This is the Accepted Manuscript version of an article published by Sage Publishing in 2018 in Language and Speech, 61(1), 31-42, available online: https://doi.org/10.1177/0023830917695853

Spectral coefficient analyses of word-initial stop consonant productions suggest similar anticipatory coarticulation for stuttering and nonstuttering adults

Santosh Maruthy

All India Institute of Speech and Hearing

Mysore, Karnataka, India

Yongqiang Feng

Key Laboratory of Speech Acoustics and Content Understanding,

Institute of Acoustics, Chinese Academy of Sciences,

Beijing 100190, China

Ludo Max

Department of Speech and Hearing Sciences, University of Washington,

Seattle, WA 98105-6246

Haskins Laboratories, New Haven, CT 06511-6695

Corresponding Author:

Dr. Santosh Maruthy
Reader in Speech Sciences
Department of Speech-Language Sciences
All India Institute of Speech and Hearing
Manasagangothri
Mysore-570006
India
E-mail: santoshm79@gmail.com
Phone: +91-0821 2512523 (Office)
Fax: +91-0821-2510515

Abstract

A longstanding hypothesis about the sensorimotor mechanisms underlying stuttering suggests that stuttered speech dysfluencies result from a lack of coarticulation. Formant-based measures of either the stuttered or fluent speech of children and adults who stutter have generally failed to obtain compelling evidence in support of the hypothesis that these individuals differ in the timing or degree of coarticulation. Here, we used a sensitive acoustic technique—spectral coefficient analyses—that allowed us to compare stuttering and nonstuttering speakers with regard to vowel-dependent anticipatory influences as early as the onset burst of a preceding voiceless stop consonant. Eight adults who stutter and eight matched adults who do not stutter produced C_1VC_2 words, and the first four spectral coefficients were calculated for one analysis window centered on the burst of C_1 and two subsequent windows covering the beginning of the aspiration phase. Findings confirmed that the combined use of four spectral coefficients is an effective method for detecting the anticipatory influence of a vowel on the initial burst of a preceding voiceless stop consonant. However, the observed patterns of anticipatory coarticulation showed no statistically significant differences, or trends toward such differences, between the stuttering and nonstuttering groups. Combining the present results for fluent speech in one given phonetic context with prior findings from both stuttered and fluent speech in a variety of other contexts, we conclude that there is currently no support for the hypothesis that the fluent speech of individuals who stutter is characterized by limited coarticulation.

Keywords

Stuttering, anticipatory coarticulation, timing, spectral analysis, acoustic

Spectral coefficient analyses of word-initial stop consonant productions suggest similar anticipatory coarticulation for stuttering and nonstuttering adults

Multiple experimental findings suggest a wide variety of differences in temporal aspects of speech articulation between individuals who stutter and individuals who do not stutter (see reviews in Bloodstein & Bernstein Ratner, 2008; Max, 2004). Often, however, those studies merely confirmed that the speech of stuttering individuals as a group is characterized by longer durations of acoustic and kinematic intervals rather than a specific timing deficit (e.g., Max & Gracco, 2005; Max & Yudman, 2003). One longstanding hypothesis that does attribute stuttering specifically to deficits in speech motor *timing* suggests that dysfluencies may occur due to problems with the timing of anticipatory, overlapping speech movements (Harrington, 1987; Stromsta, 1965, 1980, 1986, 1987; Van Riper, 1982).

Originally, this hypothesis that stuttering results from a lack of anticipatory, context-based adjustments in articulatory positioning (i.e., anticipatory coarticulation) developed from claims, based on perceptual observations, that stuttered consonant-vowel (CV) syllable repetitions sometimes sound like the speaker produced a neutral vowel /ə/ instead of the syllable's correct vowel that occurs in the eventual fluent production of the syllable (Montgomery & Cooke, 1976; Van Riper, 1971, 1982). Based on spectrographic analyses of a small number of dysfluencies ("at least one sample" per child, Stromsta, 1965, p. 317) produced by 24 children who later persisted in their stuttering, Stromsta (1965, 1986, 1987) argued that coarticulation problems are also reflected in atypical formant transitions in the stuttered repetitions of a syllable as compared with the final, fluent production of that same syllable. It is important to realize, however, that Stromsta's work(1965, 1986) was based on narrow-band spectrographic analyses[1], and that transitions in the frequency components shown in this type of acoustic analysis reflect changes in the harmonics of the fundamental frequency (laryngeal information) rather than in the formants (articulatory information).When a wide-band spectrogram example of a single dysfluency was provided retrospectively (Stromsta, 1986, 1987), the chosen syllable was the pronoun “I” (diphthong /ɑɪ/) which has a clear formant transition only in the second half of its total duration, and the supposedly “missing” formant transition in stuttered repetitions of this syllable was due to each repeated attempt being terminated before the onset of the formant transition.

