Kinematic Analysis of Tongue Movement Control in Spastic Dysarthria

Heejin Kim[1], Panying Rong[2], Torrey M. Loucks[1,2], Mark Hasegawa-Johnson[1,3]

[1]Beckman Institute, University of Illinois, Urbana, USA
[2]Department of Speech and Hearing, University of Illinois, Urbana, USA
[3]Department of Electrical and Computer Engineering, University of Illinois, Urbana, USA

{hkim17, prong2, tloucks, jhasegaw}@illinois.edu

Abstract

This study provided a quantitative analysis of the kinematic deviances in dysarthria associated with spastic cerebral palsy. Of particular interest were tongue tip movements during alveolar consonant release. Our analysis based on EMA measures indicated that speakers with spastic dysarthria had a restricted range of articulation and disturbances in articulatory-voicing coordination. The degree of kinematic deviances was greater for lower intelligibility speakers, supporting an association between articulatory dysfunctions and intelligibility in spastic dysarthria.

Index Terms: dysarthria, kinematic analysis, EMA

1. Introduction

Dysarthria is a neuromotor speech disorder, which can result from a variety of causes, including traumatic brain injury (TBI), degenerative diseases (e.g., amyotrophic lateral sclerosis, Parkinson's disease), or congenital disorders (e.g., cerebral palsy). Muscular control of speech production is affected in dysarthria, thereby exhibiting deviant speech patterns such as imprecise consonants, distorted vowels, hypernasality, and excessive or monotonic intonation [1]. A large body of research in dysarthria has examined acoustic correlates of speech production. Findings include longer segmental duration, slower transitions from a consonant to the following vowel, less prominent spectral peaks for the fricative /s/, and reduced vowel space compared to normal speech [2-12]. These studies have contributed to advance our understanding of the acoustic attributes of general deviances in dysarthric speech. However, relatively little is known about the underlying articulatory mechanisms responsible for imprecise acoustic outputs.

Studies using the electromagnetic articulography system (EMA) have demonstrated that kinematic measures obtained by EMA are successful in capturing the dynamic aspects of speech movements in normal and disordered speech [13-17]. For example, [15] examined tongue movements of /t, s, k/ produced by speakers with TBI. Findings suggested a relationship between longer consonant durations and speed generation capabilities for some speakers; but for other speakers, larger displacement was a major factor in the perception of longer duration. A study comparing tongue movements of /t, k/ in onset positions across severe TBI vs. mild TBI vs. control groups found a significant difference between the severe TBI and control groups for the release phase of the /t/ consonant [16]. Specifically, a significantly larger maximum acceleration was found for the severe TBI group compared to the control group. These studies suggest that EMA measures can provide an objective quantification of distorted articulatory control in dysarthria. As [13, 18] suggested, articulatory deviances will likely vary depending on the type of disorder, intelligibility levels, and speech task.

The purpose of the current study was to identify kinematic deviances in tongue tip movements associated with spastic cerebral palsy (CP). Spastic dysarthria is noted as the most common type of severe, chronic speech disorder [19]. Our focus was on alveolar consonants in syllable onset positions in American English. The research questions were motivated by perceptual and acoustic reports concerning consonants in dysarthria. First, given the general perception of slower speech and acoustic findings of longer consonant durations, spastic dysarthric speakers might exhibit deviant lingual velocity and acceleration compared to control speakers. However, it is not clear whether the slower speech in spastic CP-related dysarthria is mainly a consequence of reduced movement speed. As shown in [15] with TBI-related dysarthria, slower speech and longer duration might be associated with larger displacement but comparable movement speed to control speakers. Second, [7] reported a less abrupt 'release' of the /s/ to the following vowel based on visual inspection of spectrogram. As the authors speculated, reduced abruptness might be due to deviances in coordinating tongue movement during the /s/ release with laryngeal control for the following vowel. Finally, given the perceptual impression of fricated stops in dysarthria, kinematic analysis of /t, d/ vs. /s, z/ may reveal a reduced distinction between stops and fricatives. Direct observation of articulatory movements is necessary to determine kinematic factors that can account for the perceptual and acoustic deviances in dysarthria.

The following research questions were addressed: (1) what type of tongue tip deviations do speakers with spastic CP exhibit? (e.g., a restricted range of movement, disturbances in coordination with phonation, reduced speed), (2) is there any evidence of reduced kinematic distinctions between stops and fricatives?, and (3) is there an association between a speaker's intelligibility level and phoneme-specific kinematic measures?

