results in vowels that are more neutralized in their production, which is associated with poorer intelligibility (Sapir et al., 2010).

VSA is a valuable acoustic measure but it also has its limitations. Some studies that used VSA have found statistical differences between dysarthric and normal speech, while other studies have not (Sapir et al., 2010). Sapir et al. (2010) also found that VSA is highly sensitive to interspeaker variability, which might mask differences between dysarthric and normal speech. The effects of coarticulation are also a limitation of VSA. Recent studies have used vowels in the context of sentences to calculate VSA (Sandoval et al., 2013; Skodda, Visser, & Schlegel, 2011). As a result of using sentence stimuli, coarticulation can greatly affect vowel formants by altering how the vowel is formed based on the phonemes that are being produced before or after it. Other aspects of speech can also greatly affect vowel formants such as speech rate, dialect, and speaking style (Fridland, Kendall, & Farrington, 2014; Tjaden & Wilding, 2004).

The Formant Centralization Ratio (FCR) was developed to quantify the distinctiveness of vowels with greater sensitivity than the VSA. FCR is able to detect more of the differences between disordered and typical speech because it is calculated differently than VSA (Sapir et al., 2010). The equation for FCR is expressed as (F2u + F2a + F1i + F1u)/ (F2i + F1a). F2u represents the second formant for the vowel /u/, F2a represents the second formant for the vowel /ɑ/, F1i represents the first formant for the vowel /i/ and so forth. A result of the FCR equation is that the formant frequencies in the numerator are going to increase while the formant frequencies in the denominator are going to decrease when vowel centralization occurs in disordered speech. Consequently, the FCR is directly focused on vowel centralization. This results in researchers being able to differentiate even mildly dysarthric speech from normal speech. Although FCR is more sensitive to vowel formant differences than VSA, a more recently developed measure, the vowel articulation index (VAI), is starting to be used to quantify the acoustic distinctiveness of vowels.

VAI is a simple mathematical reformulation of the FCR (Roy, Nissen, Dromey, & Sapir, 2009). The VAI for the four corner vowels is determined by the following equation (F2i+F1a)/ (F2u + F2a+F1i+ F1u). The VAI decreased with vowel formant centralization and increased with decentralization. This means that as the speaker's vowel space becomes more centralized, their VAI decreases and as vowel space becomes more expanded the VAI increases. Thus, the VAI is more intuitive to interpret than the FCR because it decreases when vowel centralization occurs and increases when the vowel space expands.

A novel formant-based acoustic measure is the articulatory-acoustic vowel space (AAVS). It differs from the VSA, FCR and VAI measures in that AAVS is based in the formant history for all vowels and diphthongs in a sentence, rather than just a few static points during the production of corner vowels (Whitfield & Goberman, 2014). All three previously discussed acoustic measures are calculated from the midpoint of the three or four corner vowels, providing a snapshot of one steady-state moment in a speech signal. AAVS tracks the trajectories of formants during connected speech. It results in a dynamic picture of how the articulators are interacting throughout an entire utterance. This way of measuring formants may allow researchers to more readily distinguish disordered speech from typical speech. Whitfield &

Goberman used AAVS to determine differences in articulatory-acoustic patterns between people with Parkinson's disease (PD) and their neurologically typical peers while they spoke in their habitual style or while employing clear speech strategies. The authors explained that they used AAVS to track clarity-related changes in articulatory function of individuals with PD. The use of AAVS allows for more sensitivity to within-speaker articulation changes as well as overall articulatory behavior.

Acoustic measures have proven valuable in research because they conveniently provide objective measures of performance. With that being said, they also have their limitations. For example, some articulatory movements result in small acoustic changes while other articulatory movements can result in large acoustic differences (Stevens & Keyser, 2010). This nonlinearity can make it difficult to interpret acoustic measures in a straightforward way.

Speech Kinematics

Speech kinematic techniques provide direct measurements of the movements of vocal tract structures. Combining kinematic and acoustic measurements provides a more complete picture of how the articulators are moving during speech. Kinematic measures can augment the information derived from acoustic analysis by revealing aspects of speech that are not perceptually noticeable (Barlow, Cole, & Abbs, 1983). Kinematic parameters can reflect speech movement patterns, which can lead to greater understanding of the differences between typical and disordered speech.

The head-mounted movement transducer system is one way to measure the kinematics of speech (Barlow et al., 1983). It is made up of a head frame with strain gauge movement transducers, which are used to measure articulator movement. The transducers are connected to the upper lip, lower lip, and jaw. Each one has an "individual strain gauge cantilever module" (1983, p. 285). These cantilevers measure the vertical displacement of the upper lip, lower lip, and jaw. One drawback of the head-mounted movement transducer system is that it measures the movement of only three, externally accessible articulators, and not the tongue. It also is not practical to use on neurologically impaired individuals who may display ballistic movements of the head and neck.

Another way to measure speech kinematics is the OptoTrak (Northern Digital, Waterloo, Ontario, Canada) camera system (Kleinow & Smith, 2006). In this system small infrared light emitting diodes (IREDs) are attached to the upper lip, lower lip, and jaw. These IREDs are then tracked via the camera system while the individual speaks. This allows the articulation of the speaker to be measured. The OptoTrak system also includes one IRED being placed on the forehead and three on a pair of modified goggles, which are used as a point of reference for the head. Like the strain gauge system, OptoTrak also lacks the capacity to track the movements of the tongue. Other kinematic transduction systems are able to track the movement of all of the articulators.

The electromagnetic articulograph (EMA) system is able to measure the movement of the articulators in three dimensions (Berry, 2011). The NDI Wave system (NDI, Waterloo, Ontario, Canada) is a type of EMA system that allows tracking of x, y, and z spatial coordinates, and the angular coordinates characterizing rotation about the transverse axis (pitch) and anterior–posterior axis (roll). This system uses a transmitter that creates a magnetic field, allowing each sensor to be tracked. Two advantages of this system are that it displays real-time movement of the articulators and automated head-movement compensation. These features allow for an accurate representation of the kinematics of the articulators by taking into account the movement of the tongue as well as the movement of the externally accessible articulators (the lips and jaw).

Stroke metrics are average measures of articulatory gestures derived from a kinematic recording of continuous speech. A stroke is a movement that occurs during the period between two successive local minima in the speed history as the articulators transition from one position to another position during speech (Tasko & Westbury, 2002). Stroke metrics include a variety of kinematic measures such as stroke count, onset speed, mean speed, peak speed, stroke distance, and stroke duration. Stroke count is a reflection of how many times the articulators changed position for individual articulatory gestures throughout the utterance being analyzed. Onset speed is the speed of the articulators when they first start to change position and begin a new movement stroke. Peak speed is the fastest speed that the articulators reach during a movement stroke. Stroke distance is how far the articulators traveled when transitioning from one position to another. Stroke duration is the amount of time between each change in position of the articulators. Stroke metrics are useful because they can summarize the movement of the articulators as they transition from one sound to the next throughout an entire utterance. In addition to stroke metrics, sentence duration and hull volume provide further information about articulation. Hull volume is the overall area in mm^2 in the x/y plane that encompasses the space in which the articulators moved during an utterance.

Purpose of the Present Study

This study investigated the articulatory acoustic and kinematic differences between soft, comfortable, and loud speech (Schulman, 1989). Previous research has shown that louder speech results in larger movement of the articulators (Dromey & Ramig, 1998). By collecting data for speech that includes a range of articulatory amplitudes, we were able to learn how the acoustic metrics (VSA, VAI, and AAVS), along with their kinematic counterparts, reflected vocal tract movements. By doing this, we were able to better understand the parallels between acoustic and kinematic indexes of speech, while also comparing metrics based on isolated vowel samples with measures from complete sentences.

Method

Participants

The participants were native speakers of American English without noticeable regional dialects, as determined by the examiners. They had no history of speech, language, or hearing disorders. The group of participants included 9 males and 10 females ranging in age from 20 to 34 with a median age of 25.

Stimuli

This experiment was part of a larger research study that addressed several different questions using recordings from the same speakers. The stimuli for this part of the study included a set of six sentences with the three corner vowels /i/, /u/ and /ɑ/ in different phonemic contexts. Each participant said the sentence three times in a row after a 20-minute adaptation period that allowed the speakers to adjust to the presence of the movement sensors on the articulators.

Table 1

Sentence Stimuli

1. It's time to shop for two new suits.
2. A good AC should keep your car cool.
3. It's never too soon to choose the right.
4. One warm morning a boy was mowing the lawn.
5. We do agree the loud noise is annoying.
6. There's no good reason they would go down there.

Instrumentation

Each sentence (Table 1) was spoken into an AKG C2000B microphone, which was positioned approximately 30 cm from the speaker's mouth. The microphone signal was sampled at 22,050 Hz using a Focusrite Scarlett 2i2 preamplifier and a Dell Optiplex 990 computer. The articulator movements were tracked via an NDI Wave system with a sampling rate of 400 Hz. The acoustic and kinematic signals were recorded with the NDI WaveFront software.

Procedure

Each participant was positioned on a chair in an Acoustic Systems sound-attenuating booth approximately 90 cm from the sentence stimuli, which were printed in black, 36 point font on white paper. The researchers used latex gloves, tongue depressors and PeriAcryl®90 cyanoacrylate tissue adhesive to the glue 5 magnets to the tongue, lips and lower incisors; the incisor attachment included Stomahesive material to avoid enamel damage from the glue. Silver DriAid pads were used to dry the tongue in preparation for the placement of the tissue adhesive. The first magnet (T1) was glued onto the superior surface of the tongue 3 cm posterior to the tip along the midline. The second magnet (T2) was glued onto the superior surface of the tongue 1 cm posterior to the tip along the midline. The third magnet (J) was attached to the lower incisor along the midline to reflect jaw movement. The fourth sensor (LL) was glued on the vermillion border of the lower lip at midline. The fifth magnet (UL) was glued onto the vermillion border of the upper lip at midline.

Data Analysis

Acoustic Measures. A custom Matlab (MathWorks, 2014) software application was used to segment out time-aligned audio and kinematic signals from the corner vowels /i/, /ɑ/ and /u/ from the sentence stimuli for soft, comfortable and loud speech. This was done for the three repetitions for each participant. Once the vowels were segmented, they were imported into Praat (Boersma & Weenink, 2015), where the formant tracks were computed. The Praat formant text file was imported into Matlab, where the F1 and F2 histories along with the time-aligned kinematic records were processed further. The software computed average F1 and F2 values for the middle 50% of each corner vowel repetition for each speaker and also the X/Y position for both tongue sensors at the same time points.

Vowel space area. The VSA for each vowel was computed based on the F1 and F2 formants for each of the three repetitions of the vowels /i/, /ɑ/, and /u/ from each participant at each intensity level. Matlab computed the area in Hz^2 based on the F1 and F2 'coordinates' for the three corner vowels.

Vowel articulation index. The F1 and F2 values from each vowel were also used to compute the VAI with the equation (F1a+F2i)/ (F2a+F1i+ F1u+F2u).

Articulatory-acoustic vowel space. Dr. Jason Whitfield from Bowling Green State University used his proprietary Matlab software to calculate AAVS from the continuous F1 and F2 for records from the selected sentence stimuli in the current dataset.

Kinematic Measures.

