2007

Lip Kinematics in Spasmodic Dysphonia before and after Treatment with Botulinum Toxin

Christopher Dromey
Brigham Young University, dromey@byu.edu

Aubrey Reese
Brigham Young University

Susan Howey
Toronto Rehabilitation Institute

Original Publication Citation

Dromey, C., Reese, A., & Howey, S. (2007). Lip kinematics in spasmodic dysphonia before and after treatment with botulinum toxin. Journal of Medical Speech-Language Pathology, 15, 263-277.

Lip kinematics in spasmodic dysphonia before and after treatment with botulinum toxin

Christopher Dromey, Ph.D.
Aubrey Reese, M.S.
Brigham Young University

Susan Howey, M.S.
Toronto Rehabilitation Institute

Contact Author
Christopher Dromey, Ph.D.
Brigham Young University
Department of Communication Disorders
133 TLRB
Provo, UT 84602
801-422-6461
dromey@byu.edu

ABSTRACT

In order to learn about articulatory behavior in adductor spasmodic dysphonia (SD), the lip movements of seven individuals with SD and three control speakers were recorded with strain gauges as they repeated ‘Buy Bobby a puppy’ in voiced and whispered conditions. Results included weaker bilabial coordination, qualitative differences in the shape of lip movement profiles, and an increased number of lip velocity peaks in the speakers with SD. These participants received percutaneous botulinum toxin injection to the thyroarytenoid muscle, and were then recorded again. The pre- vs. post-treatment data revealed that improved vocal activity was associated with fewer articulatory disturbances. These findings lend support to the hypothesis that articulatory function in SD may at least in part be disturbed because of the coupling between the larynx and the lips during speech.

INTRODUCTION

Spasmodic dysphonia (SD) is characterized by dystonic spasms of laryngeal muscles during speech, which result in pitch and voice breaks, stutter-like blocks, and strained-strangled phonation in the more common adductor variety of the disorder (Sapienza, Murry, & Brown, 1998; Zwirner, Murry, & Woodson, 1993). In most cases, the frequency of spasms varies with the speech context. They are much less prevalent in utterances containing mostly voiceless consonants, and also tend to subside during whisper, falsetto phonation, laughter, and coughing. In utterances with many voiced sounds, vocal spasms occur more frequently (Ludlow & Connor, 1987). Injection of botulinum toxin (BT) into the thyroarytenoid (TA) muscle is currently the primary treatment for SD. BT injections block acetylcholine release at the neuromuscular junction, which results in a temporary chemical denervation and consequently, a reduction in the frequency and severity of SD spasms (Fisher, Scherer, Guo, & Owen, 1996).

Although articulatory deficits are not part of the clinical description of SD (Aronson, Brown, Litin, & Pearson, 1968), and it is viewed primarily as a laryngeal phenomenon, some researchers have examined speech characteristics other than voice quality. Cannito et al. (1997) employed visual analog scales to have listeners rate the perceived fluency of the speech of individuals with SD. They rated them as being substantially less fluent than a control group. These speakers with SD also produced significantly prolonged interword intervals and articulation times, which resulted in reduced rates of speech. The disfluencies most often associated with SD were tense pauses, part-word repetitions, and prolonged acoustic segments. Problems with voice quality (i.e., intermittent laryngeal muscular tension and struggle) typical of SD may lead to this perception of disfluency (Aronson et al., 1968). More recent studies that examined the speech of individuals with SD after BT injection have found improved intelligibility (Bender, Cannito, Murry, & Woodson, 2004) and fluency (Cannito, Woodson, Murry, & Bender, 2004), suggesting that more than just phonation can improve as a result of treatment. These findings are consistent with those of another study that documented changes to the configuration of the vocal tract following BT injection to the true vocal folds of speakers with SD (Crary, Kotzur, Gauger, Gorham, & Burton, 1996). However, different results were found in a study of phonatory and vocal tract steadiness (Zwirner, Murry, & Woodson, 1997). These authors found that BT injection to the vocal folds resulted in improved phonation, but that formant frequency measures showed no improvements in the steadiness of the supraglottal vocal tract. One study found that diadochokinetic performance in speakers with SD did not improve for whispering (Cannito, Ege, Ahmed, & Wagner, 1994), where the absence of spasms would be predicted to result in more normal articulatory function. Another study that entirely circumvented the vocal folds by having speakers with SD use an electrolarynx (Cannito, 1989) found increased second formant variability compared to controls, suggesting a problem with vocal tract steadiness.

Several possible explanations exist for disordered articulatory function in SD. One of these would be that there is a coexisting impairment of the articulatory subsystem, in which dystonic activity of the articulatory muscles affects the movements of the tongue, lips, and jaw. Previous studies (Cannito, 1989; Cannito & Kondraske, 1990) have reported an increased incidence of abnormal extralaryngeal motor function in speakers with SD, and Schaefer et al. (1992) documented electromyographic abnormalities in the tongue and velum, as well as the larynx. Such findings as tremor and focal dystonia, along with involuntary lingual, velar, neck and hand movements during speech, are suggestive of pyramidal and extrapyramidal motor system involvement (Cannito, 1989; Cannito & Kondraske, 1990). Brain imaging studies have supported the notion of supranuclear impairment in SD. Specifically, brain electrical activity mapping, single photon-emission tomography, and nuclear magnetic resonance imaging have revealed abnormal electrical activity in the cortex, reduced regional cerebral blood flow and/or scattered structural lesions in the deep cerebral nuclei and surrounding white matter (Cannito & Kondraske, 1990). Examination of supralaryngeal structures has suggested that SD affects the behavior of the hypopharynx, tongue and soft palate in some individuals with SD. Cannito (1989) stated, "This would suggest the possibility of a secondary articulatory abnormality in SD, masked perhaps by the more salient presenting voice disorder" (p. 244).

Another potential explanation for abnormal articulation in SD may be found in a consideration of laryngeal-articulatory coupling. In any speaker, the action of the larynx must be closely coordinated with the movements of the articulators in order to accurately produce sequences of voiced and voiceless sounds (Gracco & Lofqvist, 1994). Evidence from studies of both normal and perturbed speech suggests that coordination of the larynx and oral articulators is consistent with the coordinative structures model of speech motor control. This model proposes that groups of muscles and articulators act synergistically to achieve phonetic goals (Munhall, Lofqvist, & Kelso, 1994). Disturbances in the movement of one structure are accompanied by changes in the movements of other structures involved in producing the same target (Gracco & Abbs, 1988). Therefore, given that the actions of the larynx and the articulators are closely coordinated during speech production, a disruption of laryngeal movement patterning caused by SD spasms could result in a change in articulatory movements. In other words, if the larynx 'stumbles,' its tight coupling with the supralaryngeal articulators may cause them to hesitate or move abnormally. Indeed, in describing the speech of individuals with SD, Aronson et al. (1968) noted “At times articulation ceased completely. Often, if a spasm occurred just prior to or simultaneously with initiation of a word, the effect was to force the patient to repeat or prolong the phoneme or syllable until it bore some resemblance to normal speech." (p. 210).

The first purpose of the present investigation was to examine the articulatory kinematic profiles of individuals with SD and control speakers to determine whether the spatiotemporal patterning of articulation was different between these groups. It was hypothesized that the presence of laryngeal spasms in SD would impact the smoothness of lip and jaw movements. It was anticipated that the differences between disordered and control speakers would be manifest most clearly in voiced speech, where spasms have been most often reported (Ludlow & Connor, 1987). The second purpose was to examine the association between the severity of laryngeal spasms and the hypothesized articulatory kinematic changes in SD. It was reasoned that if laryngeal-articulatory coordination contributes to these articulatory disturbances, then decreased frequency and severity of laryngeal spasms should be associated with fewer articulatory disturbances. Specifically, fewer articulatory disturbances should occur following medical voice treatment and during whispered speech. In both of these conditions, fewer laryngeal spasms would be expected. If, on the other hand, the articulatory disturbances in speakers with SD are due to a concomitant motor disorder affecting the supralaryngeal articulators, no significant improvements would be predicted to accompany treatment-related changes in laryngeal function.