2. Method

2.1. Participants

Four American English native speakers participated in the current study: 2 speakers diagnosed with CP (1 male and 1 female) and 2 male control speakers, ranging in age from 22 to 43. Both speakers in the CP group were diagnosed as principally demonstrating spastic dysarthria. Control speakers reported no history of a language or speech disorder. The intelligibility scores of participants were based on five naive listeners' word transcription in our previous work [20]. Intelligibility scores for the participants in this study were 39% for the male speaker (M01) and 29% for the female speaker (F01). Control speakers' intelligibility scores were between 96% and 98%.

2.2. Recording procedures and material

All kinematic data were collected using the EMA system (AG500, Carstens, Germany) at a sampling rate of 200Hz. In addition to EMA, electroglottograph (EGG) and surface electromyograph (EMG) data were also collected simultaneously. Since this report mainly concerns kinematic and acoustic measures, instrumentation and recording descriptions are only reported for EMA. Two receiver sensors were attached to the tongue at 1cm and 3cm from the tongue tip, respectively. Three sensors were affixed to reference points: on the nose bridge, and the right and left tragus, respectively. Five additional sensors were attached to other articulators (i.e., jaw, upper lip, lower lip, left lip corner, and right lip corner). One sensor was used at the end for palate plane tracing. Prior to attaching the sensors, an experimenter marked the points using a surgical marker. This method ensured the consistency in sensor placement across participants and recording sessions. The three reference sensors were used for normalizing the sensor movements with respect to head movements.

The stimuli were from a large corpus used for developing automatic speech recognition systems for dysarthric speech, which included a comprehensive list of the following words: 10 digits (e.g., zero, one), 26 international radio alphabet letters (e.g., alpha, bravo), 26 computer commands (e.g., enter, delete), 200 phonetically-balanced words (e.g., are, bad), and 76 minimal pairs (e.g., buy, die, guy). The phonetically-balanced word list was adapted from the American National Standards Institute (ANSI) standards. While a speaker sat inside the EMA cube, he or she was asked to read a word displayed on a Power-Point slide on a laptop monitor. An experimenter sat three feet in front of the cube and advanced the slides after the speaker articulated each word. Speakers with CP participated in two separate recording sessions and they repeated the word list 2-3 times per session. Each session lasted approximately two hours. Control speakers participated in one recording session and repeated the word list three times. After recording, the EMA data were corrected for head motions and normalized using the NormPos software (Carstens, Germany).

The target consonants of this study were /t/, /d/, /s/, and /z/ in the context of one-syllable words of CV or CVC, wherein the first C was one of the target consonant and the V=[ɑɪ] or [æ]. The resulting word list included the following words: *tie, tab, die, dike, size, sigh, zion*. Each speaker produced these words 3-4 times, yielding a total of 6-8 tokens per target consonant phoneme.

2.3. Analysis

The normalized data were exported to MATLAB (version R2009a, The MathWorks, Inc). First, the data were filtered at 10Hz with a low-pass phase-corrected filter. Then, the velocity and acceleration of tongue tip movements in the inferior-superior dimension were determined by taking the integral of the displacement and resulting velocity signals. Tongue-tip displacement (Inf-Sup) and the corresponding velocity and acceleration curves were viewed using a custom-written MATLAB interface. The opening phase of the target consonant was defined as follows. The movement onset point was determined as the point of maximum acceleration. For stop consonant, the beginning point was also marked by the acoustic burst onset. The end of the opening phase was defined as the point of maximum deceleration. For fricatives, the frication phase was also marked. The beginning of the frication phase corresponded to the onset of visual frication in the acoustic waveform. The end
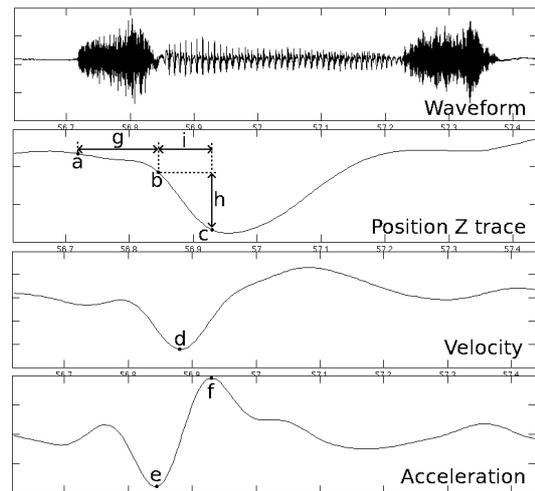

Figure 1: *Example of kinematic measures in the opening phase of a tongue tip movement produced by speaker 'Control1' for the /s/ in the word 'size': (a) beginning of frication, (b) beginning of the opening phrase, (c) end of the opening phase, (d) MaxV, (e) MaxA, (f) MaxD, (g) DurF, (h) DisO, (i) DurO.*

of the frication phase coincided with the beginning of the opening phase. For the opening phase, the following measures were obtained.