Kinematic-vowel space area. The kinematic equivalent of VSA, called kinematic-vowel space area (KVSA) was computed in Matlab. The articulatory working space for both the T1 (mid-tongue) and T2 (front of the tongue) sensors was based on the X/Y position records for the tongue markers at the same time points used for the VSA format-based metric. As a kinematic parallel of the acoustic VSA, the area of the triangle formed by the corner vowels was computed in mm^2.

Kinematic-vowel articulatory index. The kinematic- vowel articulatory index (KVAI) was computed from the T1 and T2 sensor coordinates, expressed in mm on the X and Y axes. The equation (iy+ux+uy+ax)/(ix+ay) was used to calculate the KVAI, which corresponds with the VAI acoustic measure. The i, u, and a indexes refer to the corner vowels and the x and y refer to the anteroposterior and vertical positions, respectively.

Articulatory-kinematic vowel space. A kinematic equivalent of AAVS, called the articulatory kinematic vowel space (AKVS), was calculated for each sentence by Dr. Whitfield, using the same algorithm as the AAVS, but with the X/Y tongue position records as input.

Stroke metrics. The speed record was computed as the square root of the squared position changes in the X and Y planes, such that it reflected the movement speed of each articulatory sensor without regard to direction; this record was based on the position of the articulators throughout the utterance (Tasko & Westbury, 2002). The speed history included alternating local minima and maxima as the articulatory movement varied in speed. Figure 1 provides an example of a speed plot with speed minima for T2 using the *Suits* sentence. The following stroke metrics were computed using an algorithm in Matlab: stroke count, onset speed (mm/s), mean speed (mm/s), peak speed (mm/sec), stroke distance (mm), and stroke duration (m/s). Additionally, sentence duration and hull volume (mm^2) were computed. Figure 2 shows the x-y movement of both tongue makers throughout the *Suits* sentence with the hull volume.

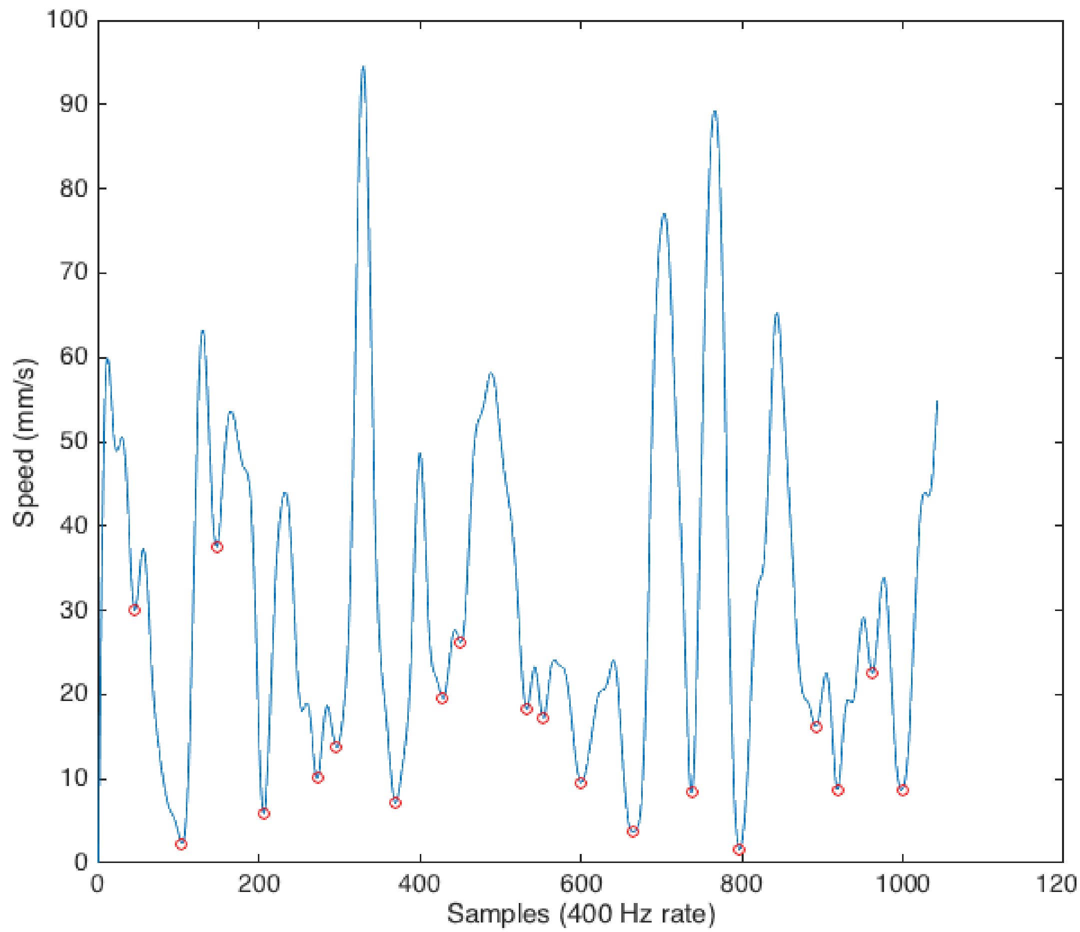

Figure 1. Speed plot for T2 using the *Suits* sentence (*It's time to shop for two new suits*) with speed minima (circles).

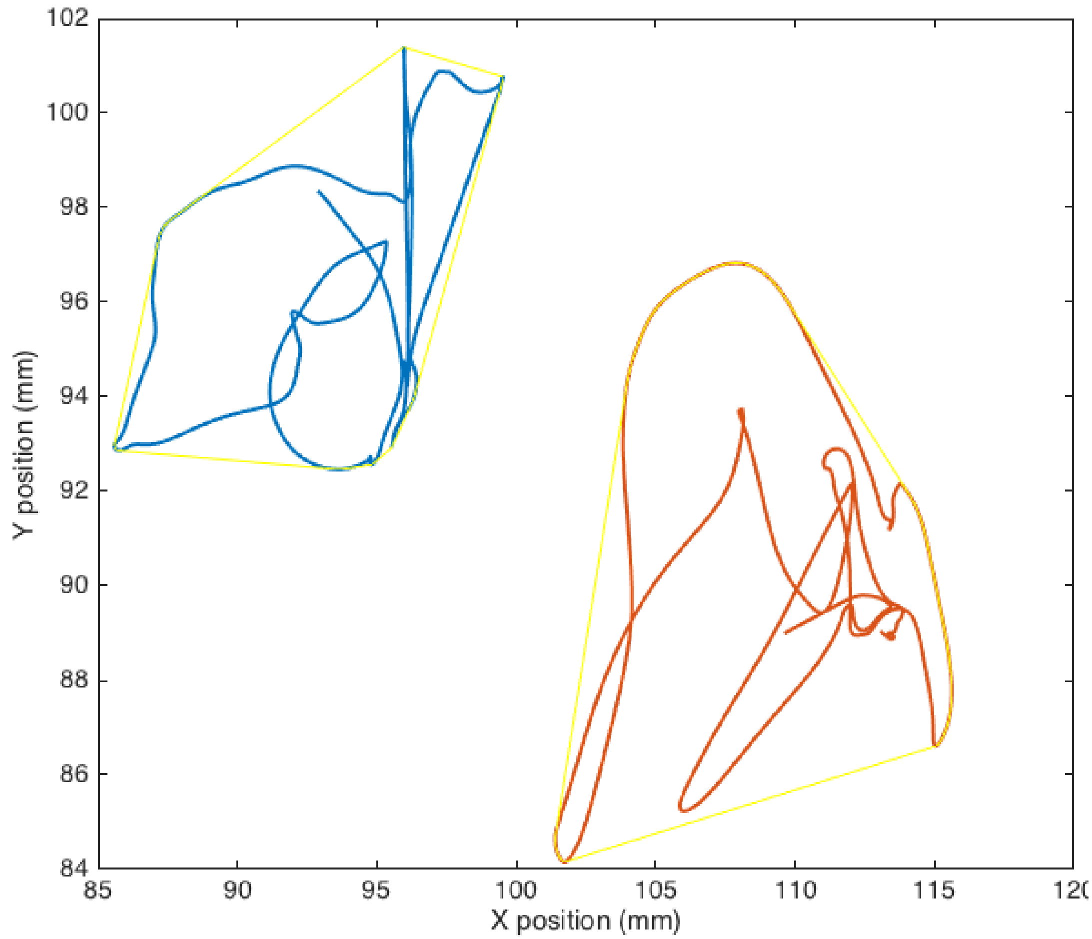

Figure 2. Movement in the x-y direction for both tongue makers throughout the *Suits* sentence (*It's time to shop for two new suits*) with the hull volume.

Results

Independent Variable: Sound Pressure Level

A repeated measures ANOVA was used to analyze the results of this study for the variables that were measured across the three vocal effort levels of soft, comfortable, and loud. Vocal intensity was the independent variable, and it was measured in dB SPL at 100 cm. The participants were asked to speak at what they judged to be soft ($M = 70.89$, $SD = 5.63$), comfortable ($M = 77$, $SD = 3.95$), and loud levels ($M = 84.64$, $SD = 4.83$). Intensity changed significantly across the experimental conditions, $F(2, 36) = 211.86$, $p < .001$. Soft, $F(1, 18) =$

79.23, $p < .001$, and loud, $F(1,18) = 203.\ 30$, $p < .001$, levels were found to be significantly different from the comfortable intensity. Since there were no significant differences in performance between men and women for the dependent variables in the study, gender was not retained as a factor for statistical analysis.

Table 2

Descriptive Statistics for Acoustic and Kinematic Measures for T1 and T2 Across the Intensities

	Soft		Comfortable		Loud	
	M	*SD*	M	*SD*	M	*SD*
VSA	129062	114984	114171	51797	126635	126635
KVSA T1	13.9	16.5	10.7	6.0	14.0	10.0
KVSA T2	14.0	13.5	10.6	9.0	15.0	10.1
VAI	.857	.084	.834	.048	.834	.050
KVAI T1	1.045	.140	1.071	.102	1.089	.076
KVAI T2	.912	.117	.935	.075	.942	.064

Table 3

Descriptive Statistics for Stroke Metrics Across the Intensities for T1 and T2

	Soft		Comfortable		Loud	
	M	*SD*	M	*SD*	M	*SD*
Sentence Duration	2.14	.25	2.10	.25	2.33	.36
Stroke Count T1	22.07	4.51	21.56	3.30	23.18	4.45
Onset Speed T1	25.78	6.47	29.16	6.42	30.73	5.92
Peak Speed T1	88.16	22.85	94.30	23.68	106.27	21.32
Mean Speed T1	53.75	12.91	58.23	13.13	64.31	12.23
Distance T1	6.38	1.40	6.83	1.28	7.65	1.43
Stroke Duration T1	109.39	11.60	108.36	10.14	108.92	12.87
Hull Volume T1	166.06	73.89	193.34	69.69	240.13	80.49
Stroke Count T2	21.93	3.82	22.01	3.68	23.32	3.24
Onset speed T2	26.78	6.28	29.26	4.44	33.18	7.95
Peak Speed T2	81.24	16.04	87.37	14.78	96.37	17.47
Mean Speed T2	50.45	9.92	54.31	8.04	60.75	11.34
Distance T2	6.13	1.24	6.43	.977	7.00	1.23
Stroke Duration T2	111.29	9.41	107.92	7.36	105.96	9.60
Hull Volume T2	162.85	66.90	194.56	68.43	245.78	89.91

Acoustic Measures

Vowel space area. The overall repeated measures ANOVA did not show significance for changes in VSA across the loudness conditions (Table 2). When the comfortable intensity was contrasted with both the soft and loud intensities, there was no statistical significance. See Table 4 for detailed statistical results.