METHOD

Participants

Four women (aged 35 to 61 years) and three men (aged 40 to 66 years) with a diagnosis of adductor SD, participated in this study. All were native speakers of English. Additionally, two women (57 and 60 years) and one man (63 years) participated as control speakers. Two of the participants with SD and one of the control speakers were recorded in Toronto, Ontario, and the remaining speakers were recorded with an identical experimental setup in Provo, Utah. Participants gave written consent to take part in this study, which was approved by the Institutional Review Boards at the University Health Network and at Brigham Young University. The participants had no history of language or hearing disorders, and the control participants had no history of speech disorders. All participants were found to have normal hearing, based on an audiometric screening carried out at the time of the experiment. Normal hearing was defined as pure-tone air-conduction thresholds of 25 dB HL 0.5, 1, 2 and 4 kHz bilaterally with earphones. All SD patients were diagnosed by consensus of a speech-language pathologist and an otolaryngologist at either the University Health Network in Toronto, or the University of Utah Medical Center. These voice team members had considerable experience in the assessment and treatment of neurolaryngological disorders. They used a combination of case history information, in-depth speech and voice evaluations, perceptual judgments, and fiberoptic examination of the larynx to diagnose SD. All participants were receiving regular EMG guided BT injections as part of their ongoing treatment for SD at the time of this study. All speakers with SD had a history of successful response to BT injection based on self-report and the perceptual judgment of clinical personnel.

Instrumentation

Articulatory kinematic recording

Inferior-superior movements of the upper lip (UL), lower lip (LL), and jaw (J) were tracked with a head-mounted strain-gauge cantilever system (Barlow, Cole, & Abbs, 1983). The wires of the cantilever beams were attached at the midline of each lip, at the vermilion border with double-sided tape. The cantilever for the jaw was attached under the chin. The transducers were visually judged to be perpendicular to the occlusal plane. The kinematic signals were low-pass filtered at 50 Hz and digitized at a rate of 1 kHz with 16-bit quantization using WINDAQ DI-720 hardware/software data acquisition system. The strain gauges were calibrated by deflecting the beams with a linear micrometer, and measuring the output voltage. Digitized signals were subsequently displayed as a displacement in millimeters (mm).

Audio recording

A condenser microphone (AKG C-420) was placed on the headset of the strain gauge system at a constant distance of 8 cm from the participants' lips. The microphone signal was digitized on-line using the WINDAQ system at a rate of 25 kHz, after low-pass filtering at 12 kHz (Frequency Devices, 9001).

Aerodynamic recording

Aerodynamic data were collected to assess changes in laryngeal function following BT injection. Oral airflow and intraoral pressure data were collected using a flow mask and pressure transducer system (Glottal Enterprises MS100-A2). The pressure transducer was attached to a short piece of tubing that extended into the mouth and rested above the tongue. The mean airflow and intraoral pressure signals were digitized at 10 kHz with WINDAQ. The aerodynamic transducers were calibrated for each recording session using the Glottal Enterprises (MCU-4) calibration system.

Procedures

The speakers with SD attended two recording sessions. The first occurred within the week prior to a BT injection, and the second 4 to 6 weeks following the injection. Because the participants had been injected previously, they were familiar with the typical response following injection, and had requested re-injection after their SD symptoms had become sufficiently severe to require treatment. It cannot be guaranteed that the residual effects of the previous injection had fully disappeared. However, because the goal was to measure speech performance that would be reflective of the speakers' typical experience, the decision was left to the participants as to when they would request an injection. The time interval between recordings was selected because at 4 to 6 weeks post-injection, patients are typically experiencing both a decrease in vocal spasms and a reduction in the vocal side-effects of BT, such as breathiness and hoarseness (Fisher et al., 1996). The control participants were recorded once.

An otolaryngologist performed all BT injections. It has been found that individual responses to injections vary; although some patients have success with bilateral injections, others react better with a unilateral injection (Miller & Woodson, 1991). Therefore, injections were made either unilaterally, or bilaterally, based on the patient's injection history and personal preference.

Speech tasks

Data were collected from participants in a single-walled sound-isolated booth (Acoustic Systems). They completed the following speech tasks:

1. Each participant produced the sentence "Buy Bobby a puppy" fifteen times at normal rate and loudness. A beep was presented every three seconds to pace the production of the tokens. Participants were asked to take a breath and repeat the utterance once each time they heard a beep.

2. Participants were asked to whisper the sentence "Buy Bobby a puppy" fifteen times. The same pacing beep procedure was used as for voiced speech.

Aerodynamic tasks

Aerodynamic data were obtained from the participants after the strain-gauge apparatus was removed. They were seated and held the flow mask over the mouth and nose. In order to minimize air leakage from around the mask, each participant was instructed to hold the mask tightly against the face. Participants produced the syllable /pɑ/ five times on one continuous breath while using constant vocal effort, after it was modeled by the researcher at a rate of approximately 90 syllables per minute. This task was performed a total of three times.

Articulatory Kinematic Analysis

MATLAB (6.1) software was used to carry out all kinematic analyses. The lip displacement recordings were digitally low-pass filtered at 10Hz and differentiated to produce velocity waveforms. The LL signal reflected the combined motion of the LL and the jaw; no subtraction was performed to separate out the jaw's contribution to the movement of the LL. From the 15 recorded utterances, the usual procedure was to make measurements from tokens 6-15, so that the first 5 would be considered practice tokens. In cases where the speaker said the wrong word or coughed during one of the selected utterances, a token from the first 5 was substituted. The mean of ten tokens was used to represent performance on each variable. The following analyses were used to evaluate articulatory kinematics in speakers with SD and control participants.

Duration

In order to determine whether changes in duration may have impacted kinematic profiles in addition to any effects associated with SD, the duration of articulatory movement for the LL was measured for a target segment of the utterance. This segment was defined from the peak velocity of the first opening movement of the /b/ to /aI/ in the word "buy" to the peak velocity of the last closing movement of the /℘/ to /p/ in the word "puppy" (see vertical lines in Figure 1).

Insert Figure 1 about here

Point measures

To evaluate whether the amplitude and velocity of movements differed between groups and across conditions, point measures for the LL displacement movement were computed. The displacement (mm) and peak velocity (mm/s) were calculated for the movement from the /ɑ/ to the medial /b/ in "Bobby" (Figure 1). This gesture was chosen for analysis because it was defined by reliably identifiable kinematic landmarks and it occurred near the middle of the utterance. Point measures are by definition limited (Smith, Goffman, Zelaznik, Ying, & McGillem, 1995), but allow at least a simple index of the scaling of articulatory movements across speaking conditions.

UL/LL correlation

A continuous labial correlation was obtained by computing a Pearson correlation coefficient for a sliding 5 sample window for the UL and LL displacement signals throughout the target utterance. Lip movements of bilabial consonant production are normally approximately 180° out-of-phase; therefore, it was reasoned that a strong negative correlation would reflect normal patterning of lip movements (Tingley & Dromey, 2000). Weaker correlation values would be predicted for less coordinated labial movements, as might occur in cases of disordered articulation.

Spatiotemporal Index (STI)

The 10 LL displacement waveforms were segmented (as defined in (1) above) and then amplitude and time normalized. The records were amplitude normalized by subtracting the mean and dividing by the standard deviation. Time normalization was achieved by Fourier analysis and resynthesis to linearly adjust all 10 records to the same length (Smith, Johnson, McGillem, & Goffman, 2000). Finally, standard deviations of normalized displacement were computed at 2% intervals throughout the records and summed, resulting in the STI (Smith et al., 1995). A higher STI value is reflective of greater variability in the kinematic signal across repetitions. For the present study, the STI was used to evaluate whether the spatiotemporal variability of articulation would change following BT injection, and in whispered vs. voiced speech.