Accordingly, subsequent spectrographic analyses have generally failed to find evidence in support of Stromsta’s suggestion that stuttered dysfluencies are associated with incorrect coarticulation as compared with the final, fluent production of the word (Harrington, 1987; Howell & Vause, 1986; note that the findings reported by Montgomery & Cooke, 1976, do not address the same question given that they compared the final *fluent* production of the stuttered word with a separate fluent production of the word). Although some stuttered repetitions do have formant transitions that do not reach the target frequency or that change in the wrong direction (Harrington, 1987; Yaruss & Conture, 1993), these cases provide no direct evidence for faulty coarticulation as they can be directly attributed to, respectively, the early time at which the transition is interrupted and the fact that the articulators are returning to their position for the initial sound rather than moving on to a different subsequent sound (Harrington, 1987).

A different approach to examining anticipatory coarticulation in the speech of individuals who stutter has focused on perceptually fluent utterances[2]. Frisch, Maxfield, and Belmont (2016) compared anticipatory coarticulation in adults who do and who do not stutter by means of kinematic data (ultra sound images) showing tongue position during velar consonant-vowel transitions with vowels that differed in tongue advancement. The obtained results revealed no between-group differences. Two acoustic studies that specifically aimed to investigate anticipatory coarticulation have relied on a locus equation approach. Locus equations are linear regression fits based on second formant (F2) measurements at the onset of a CV formant transition and in the steady state portion of the vowel (i.e., the "offset" or "target" frequency) across syllables with the same consonant but different vowels (Lindblom, 1963, 1998; Sussman, McCaffrey, & Matthews, 1991). The rationale is that a slope of zero (i.e., transition onset is not correlated with vowel-specific transition offset) indicates an absence of anticipatory coarticulation in the consonant, and that the actual slope value (with a maximum of 1) indicates the strength of coarticulation. A study on the fluent speech of children who stutter (Chang, Ohde, & Conture, 2002) as well as a study on both dysfluent and fluent productions of adults who stutter (Sussman, Byrd, & Guitar, 2011) failed to find atypical locus equation slopes in the CV syllables of stuttering speakers.

Besides the fact that the validity of locus equation slopes as a measure of coarticulation has been questioned based on the absence of a measurable link between such slopes and direct kinematic measures of articulatory overlap (Löfqvist, 1999; Tabain, 2000; Tabain, 2002), this approach is also limited in its applicability given that, typically, only voiced consonants are analyzed. The latter limitation is related to great difficulties involved in determining, from spectrographic displays, the F2 onset frequency for voiceless consonants where most of the formant transition occurs during the unvoiced interval (e.g., during the interval corresponding to the voice onset time of a voiceless stop consonant). Hence, it would be beneficial to be able to determine the anticipatory influence of a vowel on a preceding consonant at the very beginning of the consonant and in such a manner that the analyses can be applied to both voiced and voiceless consonants.

Recently, work from our own group has demonstrated that this goal can be accomplished by calculating the first four spectral coefficients (mean, standard deviation, skewness and kurtosis) at the onset of initial stop consonants in $C_1V(C_2)$ syllables (Feng, Hao, Xue, & Max, 2011). Although others had already used the first spectral coefficient to quantify anticipatory coarticulation in clinical populations (Tjaden, 2003; Tjaden and Wilding, 2005), the work by Feng et al. (2011) demonstrated the sensitivity of an approach combining all four spectral coefficients. For each coefficient, the analyses revealed strong and statistically significant anticipatory effects of vowel context on the initial burst and aspiration noise of the stop consonants /k,t,p/. However, different coefficients resulted in a different ranking of the three consonants in terms of degree of coarticulation. Thus, the findings indicate that the combined use of all four spectral coefficients offers the most sensitive approach for detecting subtle vowel-related anticipatory adjustments.

Here, we applied the analyses described by Feng et al. (2011) to compare stuttering and nonstuttering adults with regard to the degree of anticipatory coarticulation in the early part of the release of a voiceless stop consonant. Using a limited but already available data set in which adults who stutter (AWS) and matched adults who do not stutter (AWNS) had produced C_1VC_2 words, we calculated the first four spectral coefficients for one analysis window centered on the burst of C_1 and two subsequent windows covering the aspiration phase. Only perceptually fluent words were included in the analyses.