Vowel articulation index. The repeated measures ANOVA did not show significance for changes in VAI across the loudness conditions. When the comfortable intensity was compared to both the soft and loud intensities, there were no statistically significant differences (Table 4).

Kinematic Measures

Kinematic-vowel space area. The repeated measures ANOVA did not show significance for changes in KVSA T1 or KSVA T2. When the comfortable intensity was contrasted with both the soft and loud intensities, there were no significant differences (Table 4).

Kinematic-vowel articulation index. The repeated measures ANOVA did not show significance for changes in KVAI T1 or KVAI T2. When the comfortable intensity was contrasted with both the soft and loud intensities, there were no significant differences (Table 4).

Table 4

ANOVA Results for VSA, KVSA, VAI, and KVAI Across the Intensities for T1 and T2

	Main ANOVA				Soft Contrast				Loud Contrast			
	df	F	*p*	η_p^2	*df*	F	*p*	η_p^2	*df*	F	*p*	η_p^2
VSA	1.24, 23.55	.403	.575	.021	1, 19	.519	.480	.027	1, 19	2.60	.123	.120
KVSA T1	1.45, 27.52	.592	.508	.030	1, 19	.739	.401	.037	1, 19	2.60	.124	.120
KVSA T2	2, 38	1.09	.347	.054	1, 19	1.28	.272	.063	1, 19	3.00	.102	.135
VAI	1.30, 24.60	1.81	.191	.087	1, 19	1.89	.185	.091	1, 19	.022	.884	.001
KVAI T1	1.62, 30.68	1.17	.316	.058	1, 19	.845	.370	.043	1, 19	.736	.402	.037
KVAI T2	1.47, 27.99	.718	.456	.036	1, 19	.948	.343	.048	1, 19	.127	.725	.007

Sentence-Level Stroke Metrics

The repeated measure ANOVA revealed statistically significant changes for all of the stroke metrics for both T1 and T2 across the intensities. These stroke metrics were taken from the entire sentence *It's time to shop for two new suits*, in contrast to the VSA and VAI measures, which were based on specific vowels extracted from two sentences.

Stroke metrics for T1. There was a significant main effect for sentence duration, onset speed, peak speed, mean speed, distance, and hull volume (Table 3). Stroke count and stroke duration did not change significantly. The comfortable intensity was contrasted with both the soft and loud intensities and statistically significant differences were found for some of the stroke metrics. Table 5 summarizes the results of the repeated measures ANOVA.

Table 5

Stroke Metrics Repeated Measure ANOVA Results for T1

	Main ANOVA				Soft Contrast				Loud Contrast			
	df	F	*p*	η_p^2	*df*	F	*p*	η_p^2	*df*	F	*p*	η_p^2
Sentence Duration	2, 36	11.99	*p*< .001	.400	1, 18	.479	.498	.026	1, 18	15.30	.001	.459
Stroke Count	2, 36	1.86	.170	.094	1, 18	.388	.541	.021	1, 18	2.47	.134	.120
Onset Speed	2, 36	6.84	.003	.275	1, 18	11.22	.004	.384	1, 18	.924	.349	.049
Peak Speed	1.58, 28.42	12.39	*p< .001*	.408	1, 18	5.31	.033	.228	1, 18	6.79	.018	.274
Mean Speed	2, 36	11.42	*p< .001*	.388	1, 18	6.67	.019	.270	1, 18	5.15	.036	.222
Distance	2, 36	12.66	*p< .001*	.413	1, 18	3.55	.076	.165	1, 18	7.12	.016	.283
Stroke Duration	2, 36	.094	*.910*	.005	1, 18	.213	.650	.012	1, 18	.045	.835	.002
Hull Volume	2, 36	20.00	*p< .001*	.526	1, 18	7.10	.016	.283	1, 18	12.64	.002	.412

Sentence duration changed significantly. The comfortable loudness sentence duration was not significantly longer than the soft condition, but the loud sentence duration was significantly longer than the comfortable sentence duration.

Onset speed also changed significantly. Contrasts revealed that it was higher for the comfortable than the soft condition, but there was no difference between the comfortable and loud conditions. Peak speed changed significantly. Compared to the comfortable loudness condition, the peak speed was slower for the soft, and faster for the loud condition. Mean speed changed significantly. Compared to the comfortable loudness condition, the mean speed was slower for the soft, and faster for the loud condition. Distance changed significantly. Contrasts revealed that it was larger for the loud condition than the comfortable condition, but there was no difference between the comfortable and soft conditions. Hull volume changed significantly. Compared to the comfortable condition, the hull volume was smaller for the soft condition, and larger for the loud condition (Figure 3).

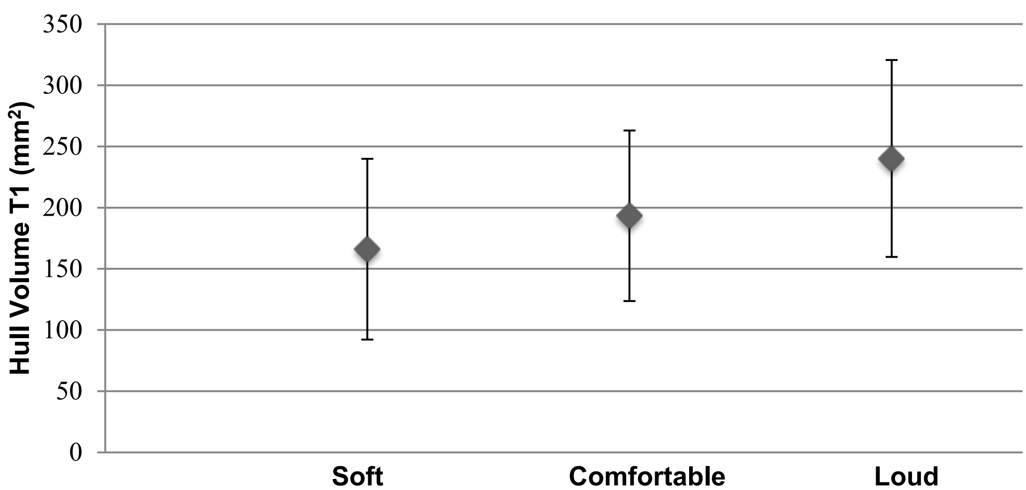

Figure 3. Hull volume for T1 across the intensities with standard deviation error bars.

Stroke metrics for T2. There was a significant main effect for sentence duration, onset speed, peak speed, mean speed, distance, and hull volume for T2. Stroke count and stroke duration did not change significantly. The comfortable intensity was contrasted with both the soft and loud intensities and statistical significant differences were found for some of the stroke metrics. Table 6 summarizes the results of the repeated measures ANOVA.

Table 6

Stroke Metrics Repeated Measures ANOVA Results for T2

	Main ANOVA				Soft Contrast				Loud Contrast			
	df	F	*p*	η_p^2	*df*	F	p	η_p^2	*df*	F	*p*	η_p^2
Sentence Duration	2, 36	11.99	$p < .001$	.400	1, 18	.479	.498	.026	1, 18	15.30	.001	.459
Stroke Count	2, 36	2.97	.064	.142	1, 18	.014	.906	.001	1, 18	3.41	.081	.159
Onset Speed	2, 36	21.74	$p < .001$	.547	1, 18	7.31	.015	.289	1, 18	10.96	.004	.378
Peak Speed	2, 36	22.72	$p < .001$	.558	1, 18	10.71	.004	.373	1, 18	11.64	.003	.393
Mean Speed	2, 36	27.81	$p < .001$	.607	1, 18	10.58	.004	.370	1, 18	13.90	.002	.436
Distance	2, 36	8.49	.001	.321	1, 18	2.96	.103	.141	1, 18	5.29	.034	.227
Stroke Duration	2, 36	3.14	.055	.148	1, 18	3.07	.097	.146	1, 18	.913	.352	.048
Hull Volume	2, 36	39.12	$p < .001$	.685	1, 18	14.21	.001	.441	1, 18	22.43	$p < .001$	.555

Sentence duration changed significantly. Contrasts revealed that it was longer for the loud condition than the comfortable condition, but there was no difference between the comfortable and soft conditions.

Onset speed changed significantly. Compared to the comfortable loudness condition, the onset speed was slower for the soft, and faster for the loud condition. Peak speed changed significantly. Compared to the comfortable loudness condition, the peak speed was slower for the soft condition, and faster for the loud condition. Mean speed changed significantly. Compared to the comfortable loudness condition, the mean speed was slower for soft condition, and faster for the loud condition. Distance changed significantly. Contrasts revealed that it was longer for the loud condition, but there was no difference between the comfortable and soft conditions. Hull volume changed significantly. Compared to the comfortable loudness condition, the hull volume was smaller for the soft, and larger for the loud condition (Figure 4).

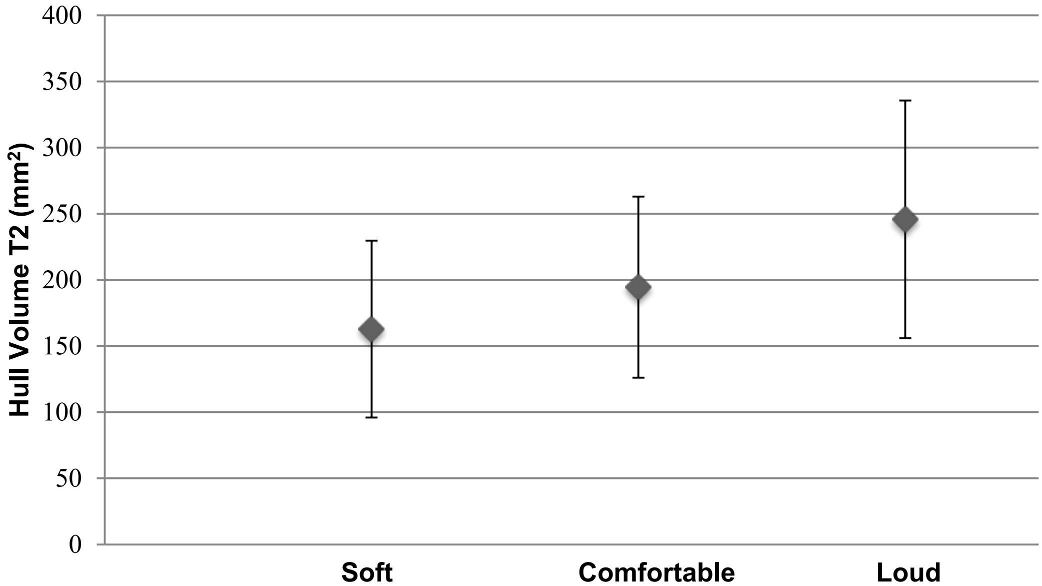

Figure 4. Hull volume for T2 across the intensities with standard deviation error bars.