Count of velocity peaks

To evaluate the smoothness of velocity profiles, the number of LL velocity peaks during the selected segment was counted and averaged across 10 repetitions. Velocity waveforms of the LL were expected to result in roughly sinusoidal shape for control speakers, whereas disordered speech has been previously reported to involve multiple velocity peaks (Adams, Weismer, & Kent, 1993).

Vocal Function

Perturbation

The microphone signal was analyzed to provide an acoustic measure of dysphonia severity. TF32 software (Milenkovic, 2005) was used to compute the percentage of jitter and shimmer, along with the signal to noise ratio (SNR). The segment for analysis was defined from peaks in the LL velocity record from the bilabial opening from /b/ to /aI/ in 'buy' to the closure into the first /b/ in 'Bobby.'

Aerodynamic measures

Aerodynamic data were collected to measure the effect of BT injections on vocal function. Laryngeal airway resistance (Rlaw) was expected to decrease following BT injections. Estimated Rlaw was calculated by dividing peak air pressure during /p/ occlusion (an estimate of subglottal pressure) by average mid-vowel airflow in the syllable task. Custom Matlab routines computed average airflow across a 100 ms window at the midpoint of the /ɑ/ vowel, and peak oral pressure during /p/ closure.

Perceptual ratings

Six graduate speech-language pathology students, who were enrolled in a voice disorders course, completed a perceptual analysis of participant vocal symptoms in order to obtain a rating of pre- and post-treatment SD severity. At this stage of their training, they had taken courses in normal speech production and had been exposed to disordered voice recordings. However, they had not yet worked in the clinic with individuals who had voice disorders. A custom Matlab routine presented recorded samples of the target utterance to listeners and allowed them to move a pointer with a mouse along a visual analog scale. One end was marked 'normal' and the other 'severely dysphonic.' One representative utterance for each speaker from the pre- and post-BT injection recordings for the 7 speakers with SD along with 3 control recordings of the sentence “Buy Bobby a puppy” were presented in a random order, for a total of 17 listening samples. Because the 10 sentences for a speaker in any given condition sounded to the investigators to be similar in terms of overall quality and the presence of spasms, the use of a single token for the perceptual task was deemed appropriate. In order to measure intra-rater reliability, an additional 8 randomly selected tokens were repeated within the evaluation, for a total of 25 ratings. Intra-rater reliability was tested with Pearson coefficients, which ranged from .92 to .99 (mean .96).

Inter-rater agreement was evaluated using intra-class correlation coefficients, which yielded a single measure correlation of .801 and an average measures value of .960 ($F=25.177$, $p < .001$).

RESULTS

Because the number of participants was small, it was not appropriate to apply inferential statistical techniques to the data set. Instead, descriptive statistics were used to present the patient group's performance in the voiced and whispered condition before and after BT injection. These values will be compared with the performance of the control group. A summary of the data is presented in Table 1.

Insert Table 1 about here

Articulatory Kinematics

Duration

Prior to the injection, the mean duration for the speakers with SD decreased by almost 100 ms for whispering compared to the voiced condition. It decreased further for both conditions after BT injection to become similar to the value for the control speakers, who showed a minimal difference between voiced and whispered utterances. Figure 2, which shows the performance of the individual speakers in the voiced condition pre- and post-treatment, reveals that three of the speakers with SD were the main contributors to the group mean decrease in duration, and that the other four were similar to the control speakers in their pre- and post-treatment duration. Listening to the recordings revealed that for those individuals with the longer durations, the first vocalic segment of the target utterance was prolonged in its production. Inspection of the lower lip waveforms revealed a longer lowering for the vowel prior to treatment (see Figure 3).

Insert Figures 2 and 3 about here

Point measures and lip trajectory features

Displacement and velocity measures for the target bilabial closure for the speakers with SD decreased slightly as a group in the pre-treatment whispered and both post-injection conditions. However, only two of the speakers exhibited clear decreases, and their values led to this modest group change. Most of the speakers with SD showed little change across conditions on these measures. Those speakers with larger displacements in any given condition also had the larger velocities. The control speakers' performance remained unchanged for voiced or whispered speech.

Qualitative examination of the displacement records revealed subtle shape changes following treatment. Figure 3 shows irregularities in the lip movement of a speaker with SD before treatment that are not visible following the injection as the lip was lowered for the first vowel and raised for the second /b/.

UL/LL correlation

The correlation was weakest for the speakers with SD in the voiced condition before injection, reflecting reduced coordination in the movement of the two lips. Whispering increased the strength of the correlation pre-treatment. After the injection the correlation strengthened for the voiced and further still for the whispered condition. The correlation was stronger for the control speakers than for the disordered group – more so in the voiced than whispered condition. Figure 4 reveals that the speakers varied in the strength of the correlation, as well as in the degree to which they changed following treatment. Figure 5 shows movements for a participant with SD before treatment. There are times in this record when the UL and LL are moving in the same direction at the same time. The lower panels of this figure shows data for the same speaker post-treatment, revealing improved labial coordination as reflected in more consistent negative correlation values.

Insert Figures 4 and 5 about here

Spatiotemporal Index (STI)

This index was highest for the patient group before treatment in the voiced condition, and their lowest value was for voiced speech after injection, reflecting reduced variability in movement across the ten repetitions. The pre- to post-treatment STI decreased for 5 of the 7 speakers with SD – substantially in some cases. Two of the speakers showed increases on this measure following treatment. The control group value for voiced speech was similar to the patients' post-injection results; their STI for whispered speech was slightly higher than for the voiced condition, reflecting mildly increased variability.

Count of velocity peaks

The mean number of velocity peaks counted in the kinematic record was highest for the patient group in voiced speech before treatment. This value decreased for whispering, then decreased again for voiced speech after treatment. Their lowest count was for whispered speech after treatment, approaching the values for the control speakers. The scatter plot in Figure 6 which combines the values for all speakers and conditions shows that speakers with longer utterance durations tended to have higher counts of velocity peaks (R^2 = 0.7764).

Insert Figure 6 about here

Vocal Function

Perturbation

Percent jitter and percent shimmer decreased by almost 50% following BT injection for the speakers with SD, but remained much higher than the values for the control speakers. The signal to noise ratio for the patients increased after injection, but remained slightly lower than for the controls.

Aerodynamic measures

For one speaker with SD, technical problems prevented the collection of usable aerodynamic data, and the results are thus for 6 rather than 7 patients. Before BT injection, pressure was slightly higher and flow somewhat lower for the speakers with SD than for the controls. This resulted in a higher value for estimated laryngeal resistance. After injection, and substantial increase in flow led to a resistance value for the patient group that was slightly below that for the controls.

Perceptual ratings

Following treatment, the speakers with SD were rated as less than half as severe as they were before BT injection. Even following treatment, their voice severity was still higher than the control values. Figure 7 shows the association between perceptual ratings of severity and the count of velocity peaks in the voiced utterances. The speakers with the higher severity ratings tended to have a higher count of velocity peaks, although the R^2 value was only 0.3702.

Insert Figure 7 about here

DISCUSSION

The aim of this study was to learn whether speakers with SD have abnormal labial kinematics compared to individuals with normal speech. Additionally, conditions were compared where more spasms were expected (pre-treatment and voiced speech) against conditions where fewer spasms were anticipated (post-treatment and whispered speech) to observe whether changes in phonatory behavior would influence lip movement measures. Acoustic perturbation and perceptual measures were used to verify that the BT injections had a beneficial effect on the voice.