Method

Subjects

The data set included 8 AWS (mean age = 27.75, age range 19-49, one female) and 8 age-matched (+/- 3 years) and gender-matched AWNS (mean age = 25.87, age range 20-48, one female) who had participated after giving informed consent. Based on the Stuttering Severity Instrument for Children and Adults – Third Edition (SSI-3, Riley, 1994), stuttering severity was very mild for one subject, mild for three, moderate for two, and very severe for two participants. Among the AWS, one participant reported being left-handed, the others reported being right-handed. Among the AWNS, all participants reported being right-handed. All participants in both groups were native speakers of American English, and all reported that they had never been diagnosed with any neurological, psychological, or communication problems (other than stuttering in the AWS group). Behavioral pure tone hearing testing at all octave frequencies from 250 to 4000 Hz revealed that all participants except for two AWS had bilateral hearing thresholds at or below 25 dB HL. These two AWS each had one unilateral 4000 Hz threshold at 40 dB HL. Detailed individual participant characteristics are provided in Table 1.

Procedure

Participants were seated inside a sound booth, and read individually displayed monosyllabic words from a computer monitor at a rate of 15 words per minute. Participants were provided with color-coded visual feedback to assist with maintaining their speech intensity between 72 and 78 dB SPL as measured 15 cm from the mouth. The overall recording session for each participant included four blocks of 180 trials. Within each block of 180 trials, three different monosyllabic words were presented in random order in each cycle of three trials (thus 60 productions of each target word in each block). The three words produced during the recording session were C_1VC_2 syllables in which C_1 was always the alveolar voiceless stop consonant /t/ and C_2 was always the velar voiceless stop consonant /k/. The vowel was the front /ɛ/, central /ʌ/, or back /ɔ/ (“tech,” “tuck,” “talk”).

Given that the data had originally been recorded for a study that involved auditory feedback manipulations (in particular formant frequency shifts) during the middle portion of each block of trials, we selected for the present analyses only those trials that had been produced during each block’s baseline phase with typical, unaltered auditory feedback. This selection included the first 30 trials from two blocks and the first 60 trials from the other two blocks. Thus, the analyses included 180 trials per subject for a total of 1440 trials. Both the first author (a certified speech-language pathologist in India who completed postdoctoral research training at the University at Washington) and a graduate student with expertise in fluency disorders (a native speaker of English serving as a research assistant at the University of Washington) independently judged all included trials to be perceptually fluent.

In terms of possible effects of the feedback manipulation on the extent of participants' coarticulation during the non-manipulated baseline trials, the key points to take into consideration are that: (a) by shifting all formant frequencies proportionally the same amount, the feedback alterations affected the perception of each vowel in an identical manner and thus the distinction of these vowels in acoustic vowel space remained intact; (b) the feedback alterations affected only voiced segments and thus did not in any way affect the burst or aspiration phase of the analysed stop consonants; (c) altered feedback was always followed by 30 trials in a post-perturbation wash-out phase during which participants heard unaltered feedback and their vowel productions returned to baseline; and (d) after the wash-out phase of each block of trials, an additional 3-5 minutes of rest were provided prior to initiation of the baseline phase for the next block.

Instrumentation and data extraction

Participants' speech was recorded with a microphone (SM58, Shure, Evanston, IL, amplified by a DI/O Preamp System II, ART, Rochester, NY) and CD recorder (XL-R5000BK, JVC, Wayne, NJ). The mouth-to-microphone distance was maintained at 15 cm. Offline, the 44100 Hz CD tracks were converted to "wav" format and down-sampled to 22050 Hz.

The burst and early aspiration noise of the word-initial C_1 consonants were used to extract all four spectral coefficients for the purpose of examining the coarticulatory influence of the subsequent front, central, or back vowel (Feng et al., 2011). Using custom scripts for the Praat software (http://www.fon.hum.uva.nl/praat/), each word was first viewed as a wide-band spectrogram, and then a zoomed view of the initial part of C_1 was displayed in order to manually mark the onset of the initial burst of C_1. The script then extracted the acoustic signal from -10 to 30 ms relative to the marker. Custom written MATLAB routines were then used to calculate the four spectral coefficients from linear frequency scale spectra obtained for three successive 11.61 ms windows (i.e., 256 samples per window) with the first window centered on the onset of the C_1 burst. After applying pre-emphasis and Hamming windowing to each of these brief segments, a 256-point Fast Fourier Transform (FFT) was computed, and the power spectrum across discrete frequency samples was derived. The first four spectral coefficients (mean, standard deviation, skewness, and kurtosis) were computed after normalization of the power spectrum. The power spectrum was normalized using the following formula:

$$p(f_k) = \frac{P(f_k)}{\sum_{k=0}^{N/2} P(f_k)}$$

Where P (fk) is the power spectrum, fk is 2fNqk/N, k = 0, 1, N/2, and N = 1024. fNq is the Nyquist frequency, which was 11025 Hz. This suggests the upper limit of the frequency range that was used to compute spectral coefficients. The first two spectral coefficients (mean and standard deviation) are represented in kHz, whereas the third and fourth coefficients (skewness and kurtosis) are dimensionless.