AAVS and AKVS Sentence-Level Metrics

The sentence-level metrics AAVS and AKVS were computed in Ohio by Dr. Jason Whitfield. He ran the audio and kinematic recordings of two sentences through his AAVS and AKVS algorithm (Whitfield & Goberman, 2014). The two sentences that were used for these measures were *A good AC should keep your car cool* and *It's time to shop for two new suits.*

A paired t-test was used to analyze AAVS and AKVS for the comfortable and loud conditions. It was not possible to compute the AAVS measure for the soft condition because the recordings lacked the acoustic intensity to adequately drive the formants throughout the voiced segments. Since the soft condition was unusable, only two of the loudness conditions were compared, resulting in the use of a paired t-test instead of a RM ANOVA that was used for the other dependent measures across the three vocal effort conditions. AAVS changed significantly for the AC sentence and AKVS changed significantly for both sentences (Table 7).

Table 7

Descriptive Statistics for AAVS and AKVS for T1 and T2 for the Comfortable and Loud Intensities

	Comfortable		Loud	
	M	SD	M	SD
AC AAVS	47623	16145	57918	18655
AC AKVS T1	19.4	7.6	23.9	9.3
AC AKVS T2	19.1	7.8	24.4	8.9
Suits AAVS	38362	14672	39465	16248
Suits AKVS T1	15.3	4.7	19.3	6.4
Suits AKVS T2	10.9	3.0	12.2	2.9

Note. AC refers to the sentence *A good AC should keep your car cool* and Suits refers to the sentence *It's time to shop for two new suits.*

Since AKVS was able to include all three intensities, a repeated measures ANOVA was also used to analyze this kinematic measure. Significant changes were found across the three conditions for both sentences (Table 8).

Table 8

Descriptive Statistics for AKVS Across the Intensities for T1 and T2

	Soft		Comfortable		Loud	
	M	SD	M	SD	M	SD
AC AKVS T1	15.8	6.4	19.4	7.6	23.9	9.3
AC AKVS T2	15.9	7.8	19.1	7.8	24.4	8.9
Suits AKVS T1	12.3	4.7	15.3	4.7	19.3	6.4
Suits AKVS T2	9.4	3.0	10.9	3.0	12.2	3.0

Note. AC refers to the sentence *A good AC should keep your car cool* and Suits refers to the sentence *It's time to shop for two new suits.*

Articulatory-acoustic vowel space. A paired samples t-test was used to compare the AAVS of the comfortable condition to the loud condition. This test revealed that the loud condition had a larger AAVS than the comfortable condition for the AC sentence, but AAVS for the Suits sentence did not differ significantly (Table 9, Figure 5).

Table 9

Paired t-test for AAVS and AKVS for T1 and T2 Comparing the Comfortable and Loud Intensities

	t	p
AC AAVS	-5.694	p<.001
AC AKVS T1	-5.884	p<.001
AC AKVS T2	-5.411	p<.001
Suits AAVS	-.720	.481
Suits AKVS T1	-4.293	p<.001
Suits AKVS T2	-2.617	.017

Note. AC refers to the sentence *A good AC should keep your car cool* and Suits refers to the sentence *It's time to shop for two new suits.*

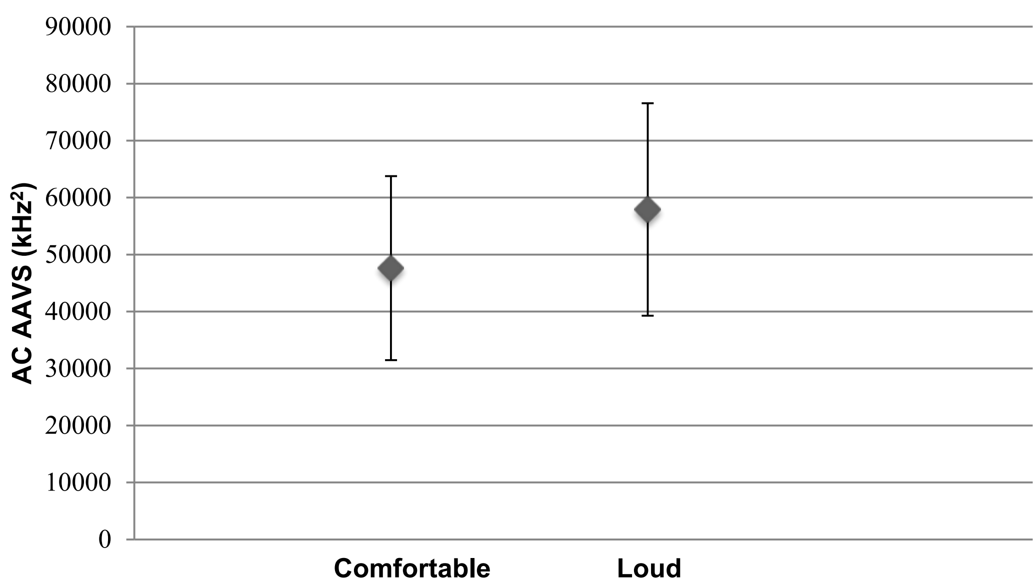

Figure 5. AC sentence (*A good AC should keep your car cool*) AAVS for comfortable and loud intensities with standard deviation error bars.

Articulatory-kinematic vowel space. A paired t-test was used to compare the comfortable condition and the loud condition for AKVS in order to equate more directly with the acoustic measure (AAVS) which was only available for comfortable and loud conditions. Significant differences were found between the two conditions. The loud condition had a larger AKVS than the comfortable condition (Table 9).

The repeated measures ANOVA showed significant changes in AKVS across the three loudness conditions. When compared to the comfortable condition, the soft condition was smaller and the loud condition was larger for both sentences (Table 10, Figure 6, Figure 7).

Table 10

ANOVA Results for AKVS Across the Intensities for T1 and T2

	Main ANOVA				Soft Contrast				Loud Contrast			
	df	F	*p*	p2	*df*	F	*p*	p^2	*df*	F	*p*	p^2
AC AKVS T1	2,36	46.9	*p*<.001	.723	1,18	28.2	*p*<.001	.611	1,18	34.6	*p*<.001	.658
AC AKVS T2	2,36	45.4	*p*<.001	.716	1,18	17.9	*p*<.001	.499	1,18	29.3	*p*<.001	.619
Suits AKVS T1	1.49, 26.9	25.0	p<.001	.581	1,18	18.0	*p*<.001	.500	1,18	18.4	*p*<.001	.506
Suits AKVS T2	2,36	14.0	*p*<.001	.437	1,18	12.0	.003	.401	1,18	6.8	.017	.276

Note. AC refers to the sentence A good AC should keep your car cool and Suits refers to the sentence It's time to shop for two new suits.

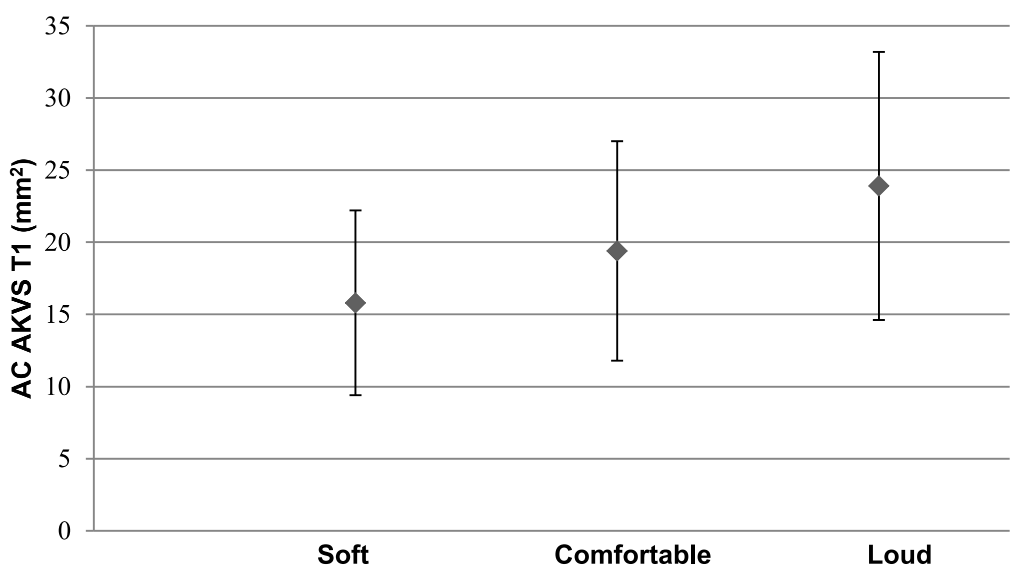

Figure 6. AC sentence (*A good AC should keep your car cool*) AKVS for T1 across the intensities with standard deviation error bars.

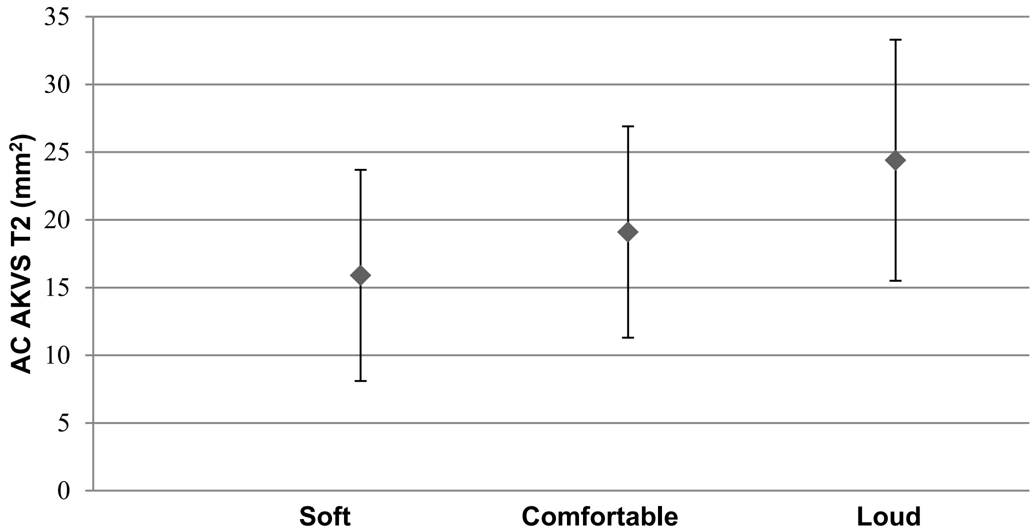

Figure 7. AC sentence (*A good AC should keep your car cool*) AKVS for T2 across the intensities with standard deviation error bars.

Discussion

The purpose of this study was to compare acoustic and kinematic measures of articulation across soft, comfortable, and loud intensities, because articulation was anticipated to change with adjustments to vocal effort. The validity of the acoustic measures in reflecting articulatory activity was evaluated by comparing them with the equivalent kinematic measures.

Acoustic Measures

Vowel space area. Previous research has used VSA to measure the articulatory working space of vowels (Sandoval et al., 2013; Sapir et al., 2010). It has been reported that VSA may not be able to detect differences in mild or moderate dysarthria when compared to normal speech (Neel, 2008). Sapir et al. (2010) explained that VSA is sensitive to interspeaker variability, which could mask the difference between typical speech and disordered speech. Another limitation of VSA is that it is based on vowel mid-points. As a result, the acoustic information that occurs between the vowel mid-points is not included in the VSA measure (Fridland et al., 2014). The current findings are consistent with previous reports that VSA may not always be able to detect differences between speech conditions (Sandoval et al., 2013; Sapir, Spielman, Ramig, Story, & Fox, 2007).