Articulatory Kinematics

Duration

The longer utterance durations for the speakers with SD than for controls are consistent with previous accounts (Cannito et al., 1997; Cannito et al., 1994). However, individual performance was variable, with three of the disordered speakers exhibiting longer durations, and four who were similar to the control speakers (see Figure 2). Thus, increased segment durations do not appear to be characteristic of all speakers with SD. For those speakers who showed a prolonged lowering of the lower lip (see Figure 3), it could be speculated that when the larynx was unable to produce sound smoothly because of a spasm, the tight coupling between the laryngeal and articulatory subsystems prevented the lips from moving ahead normally because the larynx had not finished producing the vowel.

Point measures and lip trajectory features

The inconsistent changes in lip displacement and velocity across speaking conditions do not allow straightforward inferences about the impact of phonatory stability on this aspect of articulatory function. The scaling of velocities relative to displacements is consistent with reports of normal speech production, in that larger velocities typically accompanied larger displacements (Schulman, 1989; Smith, 1992). Other reports have highlighted the limitations inherent in point measures of articulation (Lucero, Munhall, Gracco, & Ramsay, 1997; Smith et al., 1995), and it may simply be that they were insensitive to the articulatory changes associated with SD. On the other hand, the qualitative differences in the lip trajectory before and after treatment suggest that more disordered phonation was linked to less smooth lip movement (Figure 3). The pre-treatment waveform is characterized by a number of extra peaks, whereas the post-treatment movement is characterized by smoother upward and downward movements. This finding is consistent with the view that laryngeal-articulatory coupling may be at least partly responsible for abnormal labial kinematics in SD. In other words, the discontinuities in phonation could directly contribute to hesitant articulatory movements because of the coupling between the two systems.

UL/LL correlation

The speakers with SD demonstrated notably weaker negative UL/LL correlations than the control participants; thus there appeared to be a less consistent coordination of UL and LL movement in the disorder group. The speakers with SD showed a stronger negative UL/LL correlation in the whispered condition compared to the voiced condition. Following BT injection, the coordination of lip movement became stronger still, with its highest value in the whispered post-treatment condition. The fluctuating phase relations between the two lips in the pre-treatment condition for a speaker with SD are apparent in Figure 5, and appear far more consistent following BT injection. These findings suggests that improved phonatory function corresponds to more stable coordination of UL and LL movements. If the presence of articulatory disturbances in SD were solely due to a concomitant dystonia of the articulatory muscles, then no change in lip movement would be predicted from injection of the thyroarytenoid. The finding that an improvement in laryngeal function is associated with more consistent lip coordination is supportive of the hypothesis that inter-system coupling is at least in part responsible for disordered articulation in SD.

Spatiotemporal Index

The higher STI for the speakers with SD than for the controls is indicative of greater lip movement variability over multiple repetitions. Previous work has reported elevated STI values in other disorder groups, such as stutterers (Kleinow & Smith, 2000) and speakers with Parkinson disease (Kleinow, Smith, & Ramig, 2001). The UL/LL correlation data suggest that the presence of spasms affects the speaker's ability to move the lips smoothly, and the inconsistent nature of spasmodic muscle activity likely led in the present findings of more variable lip movement. The lower post-treatment STI values lend further support to the hypothesis that abnormal laryngeal activity may contribute to the articulatory disturbance in SD.

Another previously reported influence on STI is speech rate. Other authors have reported elevated STI measures for individuals with normal or disordered speech when they talk slowly (Kleinow et al., 2001; Smith et al., 1995). The three speakers who showed a decrease in utterance duration following treatment also showed clear decreases in STI, and thus a rate increase may have contributed to their articulatory consistency. On the other hand, two speakers with substantial decreases in STI did not change their rate following treatment, and thus the laryngeal-articulatory coupling explanation could be tendered as a tentative explanation for these individuals.

Count of velocity peaks

The finding of more LL velocity peaks for participants with SD may have been linked to the more variable lip movements that are reflected in the correlation and STI measures. However, it is possible that the increased count of velocity peaks resulted from slower speech prior to treatment, rather than from a direct impact of laryngeal disturbances on articulatory movements. The scatterplot in Figure 6 shows that these variables are clearly correlated. Adams et al. (1993) investigated the impact of speaking rate on velocity profiles. They found that decreases in speaking rate were associated with an increase in the number of velocity peaks. Therefore, the occurrence of more LL velocity peaks may result from reduced speech rate rather than as a direct consequence of the disorder itself. On the other hand, the SD speakers with the highest severity ratings also tended to have an increased count of LL velocity peaks. This finding supports the hypothesis that disturbances to laryngeal movement patterning may underlie the disturbances to supralaryngeal articulatory movement.

Vocal Function

The acoustic, aerodynamic and perceptual ratings of vocal function were made in order to verify that the BT injections had the intended effect of reducing spasm severity. Following treatment, the voices of the speakers with SD were better on each of these measures, but still were not as good as those of the control participants. Nevertheless, it was important to establish that vocal function had improved in these speakers, because BT injections are not always completely successful.

General Discussion

The present data suggest that differences exist between the labial kinematic behavior of some individuals with SD and control speakers. Specifically, results indicated weaker UL/LL correlations, qualitative differences in the shape of LL movement profiles, and an increased number of LL velocity peaks in the disordered speakers. The pre- vs. post-treatment data revealed that improvements in vocal function were associated with fewer articulatory disturbances. The data from the present study are consistent with reports of improvements to the intelligibility and fluency of speech following BT injection in SD (Bender et al., 2004; Cannito et al., 2004; Crary et al., 1996). They are also congruent with previous research that details the important linkages between laryngeal and articulatory activity (Gracco & Abbs, 1988; Gracco & Lofqvist, 1994; Munhall et al., 1994). They support the hypothesis that laryngeal-articulatory coordination contributes to articulatory disturbances in SD speech.

However, the present findings must be considered in the context of previous work which has reported concomitant articulatory disturbances in SD, even when speakers whispered or used an artificial larynx to excite the vocal tract (Cannito, 1989; Cannito et al., 1994). Therefore, it appears that disordered articulation in this population cannot be attributed to a single cause. Future work with a much larger sample of speakers might reveal subgroups of speakers with SD, who may have a variety of articulatory characteristics.

Reference List

Adams, S. G., Weismer, G., & Kent, R. D. (1993). Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research, 36,* 41-54.

Aronson, A. E., Brown, J. R., Litin, E. M., & Pearson, J. S. (1968). Spastic dysphonia. I. Voice, neurologic, and psychiatric aspects. *Journal of Speech and Hearing Disorders, 33,* 203-218.

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.

Bender, B. K., Cannito, M. P., Murry, T., & Woodson, G. E. (2004). Speech intelligibility in severe adductor spasmodic dysphonia. *Journal of Speech, Language, and Hearing Research, 47,* 21-32.

Cannito, M. P. (1989). Vocal tract steadiness in spasmodic dysphonia. In K. M. Yorkston & D. R. Beukelman (Eds.), *Recent advances in clinical dysarthria* (pp. 243-262). Boston: Little, Brown and Co.

Cannito, M. P., Burch, A. R., Watts, C., Rappold, P. W., Hood, S. B., & Sherrard, K. (1997). Disfluency in spasmodic dysphonia: a multivariate analysis. *Journal of Speech, Language, and Hearing Research, 40,* 627-641.

Cannito, M. P., Ege, P., Ahmed, F., & Wagner, S. (1994). Diadochokinesis for complex trisyllables in individuals with spasmodic dysphonia and nondisabled subjects. In J. A. Till, K. M. Yorkston, & D. R. Beukelman (Eds.), *Motor speech disorders: Advances in assessment and treatment* (pp. 91-100). Baltimore: P.H. Brooks Pub. Co.

Cannito, M. P. & Kondraske, G. V. (1990). Rapid manual abilities in spasmodic dysphonic and normal female subjects. *Journal of Speech and Hearing Research, 33,* 123-133.