(1) DisO: opening displacement (= tongue tip height difference between the beginning and end points)
(2) MaxV: maximum velocity
(3) MaxA: maximum acceleration
(4) MaxD: maximum deceleration
(5) DurO: acoustic duration of the opening phase

For fricatives, the following additional parameters were obtained.

(6) Variability: a measure of the tongue tip stability during frication (= mean square error of the z-curve during the frication phase, relative to a regression line)
(7) AccLag: time difference between MaxA and the onset of voicing for the following vowel (= voicing starting time – MaxA time)
(8) DurF: Acoustic duration of the /s, z/ noise

Variability was measured to assess the relative stability in maintaining the articulatory target during frication. Large values would indicate the fluctuation of the tongue tip height during frication. AccLag was obtained as a measure for the timing relation between tongue lowering movement and the initiation of voicing for the following vowel. Thus, AccLag measures were limited to the voiceless fricative /s/. The onset of voicing in the vowel was determined by visual inspection of waveform and spectrograms. Large negative values of AccLag indicated that acceleration for the tongue tip lowering occurred at a later time point relative to voicing onset for the subsequent vowel. The kinematic parameters (1-5) were tested with two-way ANOVA analyses, with the independent factors of Speaker and Manner (i.e., stop vs. fricative). Fricative specific parameters were tested with one-way ANOVA, with the independent factor of Speaker. Correlation between speakers' intelligibility scores and each of the kinematic parameters was examined using Nonparametric Kendall tau-b analysis.

Table 1: *Mean values of kinematic and durational measures as a function of Speaker and Manner. Values in parentheses represent standard deviations.*

	Control1		Control2		M01		F01	
	stop	fricative	stop	fricative	stop	fricative	stop	fricative
DisO	10.21	7.33	10.50	7.76	6.61	5.50	3.97	4.22
	(2.92)	(1.53)	(3.20)	(0.88)	(4.19)	(2.28)	(4.03)	(2.70)
MaxV	140.17	110.42	150.73	113.49	91.07	76.59	50.13	62.95
	(31.94)	(19.44)	(55.84)	(14.46)	(48.44)	(29.71)	(41.89)	(42.65)
MaxA	2151.02	1903.47	2804.68	1972.63	1478.37	1286.82	973.22	1587.55
	(552.19)	(414.03)	(1063.60)	(251.35)	(789.66)	(552.13)	(660.18)	(1043.87)
MaxD	2566.78	2384.62	3092.45	2291.02	1682.77	1273.88	817.11	1242.62
	(1249.67)	(748.50)	(1836.55)	(550.13)	(1028.54)	(372.24)	(743.65)	(1282.75)
DurO(ms)	92.50	85.00	84.44	85.83	98.33	93.33	138.33	86.36
	(13.57)	(4.33)	(16.48)	(3.76)	(34.13)	(30.21)	(84.70)	(26.47)
Variability	-	.034	-	.050	-	.014	-	.267
	-	(.019)	-	(.024)	-	(.009)	-	(.452)
AccLag	-	3.42	-	10.60	-	-22.20	-	-46.99
	-	(4.29)	-	(7.92)	-	(37.29)	-	(36.32)
DurF(ms)	-	121.67	-	140.83	-	101.67	-	450.00
	-	(23.58)	-	(23.33)	-	(53.91)	-	(81.12)

3. Results

Table 1 shows the mean values and standard deviations of kinematic and durational measures for each speaker. In general, speakers with CP exhibited smaller DisO, MaxV, MaxA and MaxD values, but greater Variability, AccLag and DurF values when compared to the control group. A significant main effect of Speaker was found for DisO ($F(3,72) = 11.816$, $p < .001$), MaxV ($F(3,72) = 16.320$, $p < .001$), MaxA ($F(3,72) = 9.3200$, $p < .001$), and MaxD ($F(3,72) = 10.256$, $p < .001$). Post-hoc Tukey's HSD tests showed that speakers with CP had significantly smaller opening displacement, lower velocity, lower acceleration and lower deceleration values compared to control speakers. A significant effect was not found for DurO. A significant main effect of Manner was only found for DisO ($F(1,72) = 5.256$, $p < .05$). The interaction Speaker x Manner was significant for MaxA ($F(3,72) = 3.077$, $p < .05$), indicating that the effect of Manner varied across speakers. Figure 2 and Figure 3 illustrate mean values for DurO and MaxA measures, respectively.
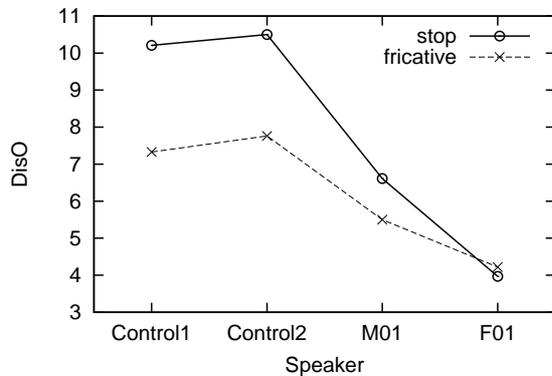