Statistical analysis

For each participant, the extracted coefficients for each analysis window were averaged across repeated productions of the same target word. Then, repeated measures ANOVA was applied with vowel (3 levels) and window (3 levels) as within-subject variables and group (AWS vs. AWNS) as a between-subject variable. The significance of all repeated measures ANOVA was determined with Huynh-Feldt epsilon-corrected degrees of freedom to correct for any violations of the sphericity assumption (Max & Onghena, 1999). Given that previous studies have generally failed to find coarticulation differences between stuttering and nonstuttering speakers, and given that the purpose of the present study was to determine whether there is at least some indication that this hypothesis warrants continued study, statistical power was maximized and Type I errors were considered less problematic than Type II errors (Kirk, 2012). Therefore, the significance of all ANOVA analyses was determined based on uncorrected α values (as correcting for the number of tests would decrease power by reducing Type I errors in favor of Type II errors). For each factor and each interaction of factors, partial eta-squared ($\boldsymbol{\eta_p^2}$) was calculated as a measure of effect size.

Results

Table 2 includes the descriptive data for all four spectral coefficients separated by Window, Vowel, and Group. Table 3 includes the results of all inferential statistics.

Overall, results replicated those reported by Feng et al. (2011) in showing that spectral coefficient analyses offer a sensitive method for detecting the anticipatory influence of a vowel on a preceding stop consonant (here always /t/). Such an influence was detected with all four spectral coefficients: statistically significant main effects of the upcoming vowel were associated with η_p^2 effect sizes of 60%, 42%, 23%, and 24% for the consonant's spectral mean, standard deviation, skewness, and kurtosis, respectively. In addition, results also replicated those of Feng et al. (2011) in confirming statistically significant main effects of analysis window (the three windows differed in their temporal distance to vowel onset) and statistically significant interactions between target vowel and analysis window. With regard to the latter interactions, vowel effects on spectral skewness and kurtosis were, in fact, largest in the first analysis window. Hence, this work confirms again that spectral coefficient analyses can reveal vowel-related coarticulation effects as early as the acoustic burst of the preceding consonant.

In contrast—and despite the intentionally liberal significance criterion—none of the main effects of Group and none of the two-way or three-way interactions involving the Group variable showed any statistically significant effects. In fact, none of these effects even showed a trend toward significance at the uncorrected α levels, and the corresponding η_p^2 effect sizes were very small (only 1 of 12 values exceeded .08). Most important for the present purposes is the absence of statistically significant interactions Group × Vowel or Group × Vowel × Window as each of these analyses were direct tests of the hypothesis that stuttering and nonstuttering speakers may differ in the degree or timing of coarticulation. Nevertheless, it is worth noting that the absence of a significant Group main effect indicates that the stuttering group also did not differ from the nonstuttering group in the detailed spectral characteristics of voiceless stop consonant production in general.

Figure 1 graphically summarizes the above findings. The vowel effects, window effects, and vowel by window interactions across the four spectral coefficients can be seen for each group. As is also clear from the figure, however, the variations in vowel effect across the three analysis windows (i.e., as the analysis window moves closer toward the onset of the vowel) are highly similar for the groups of stuttering and nonstuttering speakers.

Discussion

The purpose of this study was to compare anticipatory coarticulation in stuttering vs. nonstuttering adults. To measure coarticulation in these two groups of speakers, we calculated the first four coefficients (mean, standard deviation, skewness, and kurtosis) of the acoustic spectrum associated with the production of a perceptually fluent voiceless stop consonant /t/ preceding front, mid, and back vowels in /t/V/k/ words. Specifically, following procedures previously documented to provide a sensitive technique for detecting true anticipatory coarticulation effects (Feng et al., 2011), we calculated the four spectral coefficients for each of three successive 11.61 ms windows, with the first window centered on the burst of the word-initial voiceless stop consonant and the next two windows extending into the aspiration phase of this consonant.

As in our previous work with only fluent speakers, statistically significant Vowel effects, Window effects, and Vowel by Window interactions revealed clear anticipatory effects of the vowel on the initial consonant, as early as the acoustic burst associated with the release of this consonant's articulatory obstruction between the tongue tip and alveolar ridge. The combination of all four spectral coefficients again proved to be most informative as the aforementioned effects varied in strength across the different coefficients. For example, based on effect size measures, the spectral mean was the most sensitive measure of the Vowel main effect, but the least sensitive measure of the Vowel by Window interaction, a finding that replicates our previous results based on only nonstuttering speakers (Feng et al., 2011).