Cannito, M. P., Woodson, G. E., Murry, T., & Bender, B. (2004). Perceptual analyses of spasmodic dysphonia before and after treatment. *Archives of Otolaryngology -- Head & Neck Surgery, 130,* 1393-1399.

Crary, M. A., Kotzur, I. M., Gauger, J., Gorham, M., & Burton, S. (1996). Dynamic magnetic resonance imaging in the study of vocal tract configuration. *Journal of Voice, 10,* 378-388.

Fisher, K. V., Scherer, R. C., Guo, C. G., & Owen, A. S. (1996). Longitudinal phonatory characteristics after botulinum toxin type A injection. *Journal of Speech and Hearing Research, 39,* 968-980.

Gracco, V. L. & Abbs, J. H. (1988). Central patterning of speech movements. *Experimental Brain Research, 71,* 515-526.

Gracco, V. L. & Lofqvist, A. (1994). Speech motor coordination and control: evidence from lip, jaw, and laryngeal movements. *Journal of Neuroscience, 14,* 6585-6597.

Kleinow, J. & Smith, A. (2000). Influences of length and syntactic complexity on the speech motor stability of the fluent speech of adults who stutter. *Journal of Speech, Language, and Hearing Research, 43,* 548-559.

Kleinow, J., Smith, A., & Ramig, L. O. (2001). Speech motor stability in IPD: Effects of rate and loudness manipulations. *Journal of Speech Language and Hearing Research, 44,* 1041-1051.

Lucero, J. C., Munhall, K. G., Gracco, V. L., & Ramsay, J. O. (1997). On the registration of time and the patterning of speech movements. *Journal of Speech, Language, and Hearing Research, 40,* 1111-1117.

Ludlow, C. L. & Connor, N. P. (1987). Dynamic aspects of phonatory control in spasmodic dysphonia. *Journal of Speech and Hearing Research, 30,* 197-206.

Milenkovic, P. (2005). TF32 [Computer software]. Madison, WI.

Miller, R. H. & Woodson, G. E. (1991). Treatment options in spasmodic dysphonia. *Otolaryngologic Clinics of North America, 24,* 1227-1237.

Munhall, K. G., Lofqvist, A., & Kelso, J. A. S. (1994). Lip-larynx coordination in speech: Effects of mechanical perturbations to the lower lip. *Journal of the Acoustical Society of America, 95,* 3605-3616.

Sapienza, C. M., Murry, T., & Brown, W. S. (1998). Variations in adductor spasmodic dysphonia: acoustic evidence. *Journal of Voice, 12,* 214-222.

Schaefer, S. D., Roark, R. M., Watson, B. C., Kondraske, G. V., Freeman, F. J., Butsch, R. W. et al. (1992). Multichannel electromyographic observations in spasmodic dysphonia patients and normal control subjects. *Annals of Otology, Rhinology and Laryngology, 101,* 67-75.

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

Smith, A. (1992). The control of orofacial movements in speech. *Critical Reviews in Oral Biology and Medicine, 3,* 233-267.

Smith, A., Goffman, L., Zelaznik, H. N., Ying, G., & McGillem, C. (1995). Spatiotemporal stability and patterning of speech movement sequences. *Experimental Brain Research, 104,* 493-501.

Smith, A., Johnson, M., McGillem, C., & Goffman, L. (2000). On the assessment of stability and patterning of speech movements. *Journal of Speech, Language, and Hearing Research, 43,* 277-286.

Tingley, S. & Dromey, C. (2000). Phonatory-articulatory relationships: Do speakers with spasmodic dysphonia show aberrant lip kinematic profiles? *Journal of Medical Speech-Language Pathology, 8,* 249-252.

Zwirner, P., Murry, T., & Woodson, G. E. (1993). Perceptual-acoustic relationships in spasmodic dysphonia. *Journal of Voice, 7,* 165-171.

Zwirner, P., Murry, T., & Woodson, G. E. (1997). Effects of botulinum toxin on vocal tract steadiness in patients with spasmodic dysphonia. *European Archives of Oto-Rhino-Laryngology, 254,* 391-395.

ACKNOWLEDGEMENT

We express our appreciation to the participants who allowed us to record their speech for this research. We are also grateful to Dr. Jonathan Irish in Toronto and Dr. Marshall Smith in Salt Lake City for allowing us access to their patients. Parts of this research were undertaken as masters theses by the second and third authors.

Table 1. Means and standard deviations for the kinematic, acoustic, aerodynamic and perceptual measures across conditions for the speakers with SD and control participants.

	Speakers with SD								Controls			
	Pre-voiced		Pre-whispered		Post-voiced		Post-whispered		Voiced		Whispered	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Duration (ms)	1225.5	308.0	1129.6	195.9	1034.7	106.5	1054.2	134.6	1001.1	316.1	1036.9	167.4
Displ. (mm)	9.5	3.2	8.3	1.0	8.6	2.6	7.7	1.9	9.2	3.8	9.4	3.3
Vel. (mm/s)	138.0	50.3	114.5	18.7	125.3	44.2	107.5	27.4	145.0	52.1	147.3	49.8
UL-LL Corr.	-0.587	0.367	-0.705	0.209	-0.724	0.167	-0.794	0.126	-0.835	0.071	-0.818	0.064
STI	13.1	5.9	12.3	5.4	10.1	3.4	12.2	3.2	10.7	0.8	11.4	2.6
Vel. Peaks	5.2	1.9	4.5	1.4	4.0	0.9	3.9	0.7	3.6	1.0	3.9	0.7
% Jitter	6.1	3.3			3.5	2.7			0.7	0.3		
% Shimmer	24.2	13.7			14.2	11.1			4.5	0.6		
SNR	9.2	4.5			14.9	4.5			16.1	3.9		
Press. (cmH2O)	7.9	2.5			7.7	2.6			6.0	0.4		
Air Flow (L/s)	0.142	0.038			0.256	0.091			0.164	0.008		
LAR (cmH20/L/s)	56.7	9.0			31.1	6.8			36.8	3.1		
Percep. Severity	69.1	13.0			27.9	17.2			13.5	11.1		

Notes
Pre- = recording before botulinum toxin injection; Post- = recording after botulinum toxin injection; Displ. = lower lip displacement; Vel. = lower lip peak velocity; Corr. = Pearson correlation; Vel. Peaks = count of velocity peaks; SNR = signal to noise ratio; Press. = estimated subglottal pressure; LAR = estimated laryngeal airway resistance; Percep. = perceptual

Figure Captions

Figure 1. The upper panel shows the LL displacement for one production of the phrase "Buy Bobby a puppy." The lower panel shows the corresponding lip velocity. The target measurement segment (for duration and STI measures) was identified from the peak velocity of the first opening movement of the /b/ to /αI / in the word "buy" to the peak velocity of the last closing movement of the /℘/ to /p/ in the word "puppy" (marked by the two vertical lines). The displacement and peak velocity point measures were taken from the movement from the /□/ to the medial /b/ in "Bobby". Vocalic events identified in the velocity trace represent peak opening velocities, rather than maximal opening points for the respective vowels.

Figure 2. Individual speaker mean utterance durations for the 10 repetitions in the pre- and post-injection voiced conditions for the speakers with SD. The horizontal gray bar is the mean duration for the control speakers.

Figure 3. Pre (two upper panels) and post-injection (two lower panels) microphone and lower lip displacement recordings for a speaker with SD showing smoother kinematic profiles after treatment. The signals represent the entire utterance, "buy Bobby a puppy."