Figure 2: *Mean values of DisO (opening displacement) as a function of Speaker and Manner.*
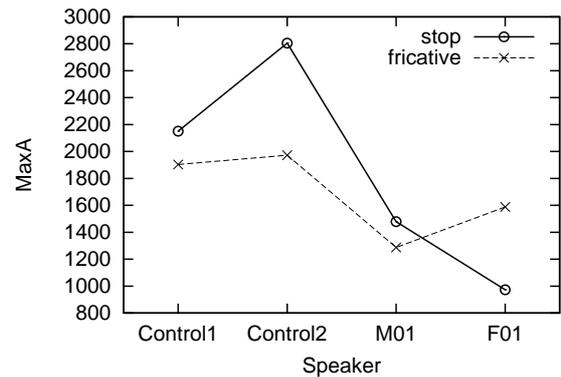

Figure 3: *Mean values of MaxA as a function of Speaker and Manner.*

Regarding the fricative-specific measures, a significant effect of Speaker was found for DurF ($F(3,31) = 89.376$, $p < .001$) and AccLag ($F(3,17) = 4.758$, $p < .05$). Subsequent post-hoc analysis revealed that the speaker F01 had a significantly longer duration of frication, and significantly larger negative AccLag values compared to the other speakers. Regarding the Variability measure, the speaker F01 had a greater mean value than the other three, indicating a greater degree of difficulty in maintaining the tongue tip height; but, it failed to reach significance. Figure 4 shows mean values for AccLag measures.

According to nonparametric Kendall tau-b analyses, a speaker's intelligibility was significantly correlated with the following measures: DisO ($r = -.49$, $p < .001$), MaxV ($r = .51$, $p < .001$), MaxA ($r = .43$, $p < .001$), MaxD ($r = .52$, $p < .001$), AccLag ($r = -.277$, $p < .01$), and DurF ($r = -.47$, $p < .001$).

4. Conclusions

This study provides a quantitative analysis of tongue tip kinematic deviations in spastic dysarthria. Our results demonstrate
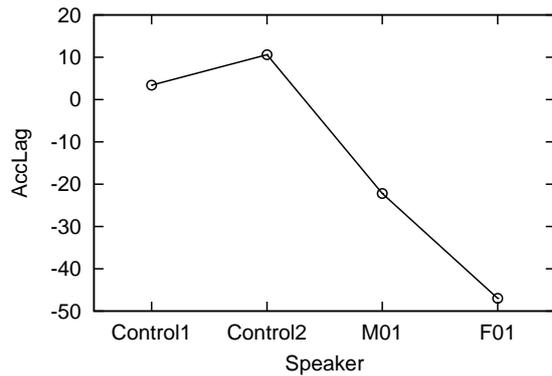
Figure 4: *Mean values of AccLag as a function of Speaker.*

that speakers with spastic CP-associated dysarthria exhibited a restricted range of movements during alveolar consonant release. Evidence for restricted motion is based on reduced displacement and slower velocity. Stark evidence for disturbed coordination was found from the AccLag measure: speakers with CP exhibited large negative values relative to the small positive values of control speakers. This indicates that acceleration for tongue tip lowing occurred later than voicing onset in the following vowel. This finding supports the speculation in [7] regarding reduced ability to coordinate supraglottal and glottal systems in dysarthria. In the current study, the onset of the glottal movements was determined by visual inspection of acoustics. A more direct observation of glottal events using EGG is currently being used in our lab to better characterize onset/offset and vibratory parameters of phonation in dysarthria.

In comparing stops vs. fricative, a greater movement range and faster speeds were generally found for stops relative to fricatives. This distinction tended to be reduced for speakers with CP, mainly by reduced movement speeds for stops. Finally, strong correlations were found between speaker intelligibility and kinematic measures. Particularly, it was evident that speaker F01, the lower intelligibility speaker, had more difficulties with articulatory movement speeds and coordination compared to other speakers. This finding suggests that phoneme specific kinematic deviations of a single articulator are associated with perceived speech intelligibility in spastic dysarthria. Future work will include further perceptual ratings of specific sounds and syllables, in order to assess the direct correlations among articulatory, acoustic and perceptual characteristics in dysarthria.