Vowel articulation index. Like VSA, some studies have shown that VAI is a sufficiently sensitive measure to reflect changes in articulatory function after treatment, while in other studies, it has failed to differentiate the speech of individuals with and without dysarthria (Roy et al., 2009). VAI was developed to reduce the influence of interspeaker variability that is associated with traditional vowel area measures. It has been demonstrated that VAI can be more effective than VSA at differentiating normal from dysarthric speech (Sapir et al., 2007). The present findings did not support the conclusions of this earlier research, in that VAI did not reflect intraspeaker articulation changes any better than VSA across a wide intensity range.

Articulatory-acoustic vowel space. AAVS is a novel acoustic measure that appears promising in its ability to reflect differences and/or changes in speech articulation. AAVS revealed differences between habitual and clear speech in speakers with PD (Whitfield & Goberman, 2014). Connected-speech VSA has recently been used to determine the peripheral vowel space area using a convex-hull algorithm. This measure yielded a more complete picture of vowel working space than the traditional quadrangle or triangle VSA measure (Sandoval et al., 2013). In the current study, AAVS detected differences in articulation for comfortable and loud speech when isolated vowel measures, such as VSA and VAI, were not sensitive enough to do this. For example, the AC sentence showed significant AAVS changes across the intensity conditions.

We also found that AAVS has its limitations. Although the AC sentence AAVS changed significantly, the Suits sentence AAVS did not change significantly across the intensities. This may be due to the fact that acoustic measures are an indirect reflection of the movements that are occurring, rather than a measure of the actual movement. The Suits sentence AAVS may not have changed significantly because it included different words and phoneme combinations than the AC sentence. For example, the AC sentence included velar sounds and a variety of vowels, while the Suits sentence included alveolar sounds and the vowel /u/.

AAVS and other continuous speech measures may be able to detect changes in speech because these measures track the transitions between vowels and consonants, whereas point measures only include an isolated phoneme that has been extracted from a word. An example of transition tracking would be to observe the movement of the tongue as it transitions from touching the alveolar ridge for the /t/ consonant in the word *to* to being in the low, back position for the /ɑ/ vowel in *shop*.

Kinematic Measures

Kinematic-vowel space area. KVSA is a new measure that was developed during this study and is the kinematic equivalent to the VSA acoustic measure. KVSA was computed from the x/y positions of the tongue sensors at the same exact time as the F1 and F2 corner vowel midpoints that were used in computing VSA. Like the acoustic VSA, the KVSA measure also lacked the sensitivity to reflect intra-speaker kinematic changes across the intensity levels. This measure may have failed to detect changes in speech movements because it relied on static articulator position data at three isolated points in time, rather than representing dynamic articulatory activity throughout an utterance. This finding is consistent with other reports that isolated measures based on one or more points in time do not provide enough information to represent the dynamic movements that occur during speech (Sandoval et al., 2013; Whitfield & Goberman, 2014).

Kinematic-vowel articulation index. KVAI is a new measure that was developed during this study; it is the kinematic equivalent of VAI. In the current study, KVAI was not sensitive to changes in articulation across the speech intensity levels. This kinematic measure may have had the same weaknesses as its acoustic equivalent. KVAI is a measure based on the same isolated corner vowels as the VSA and VAI, and only reflects the position of the tongue at specific points during the utterance. This is a weakness because articulation involves the dynamic movements of many structures, such as the lips, tongue, and jaw, rather than a series of isolated postures of the articulators (Tasko & Westbury, 2002).

Articulatory-kinematic vowel space. AKVS is a new measure that was developed during this study and is the kinematic equivalent of AAVS. Unlike the kinematic measures based on individual vowels, AKVS is a continuous speech measure of articulation throughout an utterance. In the present study, AKVS was found to change significantly across the three intensities for both sentences. AKVS revealed changes in speech for the comfortable and loud conditions for both sentences, while AAVS only showed these changes in the AC sentence. A limitation of AAVS is that it is based on formant tracks, which can be difficult to compute for many consonants. The AKVS, on the other hand, includes the movements involved in producing all sounds in an utterance. This may explain the significant AKVS findings in both sentences, compared to the single sentence for AAVS.

Stroke Metrics

Along with AAVS and AKVS, the stroke metrics were based on the entire sentence, rather than on three individual vowels. Many of the stroke metric changes were highly significant across the three intensities. Previous research has highlighted the advantages that stroke metrics provide. For example, stroke metrics can be used to compare the speech of different languages because this approach does not require the stimulus to include specific phonemes (Nissen, Dromey, & Wheeler, 2007). Another advantage is that long and heterogeneous samples of speech movement can be compared across a variety of speakers, allowing for studies that include a large number of participants (Tasko & Westbury, 2002). These metrics are also simple to compute, because vowels do not need to be individually identified and extracted from words; instead, a whole sentence can be analyzed quickly by the software algorithm.

A study by Dromey and Ramig (1998) analyzed lip opening and closing during the production of the syllable /pæp/ across the following five vocal effort conditions: ¼ loudness, ½ loudness, normal loudness, 2x as loud, and 4x as loud. The authors measured upper and lower lip displacement and peak velocity. They found that for the opening and the closing movements of the lips, the upper and lower lip displacement increased as the speaker's intensity increased. They also found that peak velocity of the upper and lower lips increased with intensity.

In the present study, onset speed, peak speed, mean speed, and hull volume all increased with intensity for both T1 and T2. This pattern showed that as speech intensity increased, the articulatory working space and the tongue movement speed increased as well. It was found that the stroke duration decreased with intensity. This is consistent with the changes in the other variables, because as the speed of tongue movement increased, the stroke duration had to decrease to allow for the articulators to transition to their new positions more quickly. The results from the present study support previous research that has shown that as speech intensity increases, articulatory working space also increases (Dromey & Ramig, 1998; Sapir et al., 2007; Schulman, 1989; Tjaden & Wilding, 2004).

Stroke metrics also reflect the dynamic behavior of the articulators in producing both consonants and vowels in connected speech, which parallels earlier studies of articulator movement in the transition from one sound to another during the production of a syllable. This contrasts with VSA, VAI, KVSA, and KVAI, which are based on the positions of the articulators at one instant in time – the corner vowel midpoint. The inclusion of consonants and vowels in the stroke metric measures might have contributed to their ability to distinguish differences between the three speech intensities for both sentences that the point metrics were not able to detect.

Stroke metrics also have some drawbacks. A disadvantage of stroke metrics is that they are averages across an utterance (Tasko & Westbury, 2002). This means that they do not provide detailed information for the transition between specific, individual phonemes or provide information about how the vowels and consonants in the sentence vary in their kinematic features. Despite this disadvantage, the present study demonstrated that stroke metrics can be used to compare a variety of speech samples, and can reveal significant changes across intensity levels.

Limitations of the Present Study and Directions for Future Research

A limitation of the present study was that participants self-selected the soft, comfortable, and loud intensities. In future studies, the SPL level could be standardized to reduce inter-speaker variability across the speaking conditions. Another limitation of the present study was that a triangle VSA was used instead of a quadrilateral VSA. The quadrilateral VSA is preferred over the triangle VSA because it includes all four corner vowels, allowing for a more comprehensive picture of the articulatory working space. The stimuli were not designed a priori to facilitate the extraction of VSA, resulting in only three of the four corner vowels being present in the stimulus sentences. Another limitation is that the present experiment did not include a clear speech condition. This condition would have provided a variable other than intensity to influence articulatory activity in order to evaluate the sensitivity of the acoustic and kinematic metrics. Another drawback of the present study is that the participants were all young adults who were healthy speakers. The study could have included participants from a variety of age groups in order to study how age and intensity affect speech. Future research could include a disordered group with dysarthria. This would allow the healthy, young adult group to serve as a control for the disordered group.

Conclusion

The current study showed the strengths and weaknesses of several acoustic and kinematic measures of speech. Because of the nonlinear relationship between speech movement and acoustic measures, this study provided data to better understand the association between these measures (Stevens & Keyser, 2010). Isolated acoustic vowel measures were not sensitive enough to detect differences between soft, comfortable, and loud speech. The kinematic equivalents of these measures were also unable to detect differences. Sentence-level acoustic and kinematic measures were able to detect significant changes in speech as it changed with intensity. Stroke metrics were able to add to these sentence-level measures by including the dynamic movement of T1 and T2 throughout the utterance instead of a just a momentary position of a vowel midpoint.

References

Barlow, S. M., Cole, K. J., & Abbs, J. H. (1983). A new head-mounted lip-jaw movement transduction system for the study of motor speech disorders. *Journal of Speech and Hearing Research, 26*, 283-288.

Berisha, V., Sandoval, S., Utianski, R. L., Liss, J. M., & Andreas, S. (2013). Characterizing the distribution of the quadrilateral vowel space area. *Journal of the Acoustical Society of America, 135*, 421-427. doi:10.1121/1.4829528

Berry, J. J. (2011). Accuracy of the NDI wave speech research system. *Journal of Speech, Language, and Hearing Research, 54*, 1295-1301. doi:10.1044/1092-4388(2011/10-0226)

Boersma, P., & Weenink, D. (2015). Praat (Version 5.4.06).

Dromey, C., Jang, G., & Hollis, K. (2013). Assessing correlations between lingual movement and formants. *Speech Communication 55*, 315-328. doi:10.1016/j.specom.2012.09.001

Dromey, C., & Ramig, L. O. (1998). Intentional changes in sound pressure level and rate: Their impact on measures of respiration, phonation, and articulation. *Journal of Speech, Language, and Hearing Research, 41*, 1003-1018. doi:1092-4388/98/4105-1003

Fridland, V., Kendall, T., & Farrington, C. (2014). Durational and spectral differences in American English vowels: Dialect variation within and across regions. *Journal of the Acoustical Society of America, 136*, 341-349. doi:10.1121/1.4883599

Kleinow, J., & Smith, A. (2006). Potential interactions among linguistic, autonomic, and motor factors in speech. *Developmental Psychobiology 48*, 275-287. doi:10.1002/dev.2 0141

MathWorks. (2014). MATLAB (Version 8.5).