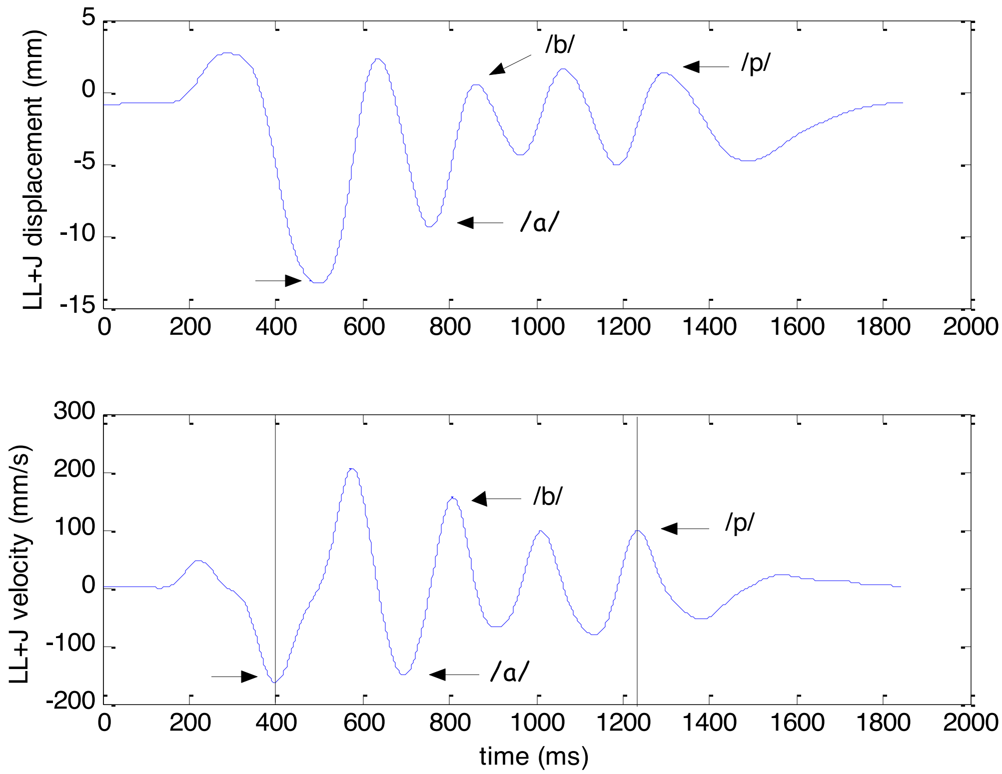
Figure 4. Individual speaker mean UL-LL correlations for the 10 repetitions in the pre- and post-injection voiced conditions for the speakers with SD. The horizontal gray bar is the mean correlation for the control speakers.

Figure 5. Upper lip (thin line) and lower lip (thick line) displacements during one repetition of the target utterance are shown in the top panel, and the continuous correlation function is in the second panel, showing poorly coordinated lip movement in a speaker with SD prior

to treatment. The lower two panels show the same data for this speaker following treatment.

Figure 6. The association between the utterance duration and the mean count of velocity peaks across all speaking conditions and speakers.

Figure 7. The association between the voice severity ratings and the mean count of velocity peaks for voiced utterance conditions across all speakers.



LL+J displacement (mm)
5
0
-5
-10
-15
/b/
/p/
/ɑ/
0
200
400
600
800
1000
1200
1400
1600
1800
2000
LL+J velocity (mm/s)
300
200
100
0
-100
-200
/b/
/p/
/ɑ/
0
200
400
600
800
1000
1200
1400
1600
1800
2000
time (ms)

Figure 1

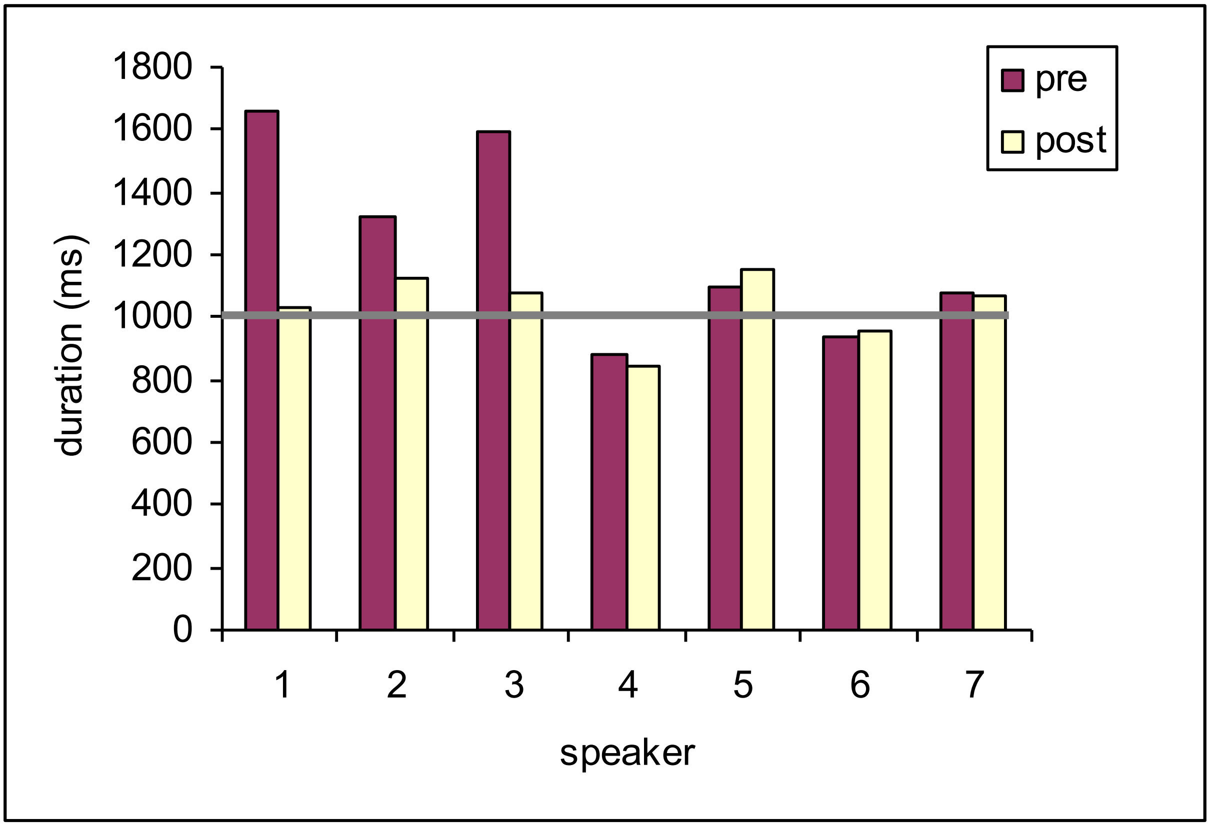

1800
1600
1400
1200
1000
800
600
400
200
0
duration (ms)
1
2
3
4
5
6
7
speaker
pre
post