It is worth noting that the obtained findings for the initial consonant's spectral mean are fully consistent with the interpretation that these effects reflect adjustments in articulatory posturing in anticipation of the upcoming vowel. In all three analysis windows, and for both groups of participants, the spectral mean was highest for the front vowel /ɛ/, lower for the central vowel /ʌ/, and lowest for the back vowel /ɔ/. The spectral mean is generally believed to correlate with the size of the cavity in front of an articulatory constriction or obstruction: the longer the anterior cavity, the lower the spectral mean (Feng et al., 2011; Nittrouer, 1995; Stevens, 1998; Tjaden, 2003). Thus, our acoustic results are compatible with a more anterior alveolar tongue tip position when /t/ is articulated before the front vowel /ɛ/ (shorter anterior cavity, higher spectral mean) and a more posterior alveolar tongue tip position when /t/ is articulated before the back vowel /ɔ/ (longer anterior cavity, lower spectral mean), with an intermediate position before the central vowel /ʌ/ (intermediate cavity length, intermediate spectral mean). Unfortunately, the link between acoustics and one or more specific vocal tract parameters remains unknown for the remaining three spectral coefficients (Feng et al., 2011).

Despite the facts that (a) the overall findings confirmed again that spectral coefficient analyses provide a sensitive technique for detecting anticipatory vowel-related effects in a preceding stop consonant, (b) the known relationship between a consonant's spectral mean and articulatory posturing fully supports an interpretation of these effects in terms of anticipatory coarticulation, and (c) all analyses were conducted with a very liberal significance threshold to maximize the probability of detecting any differences between stuttering and nonstuttering speakers, no such between-group effects were found for any of the four coefficients. There were no statistically significant results, or even trends toward statistical significance, for the Group main effect, the Group by Vowel interaction, or the Group by Vowel by Window interaction. Thus, based on the methods used here, the stuttering and nonstuttering speakers were indistinguishable not only in terms of the overall spectral characteristics of their voiceless alveolar stop consonant productions (Group analysis) but also in terms of their implementation of vowel-to-consonant anticipatory coarticulation (Group by Vowel and Group by Vowel by Window analyses).

We do recognize, of course, that word-level analyses of repeated simple, fluent C_1VC_2 utterances provide only limited information as compared with analyses of more complex target words or words that occurred in sentence-level productions or narratives, and further studies should include such utterances. As the present study was carried out only with one consonant across three vowel contexts, it provides no data regarding other initial consonants and following vowels. In addition, it is also possible that—similar to other subtle sensorimotor differences between stuttering and nonstuttering individuals—limitations in coarticulation may become observable in stuttering speakers only in conditions characterized by psychological or cognitive stress (see Caruso, Max, McClowry, & Chodzko-Zajko, 1998; van Lieshout, Ben-David, Lipski, & Namasivayam, 2014).

Nevertheless, combining our new results from spectral analyses of fluent single-word utterances in one specific phonetic context with others' findings for fluent speech analyzed in the acoustic domain with locus equations (e.g., Chang et al., 2002; Sussman et al., 2011) or in the kinematic domain with ultrasound imaging (Frisch et al., 2016), as well as formant transition analyses of dysfluent speech (Harrington, 1987; Howell & Vause, 1986), the evidence to date strongly suggests that anticipatory coarticulation is an aspect of speech motor control that does not differ in stuttering vs. nonstuttering individuals. Considered in a broader theoretical context, measures of anticipatory coarticulation quantify processes directly and specifically related to the relative timing of overlapping postures and movements (in the present study the onset and progression of vowel-related articulatory movements relative to those involved in creating the explosive burst for the preceding stop consonant). Thus, the presented data also contribute additional information to the long-standing debate about stuttering individuals' speech movement timing in general. As we have discussed in greater detail elsewhere, although the speech movements of stuttering individuals typically have longer durations than those of nonstuttering individuals (i.e., are slower), studies that investigated not just the duration from one event to another event but more direct measures of relative timing or rhythmic timing have usually failed to find statistically significant between-group differences (see Hilger, Zelaznik, & Smith, 2016; Max, 2004; Max & Gracco, 2005; Max & Yudman 2003a, 2003b). Our present coarticulation results are consistent with that extensive body of literature, and the integrated findings warrant a greater focus on theoretical perspectives that postulate a critical role for other aspects of speech sensorimotor control in the fluency breakdowns experienced by individuals who stutter (i.e., the weighting of feedforward vs. feedback control systems; insufficient priming of sensory systems during movement planning; learning stable sensorimotor representations or internal models of the vocal tract; etc.).

Acknowledgment

This research was supported in part by grants R01DC007603, R01DC014510, and P30DC004661 from the National Institute on Deafness and Other Communication Disorders, grant MOP-137001 from the Canadian Institutes of Health Research, and grant Y154221431 from the Institute of Acoustics, Chinese Academy of Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. The authors thank Caitlin J. Baldwin for data collection.