Neel, A. T. (2008). Vowel space characteristics and vowel identification accuracy. *Journal of Speech, Language, and Hearing Research, 51*, 574-585. doi:10.1044/1092-4388(2008/041)

Nissen, S. L., Dromey, C., & Wheeler, C. (2007). First and second language tongue movements in Spanish and Korean bilingual speakers. *Phonetica, 64*, 201-216. doi:10.1159/000121373

Roy, N., Nissen, S. L., Dromey, C., & Sapir, S. (2009). Articulatory changes in muscle tension dysphonia: Evidence of vowel space expansion following manual circumlaryngeal therapy. *Journal of Communication Disorders, 42*, 124-135. doi:10.1016/j.jcomdis.2008.10.001

Sandoval, S., Berisha, V., Utianski, R. L., Liss, J. M., & Spanias, A. (2013). Automatic assessment of vowel space area. *Journal of the Acoustical Society of America 134*, EL477-483. doi:10.1121/1.4826150

Sapir, S., Ramig, L. O., Spielman, J. L., & Fox, C. (2010). Formant centralization ratio: A proposal for a new acoustic measure of dysarthric speech. *Journal of Speech, Language, and Hearing Research, 53*, 114-125. doi:10.1044/1092-4388(2009/08-0184)

Sapir, S., Spielman, J. L., Ramig, L. O., Story, B. H., & Fox, C. (2007). Effects of intensive voice treatment (the Lee Silverman Voice Treatment [LSVT]) on vowel articulation in dysarthric individuals with idiopathic Parkinson disease: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research, 50*, 899-912. doi:10.1044/1092-4388(2007/064)

Schulman, R. (1989). Articulatory dynamics of loud and normal speech. *Journal of the Acoustical Society of America, 85*, 295-312. doi:10.1121/1.397737

Skodda, S., Visser, W., & Schlegel, U. (2011). Vowel articulation in Parkinson's disease. *Journal of Voice, 25*, 467-472. doi:10.1016/j.jvoice.2010.01.009

Stevens, K. N., & Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics 38*, 10-19. doi:doi:10.1016/j.wocn.2008.10.004

Tasko, S. M., & Westbury, J. R. (2002). Defining and measuring speech movement events. *Journal of Speech, Language, and Hearing Research, 45*, 127-142. doi:10.1044/1092-4388(2002/010)

Tjaden, K., & Wilding, G. E. (2004). Rate and loudness manipulations in dysarthria: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research, 47*, 766-783. doi:10.1044/1092-4388(2004/058)

Whitfield, J. A., & Goberman, A. M. (2014). Articulatory-acoustic vowel space: Application to clear speech in individuals with Parkinson's disease. *Journal of Communication Disorders 51*, 19-28. doi:10.1016/j.jcomdis.2014.06.005

Appendix A

Annotated Bibliography

Barlow, S. M., Cole, K. J., & Abbs, J. H. (1983). A new head-mounted lip-jaw movement transduction system for the study of motor speech disorders. *Journal of Speech and Hearing Research, 26*, 283-288.

Objective: The purpose of this paper was to describe the features of the head-mounted lip-jaw movement transduction system, how it measures speech movements, and the advantages and disadvantages of using this system.

Method: The head-mounted lip-jaw transduction system measures speech movements through the use of strain gauges that are attached to lips and jaw using a skin bond. The participants were asked to move their head in different directions as well as repeat the syllable /pɑ/ at a rate of 4 repetitions per second. The researchers used the head-mounted lip-law movement transduction system to compare the movements of disordered speech to typical speech.

Results: The researchers found that head movement was reduced for participants with disordered speech. They also found that speakers with spastic cerebral palsy had a great deal of variability in their speech movements between repetitions, specifically in the areas of range, slope, and smoothness.

Conclusion: This study supports the use of kinematic measurements to gain greater insight into how disordered speech differs from typical speech. Knowing more about disordered speech can lead to improved, physiologically-based rehabilitative speech programs.

Relevance to the current work: This paper described the head-mounted lip-jaw movement transduction system in great detail. The current work represents a progression, in that more sophisticated kinematic technology was used than in this 1983 study.

Berisha, V., Sandoval, S., Utianski, R. L., Liss, J. M., & Andreas, S. (2013). Characterizing the distribution of the quadrilateral vowel space area. *Journal of the Acoustical Society of America, 135*, 421-427. doi:10.1121/1.4829528

Objective: The objective of this study was to explain the characteristics of the quadrilateral vowel space area (VSA).

Method: This study had 630 participants from 8 different dialect regions. This study included recordings of the four corner vowels to compute the quadrilateral VSA using F1 and F2. Higher-level statistics were used to analyze the vowel information to show the relationship between VSA and speech intelligibility. The researchers compared the empirical estimates from the study to the theoretically predicted values that are present in the Hillenbrand and TIMIT data sets.

Results: The results showed that the Hillenbrand data and the TIMIT data were similar to the data in this study. This supports the idea that the empirical estimates were very similar to the theoretically predicted values. This establishes the validity of the results from the study.

Conclusions: This study found that using higher-order statistics led to understanding the robustness of the VSA. The higher-order statistics presented more characteristics of VSA by better capturing the shape of the vowel distribution.

Relevance to the current work: This study demonstrated the contribution of VSA, including its relationship to speech intelligibly. This paper was used in the current study to explain how VSA is calculated, and to establish VSA as a valid acoustic speech measure.

Berry, J.J. (2011). Accuracy of the NDI wave speech research system. *Journal of Speech, Language, and Hearing Research*, 54, 1295-1301. doi:10.1044/1092-4388(2011/10-0226)

Objective: The NDI Wave system measures articulator movement using three dimensional tracking in the x, y, and z planes. This study examined the NDI Wave system's accuracy for static and dynamic tracking of position of the articulators during speech.

Method: This study included three experiments: static rigid-body tracking across different locations in the electromagnetic field, dynamic rigid-body tracking across different locations in the electromagnetic field, and human jaw movement during speech. The electromagnetic field was set up to allow for the tracking for all three experiments. For the static rigid-body experiment, six sensors were attached to a measurement surface that was placed oblique to the surface of a table. The six dynamic rigid-body sensors were glued to a kinematic chain made of Lego building blocks. The five speech movement sensors were glued to the tongue, lips, and jaw. The sixth sensor was glued between the eyebrows. The magnetic field generator was placed 5 cm from the left ear of each subject while they read the Farm Passage. The NDI Wave system tracked the movement of the sensors as the participants read the passage.

Results: The results showed that the tracking errors of the NDI Wave system might be larger than the tracking errors that are found when using the Carstens AG500 EMA system. The static tracking accuracy was found to be superior to the dynamic tracking accuracy. The tracking accuracy of human jaw movement was superior to the dynamic tracking accuracy.

Conclusions: Although the NDI wave system may not be as accurate as the Carstens AG500 EMA system, it is still a valid system for measuring position. It is easy to use and produces acceptably accurate kinematic measurements.

Relevance to the current work: This study explained the NDI Wave system in great detail. Since the current work used this EMA system, it was appropriate to discuss the characteristics of the EMA system and explain how it functions.

Dromey, C., Jang, G., & Hollis, K. (2013). Assessing correlations between lingual movement and formants. *Speech Communication, 55*, 315-328. doi: 10.1016/j.specom.2012.09.001

Objective: First and second formants are often used to study normal and disordered speech. This study examined how well formants are able to reflect lingual movement during the production of a diphthong.

Method: The study had 20 native American English speakers from the western United States. Each participant produced four diphthongs in a sentence while their tongue movement was tracked using a magnetic tracking system. The researchers correlated the vertical tongue movement with F1 and anteroposterior movement with F2 during the transition phase of the diphthong.

Results: The results showed that for the most part, the formants were an accurate reflection of the kinematic movements throughout the diphthong. There were some instances where the formants were not a good reflection of the kinematics.

Conclusions: When the researchers looked more into the reason for some discrepancies between the formants and the kinematics, they found that other factors affected the formant production. Among these factors were coarticulation and motor equivalence. Nonlinearities between the kinematic and acoustic measures were also a factor.
Relevance to the current work: The current work examined how well acoustic measures reflect the kinematic measures for the corner vowels and in connected speech. This study gives background information about vowel formants and supports the use of acoustic and kinematic measures for speech. This study also pointed out some of the problems that we might find when comparing acoustic and kinematic measures.

Dromey, C., & Ramig, L. O. (1998). Intentional changes in sound pressure level and rate: Their impact on measures of respiration, phonation, and articulation. *Journal of Speech, Language, and Hearing Research, 41*, 1003-1018.
Objective: The purpose of this study was to analyze how changes in sound pressure level (SPL) and rate affect respiration, phonation, and articulation during sentence production.
Method: This study included 5 male and 5 female participants. The participants were asked to say the sentence *I sell a sapapple again* under 5 SPL and 5 rate conditions. The following measures were then made: lung volume, SPL, fundamental frequency, semitone standard deviation, upper and lower lip displacement, and peak velocities.
Results: Loud speech led to increases in lung volume initiation and termination, fundamental frequency, semitone standard deviation, and articulatory displacement and peak velocities for both lips. The study also found that lung volume decreased as rate increased and lower lip displacement was smaller for faster speech.
Conclusions: The study concluded that respiration, phonation, and articulation change with SPL and rate of speech during sentence production.
Relevance to the current work: The present study used loudness as the independent variable. This study supported the use of loudness as a variable because it found changes in speech kinematics with loudness.

Fridland, V., Kendall, T., & Farrington, C. (2014). Durational and spectral differences in American English vowels: Dialect variation within and across regions. *Journal of the Acoustical Society of America, 136*, 341-349. doi: 10.1121/1.4883599
Objective: Spectral differences have been studied across different dialects of American English but duration variability has not been studied extensively. This study examined how duration variability affects spectral changes in different dialects.
Method: The participants in this study were taken from universities in different regions of the United States. The 14 southern dialect participants were recruited from the University of Memphis. The 10 western dialect participants were recruited from the University of Nevada at Reno. The 14 northern dialect participants were recruited from the University New York at Oswego. Across all three Universities, 34 of the 38 participants were under the age of the 30. The participants were recorded with a head-mounted microphone in a quiet university office. The participants were recorded while reading a passage and word lists. The vowel measurements were then made in Praat.
Results: The researchers found that when using F1 and F2, the low back merger of the vowels /ɑ/ and /ɔ/ was mainly a characteristic found in Western dialects. They also found that Western speakers had a larger duration difference for the vowels /ɑ/ and /ɔ/ than the other

regions. For mid front vowels, the /ɛ/ was significantly longer for southern speakers. The researchers were able to find duration differences between all three dialects.

Conclusion: The overall results showed that there were vowel shifts for all three different dialect regions. The southern dialect was more distinct in both formant position and vowel duration than the other regions.

Relevance to the current work: In the current work, this study was used to describe how dialect can alter vowel formants.

Kleinow, J., & Smith, A. (2006). Potential interactions among linguistic, autonomic, and motor factors in speech. *Developmental Psychobiology, 48*, 275-287. doi:10.1002/dev.2 0141

Objective: The objective of this study was to observe how changes in the neural system affect speech. The researchers specifically wanted to analyze how autonomic arousal affected linguistic processing and speech coordination.

Method: This study had 20 participants in two different age groups. Ten of the participants were undergraduate age. The second group of 10 was composed of 9 and 10 year old children. All of the participants were native American English speakers who performed within the normal range on standardized language tests. Each participant read sentences aloud that were shown on a computer screen. For the first condition, they read the normal sentence. For the second condition, they read the sentence and then completed the Stroop Color Word Task. An OptoTrak camera system was used to track the movement of the upper lip, lower lip, and jaw while the participants completed the speaking tasks. The lip aperture variability index was calculated for the high- and low- arousal conditions.

Results: The results showed that the child group had higher lip aperture variability while completing all of the speech tasks. Increases in sentence complexity and utterance length led to more variability in speech motor coordination in both groups of participants. The Stroop task resulted in increased autonomic arousal and variability in speech motor coordination in both groups.

Conclusion: The paper explained that higher levels of autonomic arousal can affect speech motor control for both adults and children. It also concluded that sentence complexity has an effect on speech motor variability.

Relevance to the current work: This paper described how the OptoTrak system is able to measure the kinematics of speech. The current work gave a history of different systems that are used to measure speech movement. This paper contained specific details about the OptoTrak system that was included in the current work.

Neel, A. T. (2008). Vowel space characteristics and vowel identification accuracy. *Journal of Speech, Lanuage, and Hearing Research, 51*, 574-585. doi:10.1044/1092-4388(2008/041)

Objective: The purpose of this study was to analyze how characteristics of vowel production affect speech intelligibility.

Method: The study included 45 men and 48 women from the Michigan/Upper Midwest dialect group taken from the Hillenbrand study. Global and fine-grained measures were used to analyze the characteristics of each vowel. Twenty listeners then identified the vowels, providing an identification score for the researchers. The global and fine-grained measures were used to predict identification scores.