5. Acknowledgements

6. References

[1] Darley, F., Aronson, A. and Brown, J., . "Motor Speech Disorders", Philadelphia: W.B. Saunders Inc., 1975.

[2] Kent, R.D. and Netsell, R., "A case study of an ataxic dysarthric: cineradiographic and spectrographic observations", J. Speech and Hearing Disorders, 40, 115-134, 1975.

[3] Kent, R.D. Netsell, R. and Abbs, J., "Acoustic characteristics of dysarthria associated cerebellar disease", J. Speech, Language, and Hearing Research, 22, 627-648, 1979.

[4] Kent, J.F., Kent, J.F., Rosembek, J., Weismer, G., Martin, R. and Sufit, R., "Quantitative description of the dysarthria in women with amyotrophic lateral sclerosis", J. Speech, Language, and Hearing Research, 35:4, 723-733, 1992.

[5] Weismer, G., Martin, R., Kent, R.D. and Kent, J.F., "Formant trajectory characteristics of males with amyotrophic lateral sclerosis", J. acoust. Soc. Am., 92:2, 91-98, 1992.

[6] Tjaden, K. and Turner, G.S., "Spectral properties of fricatives in amyotrophic lateral sclerosis", J. Speech, Language, and Hearing Research, 40, 1358-1372, 1997.

[7] Chen, H. and Stevens, .N., "An acoustical study of the fricative /s/ in the speech of individuals with dysarthria", J. Speech and Hearing Research, 44, 1300-1314, 2001.

[8] Turner, G.S., Tjaden, K. and Weismer, G., "The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis", J. Speech and Hearing Research, 38, 1001-1013, 1995.

[9] Liss, L.M., Spitzer, S.M., Cavinsess, J.N., Adler, C. and Edwards, B.W., "Lexical boundary error analysis in hypokinetic and ataxic dysarthria", J. acoust. Soc. Am., 107, 3415-3424, 2000.

[10] Weismer, G., Laures, J.S., Jeng, J.-Y., Kent, R.D. and Kent, J.F., "Effect of speaking rate manipulations on acoustic and perceptual aspects of dysarthria in amyotrophic lateral sclerosis", Folia Phoniatrica et Logopaedica, 52, 201-219, 2000.

[11] Weismer, G., Jeng, J.Y., Laures, J.S., Kent, R.D. and Kent, J.F., "Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders", Folia Phoniatrica et Logopaedica, 53, 1-18, 2001.

[12] Higgins, C.M. and Hodge, M.M., "Vowel area and intelligibility in children with and without dysarthria", J. Medical Speech-Language Pathology, 10, 271-277, 2002.

[13] Van Lieshout, P. and Moussa, W., "The assessment of speech motor behavior using electromagnetic articulography", The Phonetician, 81:I, 2000.

[14] Goozee, J.V., Murdoch, B.E., Theodoros, D.G. and Stokes, P.D., "Kinematic analysis of tongue movements in dysarthria following traumatic brain injury using electromagnetic articulography", Brain Injury, 14:2, 153-174, 2000.

[15] Murdoch, B.E. and Goozee, J.V., "EMA analysis of tongue function in children with dysarthria following traumatic brain injury", Brain Injury, 17:1, 79-93, 2003.

[16] Kuruvilla, M., Murdoch, B. and Goozee, J., "Electromagnetic articiulography assessment of articulatory function in adults with dysarthria following traumatic brain injury", Brain Injury, 21:6, 601-613, 2007.

[17] Jaeger, M., Hertrich, I., Stattrop, U., Schonle, P. and Ackermann, H., Speech disorders following severe traumatic brain injury: kinematic analysis of syllable repetitions using electromagnetic articulogrphy, Folia Phoniatrica et Logopaedica, 52, 187-196, 2000.

[18] Kent, R.D., Kent, J.F., Weismer, G., and Duffy, J.R., "What dysarthria can tell us about the neural control of speech", J. Phonetics, 28, 273-302, 2000.

[19] Love. R.J., "Childhood Motor Speech Disability", Allyn and Bacon, Boston, 1992.

[20] Kim, H., Hasegawa-Johnson M, Perlman A., Gunderson J., Huang T., Watkin K. and Frame, S. "Dysarthric speech database for universal access research", Interspeech Proc., 2008.