Figure 2

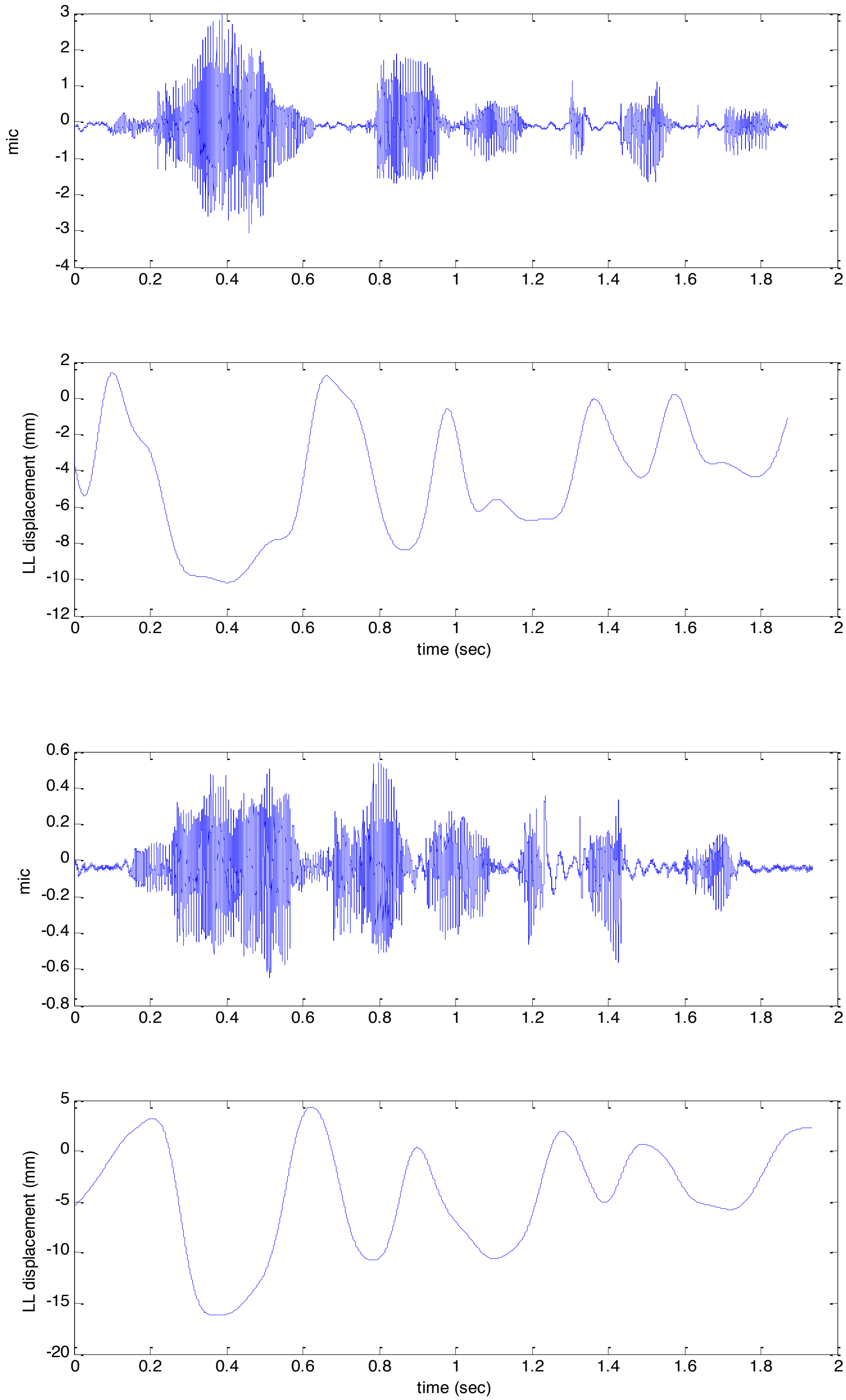
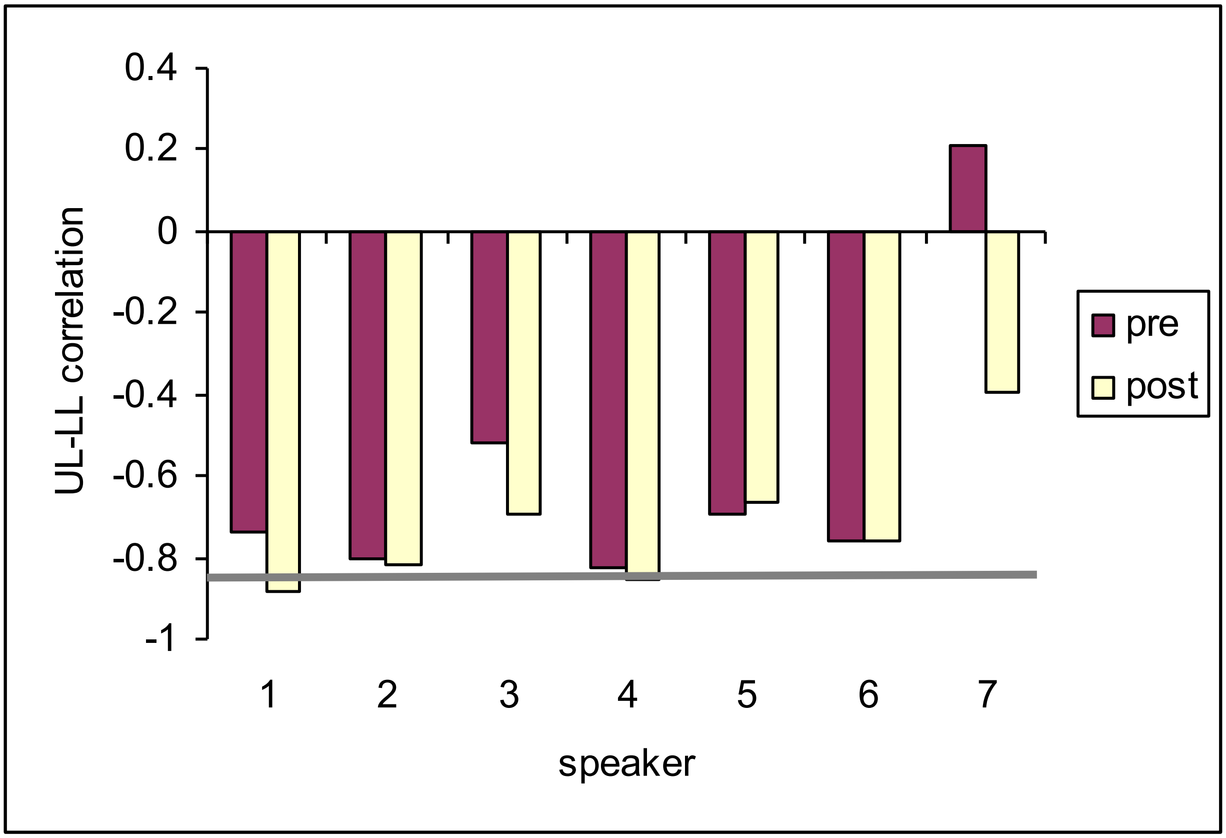

mic
3
2
1
0
-1
-2
-3
-4
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
LL displacement (mm)
2
0
-2
-4
-6
-8
-10
-12
time (sec)
0.6
0.4
0.2
-0.2
-0.4
-0.6
-0.8
5
-5
-10
-15
-20

Figure 3



0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
UL-LL correlation
1
2
3
4
5
6
7
speaker
pre
post