Results: The results showed that most of the identification scores for the speakers were greater than 95% correct. Both the global and fine-grained measures accounted for less than ¼ of the variance in the identification scores. The results also showed that VSA accounted for 9% -12% of the variance. This variance was mainly due to poor identification of some of the vowels. The vowels that were well defined were distinctive in their formant frequency, duration, and formant movement.

Conclusion: The author concluded that distinctiveness in neighboring vowels is more important in determining vowel intelligibility than VSA. She also concluded that an acoustic comparison of confused vowels is more useful than VSA when studying the intelligibility of normal and disordered speech.

Relevance to the current work: The present study explored the advantages and disadvantages of using VSA. This study included information about how VSA has been used to study the speech of normal speakers and those with mild or moderate dysarthria.

Nissen, S. L., Dromey, C., & Wheeler, C. (2007). First and second language tongue movements in Spanish and Korean bilingual speakers. *Phonetica, 64*, 201-216. doi: 10.1159/000121373

Objective: The objective of this study was to examine how tongue activity changed as participants spoke their native language versus speaking English.

Method: The study included 10 native Spanish speakers from Mexico and 10 native Korean speakers who spoke English as their second language with high proficiency. The participants ranged in age from 19 to 30 with an equal number of men and women for both native languages. A magnet was glued to the superior surface of the tongue of each participant in order to track tongue movement using a head-mounted JT-3 magnet tracking system. The speech acoustic signal was also recorded. The participants completed three speaking tasks. The first task was reading a prepared paragraph from a magazine article in English as well as a translation of the passage in the speakers' native language. The second task was to provide a monologue in response to a question for approximately one minute in English and then the participants native language. The last task was to give a 30-second description of a picture in English and then their native language. Stroke measures were then used to analyze the tongue movement of these different speaking tasks.

Results: The results showed a slower stroke peak and average speed when speaking English for both the native Korean and Spanish speaking participants. The participants had a longer stroke duration for English than their native language.

Conclusion: The study concluded that increased effort is required for a person to speak a second language because articulatory flexibility and automaticity have not been established. When speaking a second language, more attention is given to motor performance than when speaking a native language.

Relevance to the current work: The current work explored the advantages and limitations of using stroke metrics. This study provided information about stroke metrics and how they can be used to compare speech across languages.

Roy, N., Nissen, S. L., Dromey, C., & Sapir, S. (2009). Articulatory changes in muscle tension dysphonia: Evidence of vowel space expansion following manual circumlaryngeal

therapy. *Journal of Communication Disorders, 42*, 124-135. doi:10.1016/j.jcomdis.2008.10.001
Objective: The objective of this study was to analyze how muscle tension dysphonia (MTD) affects vowel articulation by using a variety of vowel acoustic measures. The study examined how manual circumlaryngeal treatment (MCT) affected vowel acoustic measures.
Method: The study had 111 female participants who were diagnosed with MTD by a speech language pathologist and otolaryngologist. Participants were from Canada, Wisconsin, North Dakota, Texas, and Utah. Each participant read the Rainbow Passage while the sentences were recorded and digitalized. The study used F1 and F2 of the four corner vowels to determine the vowel space area (VSA) and the vowel articulation index (VAI). Theses acoustic measures were taken before and after MCT. Five listeners then judged each participant's pre- and post- treatment MTD severity.
Results: The results showed that the severity of MTD decreased. The results also showed significant differences for VSA and VAI for pre- and post-treatment. Both VSA and VAI increased after treatment.
Conclusion: This study showed that MCT can improve articulation in people with MTD. Both acoustic and perceptual measures were improved post-treatment. These conclusions are congruent with previous studies with acoustic and perceptual measures of dysarthric speech.
Relevance to the current work: The study explained how VSA and VAI are measured and included the formulas for each measure. It also described how VSA and VAI have been used to study disordered speech. The current study included both VSA and VAI as acoustic parameters.

Sandoval, S., Berisha, V., Utianski, R. L., Liss, J. M., & Spanias, A. (2013). Automatic assessment of vowel space area. *Journal of the Acoustical Society of America, 134*, EL477-483. doi:10.1121/1.4826150
Objective: This study used connected speech, rather than single syllables, to estimate the vowel working space. The researchers compared traditional VSA to the vowel working space of connected speech.
Method: The speech samples that were used in this study were taken from the TIMIT database. Praat was used to extract F1 and F2 from the speech stimuli. The vowels that were used to determine VSA were hand-segmented. The connected speech condition used the automated formant estimation algorithm which analyzes the acoustic signal frame-by-frame and then estimates the F1 and F2 for each voiced frame. VSA and the connected speech condition were analyzed and then compared.
Results: They researchers found that connected speech yielded more accurate results than VSA. This may have been due to the fact that the connected speech condition was not limited to measuring just the corner vowels. The connected speech measure was able to analyze the F1 and F2 for many sounds throughout the utterances, resulting in a more accurate picture of the vowel working space.
Conclusion: The vowel working space for the connected speech condition is a good measure of speech intelligibility because it takes into account more acoustic information than just the four corner vowels that are used in VSA. Since the vowel working space in connected

speech more accurately displays what is occurring during speech, it is a metric that should be used in speech analysis.

Relevance to the current work: The current work compares the corner vowel metrics, VSA and VAI, to continuous speech metrics. This paper supports the idea that continuous speech metrics can give more information about what is occurring during speech than just point metric alone.

Sapir, S., Ramig, L. O., Spielman, J. L., & Fox, C. (2010). Formant centralization ratio: A proposal for a new acoustic measure of dysarthric speech. *Journal of Speech, Language, and Hearing Research, 53*, 114-125. doi: 10.1044/1092-4388(2009/08-0184)

Objective: Previous research has shown that formant centralization ration (FCR) is more accurate than VSA at distinguishing dysarthric speech from healthy speech. The objective of this study was to test FCR as an alternate measure to VSA.

Method: This study had 38 participants with idiopathic Parkinson's disease (PD) and dysarthria, and 14 healthy controls who were age- and gender-matched for the participants with PD. Half of the participants with PD received the Lee Silverman Voice Treatment while the other half did not receive treatment. All of the participants were native American English speakers from Arizona and Colorado. The participants read simple sentences multiple times before and after treatment. They were recorded using a head-mounted condenser microphone positioned 6 cm from the lips. VSA and FCR were both used to analyze each participant's speech sample.

Results: The results showed that for the pre-treatment analysis, the mean for VSA was smaller and the mean for FCR was larger in the treatment group and non-treatment group when compared to healthy control group. This pattern is consistent with vowel centralization that occurs in patients with PD. Significant differences were found between both of the PD groups and the healthy control group for FCR but not for VSA.

Conclusion: FCR was a more robust measure than VSA when comparing disordered speech to normal speech. FCR was able to robustly differentiate the dysarthric groups from the normal group when VSA was not able to detect these differences.

Relevance to the current work: The current study involved different vowel acoustic as well as kinematic measures. This study pointed out some of the weaknesses that can be found when using VSA. It also described how FCR is measured and concluded that FCR is a more robust measure to use than VSA when comparing disordered and normal speech. This information was included in the current work when comparing different types of vowel acoustic measures.

Sapir, S., Spielman, J. L., Ramig, L. O., Story, B. H., & Fox, C. (2007). Effects of intensive voice treatment (the Lee Silverman Voice Treatment [LSVT]) on vowel articulation in dysarthric individuals with idiopathic Parkinson disease: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research, 50*, 899-912. doi:10.1044/1092-4388(2007/064)

Objective: The purpose of this study was to evaluate the effects of using the Lee Silverman Voice Treatment (LSVT) on the speech of people with idiopathic PD.

Method: This study included 14 participants with PD who received the LSVT and 15 participants who did not receive the LSVT. A control group of 14 age-matched healthy individuals was also included. The speech of the three groups was recorded and vowels were extracted from the recordings. Vowel space area (VSA) was calculated from these vowels. Perceptual ratings of vowel goodness were then performed by trained raters.

Results: The results showed that for the treatment group, SPL and the ratio for F2i/ F2u greatly increased after completing the LSVT. The vowel triangle area also increased in the treatment group when it did not increase in the non-treatment group and the control group. Perceptual ratings of vowel goodness significantly improved for the participants who received the LSVT but not for the non-treatment and the control group.

Conclusion: This study concluded that vocal and articulatory functions improved with the use of the LSVT. Specifically, articulatory working space and vowel production improved with the LSVT by becoming more similar to normal speakers. This study also concluded that LSVT can lead to greater respiratory and laryngeal function in people with PD.

Relevance to the current work: The current study used VSA as an acoustic measure. This study provided information about VSA and how it has been used to study disordered speech.

Schulman, R. (1989). Articulatory dynamics of loud and normal speech. *Journal of the Acoustical Society of America, 85*, 295-312. doi:10.1121/1.397737

Objective: The objective of this study was to compare the production of normal and loud bilabial stops and stressed vowels. The researchers analyzed lip movement as well as acoustic features of the audio recording.

Method: The participants in this study were native speakers of Swedish between the ages of 22 and 30. A magnetometer system was used to record lip and jaw movement. An electroglottograph was used to record the opening and closing of the vocal folds. Six lists of words were read at a normal volume and then read again while shouting. The researchers then analyzed how displacement, velocity, and coarticulatory interactions changed with speech intensity.

Results: The results showed that articulator movements became larger and more exaggerated with louder speech. They found that these increases in movement changed linearly from the normal intensity to the loud intensity. Lip rounding and spreading also increased with loudness.

Conclusion: Loud speech requires more vocal and articulatory effort. As subglottic pressure increases, so does the articulatory working space during loud speech.

Relevance to the current work: The current study analyzed speech across the intensity scale including soft, normal, and loud speech. This study showed evidence that articulation patterns do change for louder speech by increasing the articulatory working space. The current study compared acoustic and kinematic measures to determine if speech patterns would vary with changes in speech intensity.

Skodda, S., Visser, W., & Schlegel, U. (2011). Vowel articulation in Parkinson's disease. *Journal of Voice, 25*, 467-472. doi:10.1016/j.jvoice.2010.01.009

Objective: The objective of this study was to compare vowel articulation in speakers with Parkinson's disease (PD) to normal speakers. The study also explored correlations between vowel articulation, stage of the disease, and global motor performance.

Method: This study included native German speakers. Sixty-eight of the participants had PD and 30 were healthy controls. Each participant read 4 complex sentences. The recordings were made in a quiet room using commercial audio software. F1 and F2 were used to determine the triangular vowel space area (tVSA) and vowel articulation index (VAI).

Results: The results showed that in the female group with PD, VAI was significantly smaller when compared to the control group. For the male PD group, both tVSA and VAI were significantly reduced compared to the control group. It was also found that tVSA, VAI, mean fundamental frequency, and fundamental frequency standard deviation did not show a correlation with disease duration.

Conclusion: This study supports the research that speakers with PD display articulatory undershoot. The male speakers with PD had reduced tVSA but the female speakers with PD did not. The study also supports that VAI may be a more sensitive measure than VSA at differentiating speakers with PD and normal speakers.

Relevance to the current work: The current study used both tVSA and VAI as acoustic measures. This study described the some advantages and disadvantages of using VSA and VAI. The characteristics of both tVSA and VAI were important to consider when they were chosen for the current study.