References

Bloodstein, O., & Bernstein Ratner, N. (2008). *A handbook on stuttering.* Clifton Park, NY: Delmar.

Caruso, A. J., Max, L., McClowry, M. T., & Chodzko-Zajko, W. J. (1998). Cognitive stress and stuttering: an experimental paradigm for connected speech. *Contemporary Issues in Communication Science and Disorders, 25,* 65-75.

Chang, S. E., Ohde, R. N., & Conture, E. G. (2002). Coarticulation and formant transition rate in young children who stutter. *Journal of Speech, Language, and Hearing Research, 45,* 676-688. Frisch, S. A., Maxfield, N., & Belmont, A. (2016). Anticipatory coarticulation and stability of speech in typically fluent speakers and people who stutter. *Clinical Linguistics & Phonetics, 25,*1-15.

Feng, Y., Hao, G. J., Xue, S. A., & Max, L. (2011). Detecting anticipatory effects in speech articulation by means of spectral coefficient analysis. *Speech Communication, 53*(6), 842-854

Harrington, J. (1987). Coarticulation and stuttering: an acoustic and electropalatographic study. In H.F.M. Peters & W. Hulstijn (Eds.), *Speech motor dynamics in stuttering* (pp. 381-392).New York: Springer-Verlag.

Howell, P., & Vause, L. (1986). Acoustic analysis and perception of vowels in stuttered speech. *Journal of the Acoustical Society of America, 79(5),* 1571-1579.

Hilger, A. I., Zelaznik, H., & Smith, A. (2016). Evidence that bimanual motor timing performance is not a significant factor in developmental stuttering. *Journal of Speech, Language and Hearing Research*, 1-12.

Kirk, R. E. (2012). *Experimental design: procedures for the behavioral sciences* (4th ed.). Thousand Oaks, CA: Sage.

Lindblom, B. (1963). *On vowel reduction.* Report No. 29, The Royal Institute of Technology, Speech Transmission Laboratory. Stockholm, Sweden.

Lindblom, B. (1998). An articulatory perspective on the locus equation. *Behavioral and Brain Sciences, 21*,274–275.

Löfqvist, A. (1999). Interarticulator phasing, locus equations, and degree of coarticulation. *Journal of the Acoustical Society of America, 106,* 2022-2030.

Max, L. (2004). Stuttering and internal models for sensorimotor control: a theoretical perspective to generate testable hypotheses. In B. Maassen, R. Kent, H.F.M. Peters et al. (Eds.), *Speech motor control in normal and disordered speech* (pp. 357-388), Oxford, UK: Oxford University Press.

Max, L., & Gracco, V. L. (2005). Coordination of oral and laryngeal movements in the perceptually fluent speech of adults who stutter. *Journal of Speech Language and hearing Research, 48(3),* 524-542.

Max, L., & Onghena, P. (1999). Some issues in the statistical analysis of completely randomized and repeated measures designs for speech, language, and hearing research. *Journal of Speech, Language, and Hearing Research, 42,* 261-270.

Montgomery, A. A., & Cooke, P. A. (1976). Perceptual and acoustic analysis of repetitions in stuttered speech. *Journal of Fluency Disorders, 9,* 317-330.

Nittrouer, S. (1995). Children learn separate aspects of speech production at different rates: evidence from spectral moments. *Journal of Acoustic Society of America, 97,* 520-530.

Riley, G. D. (1994). *Stuttering Severity Instrument for children and adults* (3rd ed.) (SSI-3). Austin, TX: Pro-Ed.

Stromsta, C. (1965). A spectrographic study of dysfluencies labelled as stuttering by parents. *De Therapia Vocis et Loquelae, 1,* 317-319

Stromsta, C. (1980). Physiological correlates of the core behaviour of stuttering. Proceedings of the XVIIIth Contress of the International Association of Logopedics and Phoniatrics (pp. 335-340). Washington DC.

Stromsta, C. (1986). *Elements of stuttering*. Oshtemo, MI: Atsmorts Publishing.

Stromsta, C. (1987). Acoustic and electrophysiologic correlates of stuttering and early developmental reactions. In H.F.M. Peters & W. Hulstijn (Eds.), *Speech motor dynamics in stuttering* (pp. 267-277). New York: Springer-Verlag.

Sussman, H. M., Byrd, C. T., & Guitar, B. (2011). The integrity of anticipatory coarticulation in fluent and non-fluent tokens of adults who stutter. *Clinical Linguistics and Phonetics, 25* (3), 169-186.

Sussman, H., McCaffrey, H., & Matthews, S. (1991). An investigation of locus equations as a source of relational invariance for stop place categorization. *Journal of the Acoustical Society of America, 90,*1309–1325.