Figure 4

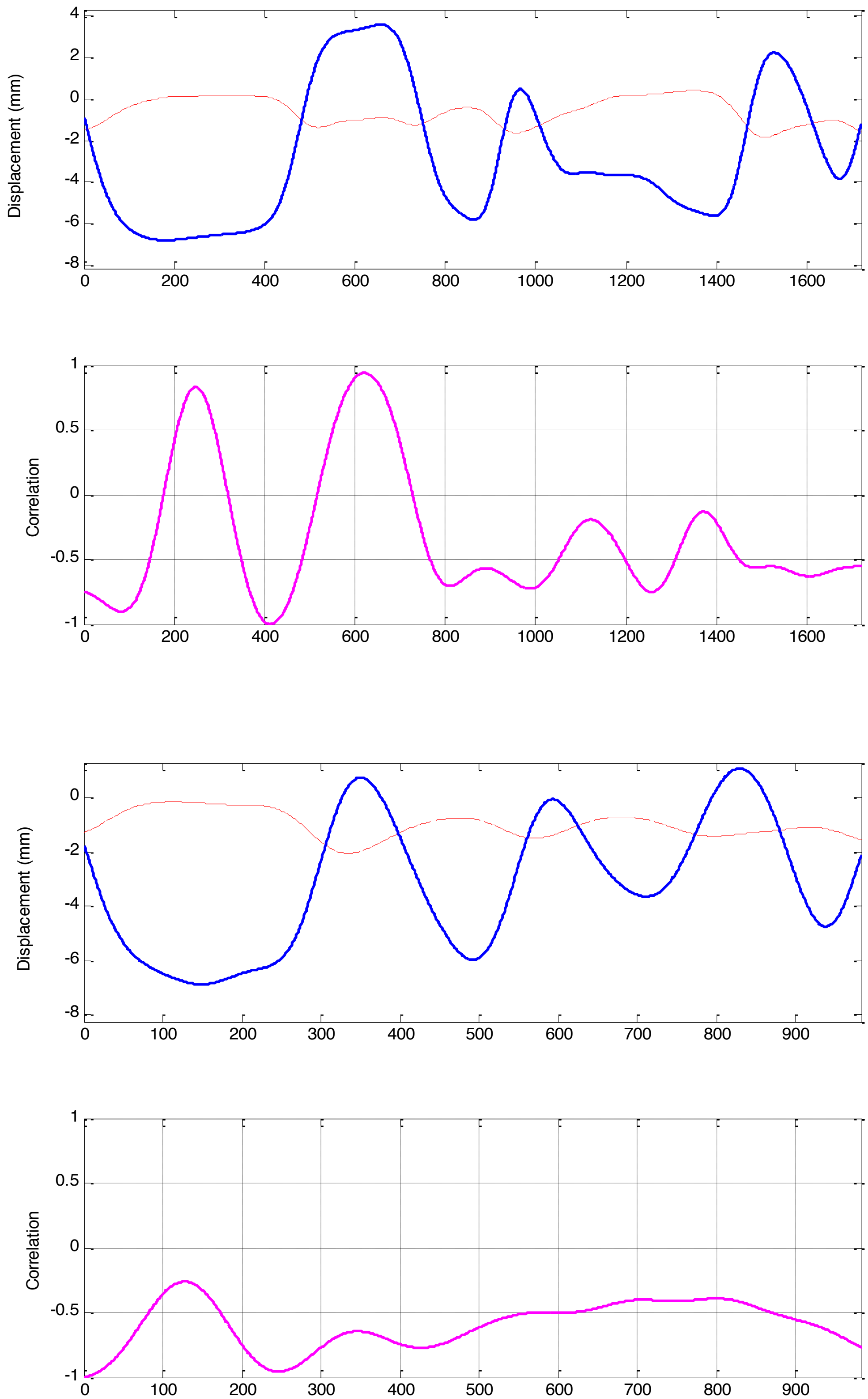
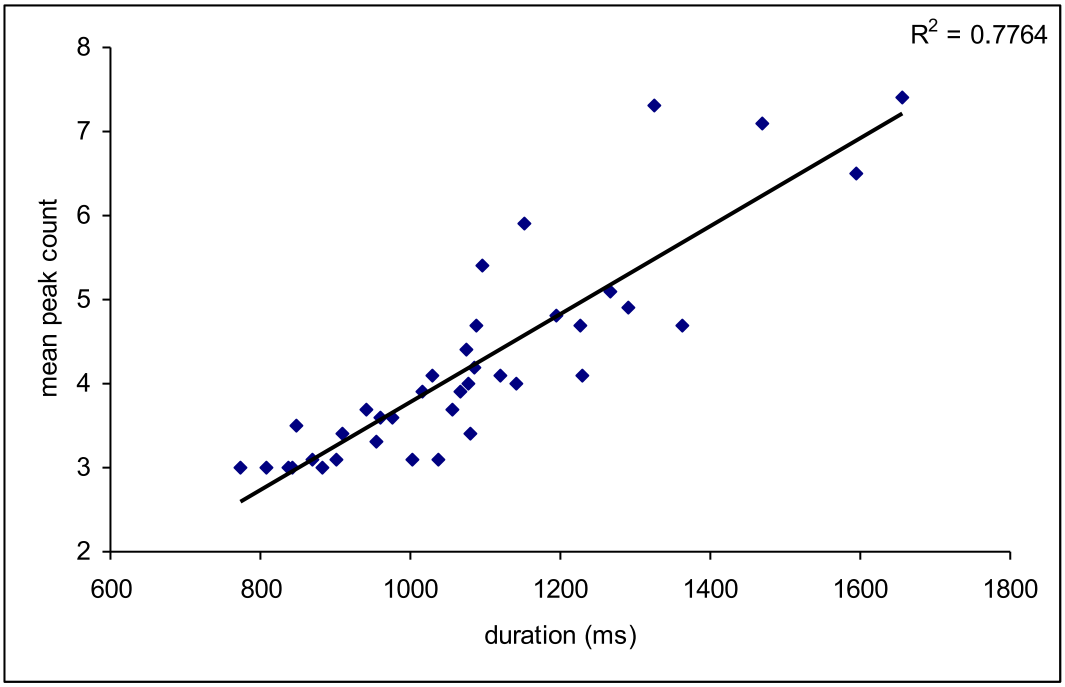

Displacement (mm)
4
2
0
-2
-4
-6
-8
0
200
400
600
800
1000
1200
1400
1600
Correlation
1
0.5
0
-0.5
-1
0
200
400
600
800
1000
1200
1400
1600
Displacement (mm)
0
-2
-4
-6
-8
0
100
200
300
400
500
600
700
800
900
Correlation
1
0.5
0
-0.5
-1
0
100
200
300
400
500
600
700
800
900

Figure 5



R² = 0.7764
8
7
6
5
4
3
2
mean peak count
600
800
1000
1200
1400
1600
1800
duration (ms)

Figure 6

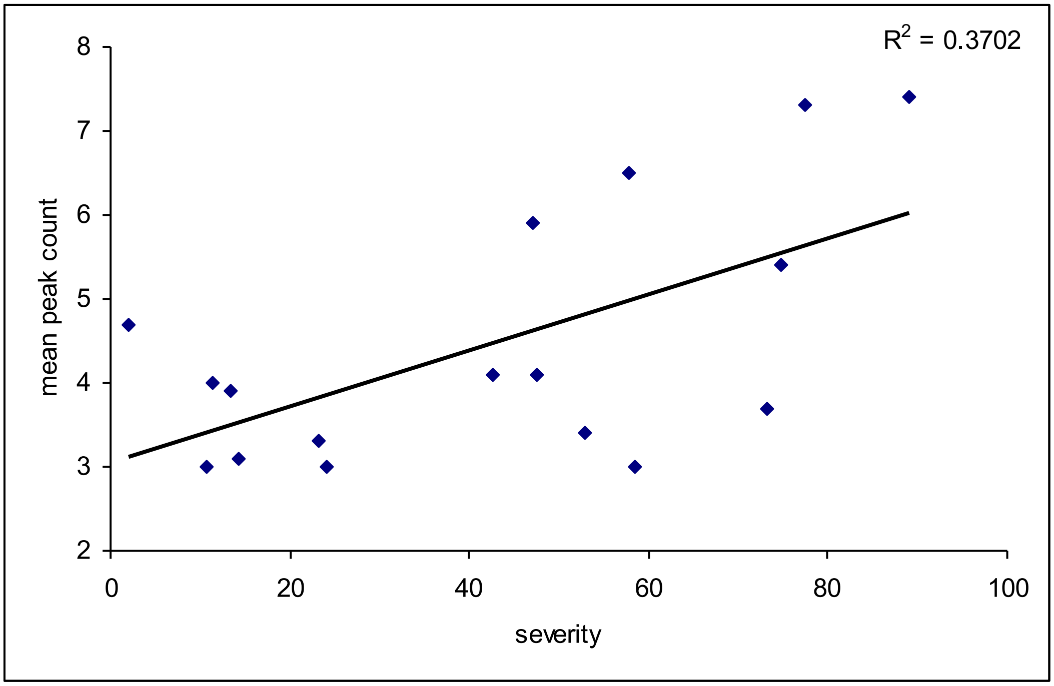


Figure 7