Stevens, K. N., & Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics, 38*, 10-19. doi:10.1016/j.wocn.2008.10.004

Objective: The purpose of this paper was to explain theories of speech production.

Method: The first theory that the paper discusses is the quantal theory, which explains that for a given speech gesture, there is a discontinuous relationship between the displacement of an articulator and the corresponding acoustical attribute that results from this articulatory movement. The second theory that the paper examined was the enhancement theory, which explains that articulatory gestures enhance the perceptual saliency of the defining attributes of a speech sound. The third theory that this paper described was overlap. The overlap theory is the idea that in connected speech, the positioning of the articulators for a speech sound affects the positioning the articulators for the neighboring sounds in the word or utterances.

Results: This paper did not have a results section.

Conclusion: This paper concludes by noting that acoustic and articulatory attributes of a speech sound can be influenced by surrounding speech sounds during connected speech.

Relevance to the current work: This paper discusses the quantal theory which explains that phonemes can have varying acoustic and articulatory features, but they are still perceived as one phoneme. This current work included information about the quantal theory and used it as a reason to explore the acoustic and kinematic features of speech

Objective: Speech is made up of many movements. Acoustic waveform measures of speech assume that speech involves successive approximations of each target sound that is sustained for a time and then the articulators move to make the next sound. These acoustic measures do not account for the many physical events that occur during speech, such as muscle contraction and movement of the articulators which occurs throughout the transitions from one sound to the next. Speech movement waveform measures display the complexity of speech production and the continuity found in speech. This study extracted stroke metrics from the movement record of the articulators, which were used to the show the different aspect of movement that are involved in connected speech. The main purpose of the study was to use stroke metrics to measure speech and compare those measures with acoustic waveform measures.

Method: Speech recordings were taken from the XRMB-SPD database; a publicly available database of speech recordings of typical English speakers. The recordings included the synchronous mid-sagittal plane motions of 8 articulator fleshpoints while the speakers performed speech and non-speech oral activities. This study used 18 recordings from the XRMB-SPD database where the speakers read the Hunter script. This stimulus was selected because evidence shows that the phonemic distribution of the text approximates the distribution of phonemes in conversational speech. Data were collected using the University of Wisconsin x-ray microbeam system. The movement of the articulators was measured from the mid-sagittal positions of small gold pellets (2-3mm in diameter) which were glued inside and around the mouth. Measurements from these gold pellets were used to calculate the stroke metrics. The following stroke metrics were measured: stroke distance (sum of the segment lengths between samples), stroke duration (duration between successive speed minima), peak stroke speed (maximum speed generated within the duration of the stroke), and boundary speed (the speed of onset and offset of the stroke).

Results: The study found that there were many small, short directional changes in the speed history. They found that the acoustic segments and the stroke metrics did not overlap in a systematic way. They also found that strokes that are short in duration and cover small distances are representative of small, short direction changes. These short strokes were particularly common around fricative segments. Ultimately, this study found that it is difficult to predict the duration of acoustic sound segments from aligned kinematic measures.

Conclusion: This study used both kinematic and acoustic measures to study speech. The researchers found that it was hard to predict kinematics from acoustic measures.

Relevance to the current work: This paper was relevant to the current work because it analyzed the correspondences between acoustic and kinematic measures of speech. The paper found that acoustic measures are not always a good reflection of what is occurring kinematically.

Tasko, S. M., & Westbury, J. R. (2002). Defining and measuring speech movement events. Journal of Speech, Language, and Hearing Research, *45*, 127-142. doi:10.1044/1092-4388(2002/010

Objective: Speech is made up of many movements. Acoustic waveform measures of speech assume that speech involves successive approximations of each target sound that is sustained for a time and then the articulators move to make the next sound. These acoustic measures do not account for the many physical events that occur during speech, such as muscle contraction and movement of the articulators which occurs throughout the transitions from one sound to the next. Speech movement waveform measures display the complexity of speech production and the continuity found in speech. This study extracted stroke metrics from the movement record of the articulators, which were used to the show the different aspect of movement that are involved in connected speech. The main purpose of the study was to use stroke metrics to measure speech and compare those measures with acoustic waveform measures.

Method: Speech recordings were taken from the XRMB-SPD database; a publicly available database of speech recordings of typical English speakers. The recordings included the synchronous mid-sagittal plane motions of 8 articulator fleshpoints while the speakers performed speech and non-speech oral activities. This study used 18 recordings from the

XRMB-SPD database where the speakers read the Hunter script. This stimulus was selected because evidence shows that the phonemic distribution of the text approximates the distribution of phonemes in conversational speech. Data were collected using the University of Wisconsin x-ray microbeam system. The movement of the articulators was measured from the mid-sagittal positions of small gold pellets (2-3mm in diameter) which were glued inside and around the mouth. Measurements from these gold pellets were used to calculate the stroke metrics. The following stroke metrics were measured: stroke distance (sum of the segment lengths between samples), stroke duration (duration between successive speed minima), peak stroke speed (maximum speed generated within the duration of the stroke), and boundary speed (the speed of onset and offset of the stroke).

Results: The study found that there were many small, short directional changes in the speed history. They found that the acoustic segments and the stroke metrics did not overlap in a systematic way. They also found that strokes that are short in duration and cover small distances are representative of small, short direction changes. These short strokes were particularly common around fricative segments. Ultimately, this study found that it is difficult to predict the duration of acoustic sound segments from aligned kinematic measures.

Conclusion: This study used both kinematic and acoustic measures to study speech. The researchers found that it was hard to predict kinematics from acoustic measures.

Relevance to the current work: This paper was relevant to the current work because it analyzed the correspondences between acoustic and kinematic measures of speech. The paper found that acoustic measures are not always a good reflection of what is occurring kinematically.

Tjaden, K., & Wilding, G. E. (2004). Rate and loudness manipulations in dysarthria: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research, 47*, 766-783. doi:10.1044/1092-4388(2004/058)

Objective: The objective of this study was to observe how articulatory rate reduction and increased loudness affect the speech of people with dysarthria.

Method: This study included 15 participants with dysarthria caused by multiple sclerosis, 12 participants with dysarthria caused by Parkinson's disease, and 15 healthy controls. The participants were asked to read a passage while using their habitual speech volume, loud volume, and a slower rate. The researchers examined how articulatory rate, sound pressure level, vowel space area, first moment difference, and F2 trajectory characteristics of diphthongs were affected in the three different conditions. The speech recordings were then scaled for intelligibility by 10 listeners.

Results: The results showed that for the slow condition, all of the speakers reduced articulatory rate. The speakers increased vocal intensity for the loud condition when compared to the habitual condition. The vowel acoustic distinctiveness was maximized in the slow condition. Stop consonant distinctiveness was maximized in the loud condition. The slope of F2 was not consistently affected by rate or loudness. Intelligibility increased with the loud condition when compared to the slow and habitual conditions.

Conclusion: Changes in both rate and loudness of speech affect the vocal tract acoustics. These changes alter the intelligibility of speech. Loud speech was more intelligible in people with dysarthria when compared to slow and habitual intensity speech.

Relevance to the current work: The current work observed speech across the intensity scale. This paper supported the idea that speech changes with intensity. This evidence provided a rationale for including intensity as an independent variable for the current study.

Whitfield, J. A., & Goberman, A. M. (2014). Articulatory-acoustic vowel space: Application to clear speech in individuals with Parkinson's disease. *Journal of Communication Disorders, 51*, 19-28. doi:10.1016/j.jcomdis.2014.06.005

Objective: The objective of this study was to use a continuous speech acoustic metric, articulatory-acoustic vowel space (AAVS), to evaluate speech between speakers with Parkinson's disease (PD) and a control group, and between habitual and clear speech conditions.

Method: This study included 12 participants with PD and 10 healthy controls. The participants were asked to read the first paragraph of the Rainbow Passage as well as other stimuli. The participants with PD were then asked to read the stimuli while speaking clearly. The healthy control group was not asked to speak clearly. The participants' speech samples were collected in a quiet room with a digital audio recorder. These samples were then analyzed with the AAVS software application. Four listeners also completed an auditory perceptual rating task for each recording.

Results: The results showed that participants with PD did not have speech that was as clear as the healthy controls. The AAVS measure was also lower for the participants with PD, meaning that they had a smaller articulatory working space than the healthy controls. AAVS was able to detect acoustic differences between habitual speech and clear speech. The auditory perceptual ratings were consistent with the AAVS measure results.

Conclusion: This study showed that AAVS is a sensitive acoustic measure that can be used to find differences in speech between typical speech and disordered speech. It can also detect differences in speech under different conditions within speakers. It was also concluded that speakers with PD can increase their articulatory range when asked to speak clearly.

Relevance to the current work: This study supports the use of AAVS as a sensitive measure that is able to track differences between speech conditions. The current study used AAVS as one of the acoustic measures because of its sensitivity to speech changes.

Appendix B

Consent Form

Consent to be a Participant

Introduction

This research study is being conducted by Christopher Dromey, a professor in the department of Communication Disorders at Brigham Young University to determine how movements of the tongue and lips change under several conditions (voicing, whispering, silently mouthing the words). You were invited to participate because you are a native speaker of English and have no history of speech, language, or hearing disorders.

Procedures

If you agree to participate in this research study, the following will occur:

- you will be seated in a sound-treated recording booth in room 106 of the John Taylor Building
- six small sensor coils will be attached with dental adhesive to your tongue, teeth, and lips and one to the frame of eyeglasses (no corrective lenses) that you will wear
- while you speak, the researchers will record the movements of these articulators and audio record your speech
- you will read sentences from a sheet in front of you under several conditions: normal speech, whispering, and silent mouthing of the words
- the total time commitment will be less than 60 minutes

Risks/Discomforts

You may feel uncomfortable having the sensors attached with dental glue inside your mouth. These may cause you to mildly lisp on some sounds at first. For several hours after the study you may be able to feel a slight residue on your tongue, which will disappear within a day. This technology has been widely used at other research centers and no problems for the research subjects have been reported.

Benefits

There will be no direct benefits to you. It is hoped, however, that through your participation researchers may learn about the way speech articulator movements may change under different voicing conditions. This may expand our understanding of the way the brain controls speech movements in healthy individuals and could lead to further work that would help people with speech disorders.

Confidentiality

The research data will be kept in a locked laboratory on a password protected computer and only the researcher will have access to the data. At the conclusion of the study, all identifying information will be removed and the data will be kept in the researcher's locked office. Arbitrary participant codes, but no names, will be used on the computer files or paper records for this project in order to maintain confidentiality. In presentations at conferences and in publications based on this work, only group data will be reported.

<u>Compensation</u>
You will receive $10 cash for your participation; compensation will not be prorated. For BYU students, no extra credit is available.

<u>Participation</u>
Participation in this research study is voluntary. You have the right to withdraw at any time or refuse to participate entirely without jeopardy to your class status, grade, or standing with the university.

<u>Questions about the Research</u>
If you have questions regarding this study, you may contact Christopher Dromey at (801) 422-6461 or dromey@byu.edu for further information.

<u>Questions about Your Rights as Research Participants</u>
If you have questions regarding your rights as a research participant contact IRB Administrator at (801) 422-1461;

A-285 ASB, Brigham Young University, Provo, UT 84602; irb@byu.edu.

Statement of Consent
I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study.

Name (Printed):____________________ Signature____________________ Date___