Tabain, M. (2000). Coarticulation in CV syllables: a comparison of locus equation and EPG data. *Journal of Phonetics, 28*(2), 137-159.

Tabain, M. (2002).Voiceless consonants and locus equations: a comparison with electropalatographic data on coarticulation. *Phonetica, 59,* 20-37.

Tjaden, K. (2003). Anticipatory coarticulation in multiple sclerosis and Parkinson’s disease. *Journal of Speech, Language, and Hearing Research, 46,* 990-1008.

Tjaden, K., & Wilding, G. E. (2005). Effects of rate reduction and increased loudness on acoustic measures of anticipatory coarticulation in Multiple Sclerosis and Parkinson’s Disease. *Journal of Speech, Language and Hearing Research, 48,* 1-17.

van Lieshout, P., Ben-David, B., Lipski, M., & Namasivayam, A. (2014). The impact of threat and cognitive stress on speech motor control in people who stutter. *Journal of Fluency Disorders, 40,* 93-109.

Van Riper, C. (1971). *The nature of stuttering.* Englewood Cliffs, NJ: Prentice-Hall.

Van Riper, C. (1982). *The nature of stuttering* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Yaruss, J. S., & Conture, E. G. (1993). F2 transitions during sound/syllable repetitions of children who stutter and predictions of stuttering chronicity. *Journal of Speech, Language, and Hearing Research, 36(5),* 883-896.

Footnotes

1. This choice of analysis method was based on an incorrect assumption that "formant structure and harmonic structure of speech as displayed spectrographically are merely alternative ways of presenting the same information" (Stromsta, 1986, p. 68).

2. Anticipatory coarticulation refers very specifically to the phenomenon of articulatory movements being affected by the movements required for subsequent sounds. Many acoustic studies examining stuttering speakers' formant transition durations, extents, and rates have provided information about articulatory movements in general but not specifically about the degree to which movement characteristics associated with subsequent speech sounds were *anticipated* during the preceding speech sounds.

Table1. Individual participant characteristics for the stuttering group. Severity of stuttering is based on the Stuttering Severity Instrument for Children and Adults – Third Edition (SSI-3, Riley, 1994). Handedness is based on self-report. Hearing status was determined with pure tone behavioural testing at all octave frequencies from 250 to 4000 Hz for both ears separately (when a specific threshold is listed for a given frequency, all other thresholds were ≤ 25 dB HL; Le = left ear, Ri = right ear).

subject	gender	age	severity	handedness	hearing
SM01	male	30	very mild	right-handed	Le 4kHz 40 dB HL
SM02	male	49	very severe	right-handed	Ri 4kHz 40 dB HL
SM03	male	21	Mild	right-handed	all ≤ 25 dB HL
SM04	male	23	very severe	left-handed	all ≤ 25 dB HL
SM05	male	24	moderate	right-handed	all ≤ 25 dB HL
SM06	male	20	Mild	right-handed	all ≤ 25 dB HL
SM07	male	20	Mild	right-handed	all ≤ 25 dB HL
SF01	female	19	moderate	right-handed	all ≤ 25 dB HL

Table 2. Means and standard deviations (in parentheses) for all spectral coefficients (Mean, Standard Deviation (SD), Skewness, Kurtosis) for consonant /t/ in the target words “talk,” “tech,” “tuck.” Data are provided for three successive analysis windows with the first window centered on the acoustic burst and the remaining windows covering the initial portion of the aspiration noise. AWS, adults who stutter; AWNS, adults who do not stutter.

Coefficient	Window	Vowel	AWS	AWNS
Mean	Window1	/ɔ/	4380.49(600.88)	4749.75 (767.81)
		/ɛ/	4779.34(488.03)	5309.13(566.52)
		/ʌ/	4529.11(507.95)	5035.50(750.90)
	Window2	/ɔ/	5829.42(588.61)	6090.64(949.04)
		/ɛ/	5980.45(480.43)	6537.92(679.59)
		/ʌ/	5913.05(624.85)	6382.87(962.94)
	Window3	/ɔ/	5425.35(796.58)	5427.47(706.99)
		/ɛ/	5598.31(647.49)	5835.00(678.93)
		/ʌ/	5527.42(866.64)	5742.71(817.06)
SD	Window1	/ɔ/	1718.40(255.78)	1901.00(492.36)
		/ɛ/	1817.77(288.96)	1859.31(331.19)
		/ʌ/	1709.38(287.51)	1851.83(435.04)
	Window2	/ɔ/	2131.55(202.85)	2010.29(247.46)
		/ɛ/	1909.46(268.32)	1809.86(244.32)
		/ʌ/	2049.08(195.74)	1879.70(307.77)
	Window3	/ɔ/	2085.59(139.37)	2055.35(193.38)
		/ɛ/	1864.26(195.52)	1779.00(264.18)
		/ʌ/	2007.23(172.88)	1908.30(313.33)
Skewness	Window1	/ɔ/	1.587(0.618)	1.351(0.778)
		/ɛ/	0.926(0.285)	0.551(0.454)
		/ʌ/	1.367(0.406)	1.034(0.722)
	Window2	/ɔ/	0.684(0.346)	0.595(0.677)
		/ɛ/	0.788(0.381)	0.514(0.436)
		/ʌ/	0.706(0.353)	0.560(0.696)
	Window3	/ɔ/	0.806(0.364)	0.798(0.352)
		/ɛ/	0.968(0.310)	0.858(0.305)
		/ʌ/	0.835(0.348)	0.826(0.482)
Kurtosis	Window1	/ɔ/	5.637(3.570)	4.691(6.732)
		/ɛ/	2.653(1.369)	2.049(2.184)
		/ʌ/	4.416(2.106)	3.748(5.241)
	Window2	/ɔ/	0.079(0.855)	0.442(1.309)
		/ɛ/	0.761(1.068)	0.374(1.577)
		/ʌ/	0.206(0.802)	0.530(1.595)
	Window3	/ɔ/	0.419(0.924)	0.587(1.144)
		/ɛ/	0.978(1.000)	0.863(1.390)
		/ʌ/	0.449(1.054)	0.864(1.719)

Table 3. *F* values, uncorrected *p* values, and partial eta-squared effect sizes (η_p^2) for all inferential statistical analyses for all spectral coefficients (Mean, Standard Deviation, Skewness, Kurtosis).Window: three successive analysis windows with the first window centered on the acoustic burst and the remaining windows covering the initial portion of the aspiration noise. Vowel: /ɔ/, /ɛ/, /ʌ/in the target words “talk,” “tech,” “tuck.” Group: adults who stutter vs. adults who do not stutter.* $p \leq 0.05$, ** $p \leq 0.001$

Factor	**Spectral coefficient**	***df, F***	***p***	η_p^2
Window	Mean	(1.564, 21.901), 77.675	< 0.001**	.85
	Standard Deviation	(1.208, 16.910), 4.381	0.046*	.24
	Skewness	(1.427, 19.978), 10.020	0.002*	.42
	Kurtosis	(1.143, 16.000), 19.844	<0.001**	.59
Vowel	Mean	(1.823, 25.528), 20.704	<0.001**	.60
	Standard Deviation	(1.781, 24.930), 9.915	0.001*	.42
	Skewness	(1.367, 19.144), 4.189	0.044*	.23
	Kurtosis	(1.306, 18.285), 4.425	0.041*	.24
Group	Mean	(1,14), 1.183	0.295	.08
	Standard Deviation	(1,14), 0.054	0.820	.00
	Skewness	(1,14), 1.000	0.334	.07
	Kurtosis	(1,14), 0.031	0.863	.00
Group × Window	Mean	(1.564, 21.901), 1.305	0.284	.09
	Standard Deviation	(1.208, 16.910), 2.600	0.121	.16
	Skewness	(1.427, 19.978), 0.751	0.442	.05
	Kurtosis	(1.143, 16.000), 0.338	0.598	.02
Group × Vowel	Mean	(1.823, 25.528), 2.424	0.113	.15
	Standard Deviation	(1.781, 24.930), 0.488	0.598	.03
	Skewness	(1.367, 19.144), 0.517	0.537	.04
	Kurtosis	(1.306, 18.285), 0.349	0.619	.02
Vowel × Window	Mean	(3.876, 54.262), 2.499	0.055	.15
	Standard Deviation	(2.409, 33.725), 9.858	<0.001**	.41
	Skewness	(2.092, 29.288), 17.210	<0.001**	.55
	Kurtosis	(1.341, 18.772), 8.180	0.006*	.37
Group × Vowel × Window	Mean	(3.876, 54.262), 0.234	0.913	.02
	Standard Deviation	(2.409, 33.725), 0.995	0.393	.07
	Skewness	(2.092, 29.288), 0.080	0.930	.01
	Kurtosis	(1.317, 18.772), 0.197	0.734	.01

Figure 1

Stuttering speakers' (left) and nonstuttering speakers' (right) spectral coefficients (M, mean; SD, standard deviation; Skew, skewness; Kurt, Kurtosis) for the initial consonant /t/ in the target words "talk," "tech," and "tuck" (represented by different symbols). Data were extracted for three successive analysis windows (w1, w2, w3) with the first window centered on the acoustic burst and the remaining two windows covering the initial portion of the aspiration noise.