DOI: 10.17469/O2103AISV000020

GIOVANNA LENOCI

Anticipatory coarticulation in the speech of people who stutter

Stuttering is a fluency disorder that manifests itself through frequent interruptions in the smooth flow of speech. The disfluencies characterising stuttering strongly suggest the presence of breakdowns in the precisely timed and coordinated articulatory movements required for fluent speech. For these reasons, coarticulation has been one of the most studied aspects in relation to stuttering. Studies investigating coarticulation in speakers with impaired speech production have obvious importance for advancing our understanding of the disorder itself, which ultimately has implications for the diagnosis and treatment of speech impairments. The purpose of this chapter is to summarize the most interesting experimental results emerged from the acoustic and articulatory study of lingual coarticulation in the speech of people who stutter. The last section of this chapter is devoted to present an ultrasound tongue imaging study developed with a group of Italian stuttering children and a matched control group. Preliminary results suggest that lingual coarticulation in the fluent speech of children who stutter presents some differences compared to normally fluent children.

Key words: stuttering, speech motor control, coarticulation, second formant transitions, Locus Equations, Ultrasound Tongue Imaging.

1. *Introduction*

In everyday life, humans accomplish a great number of different and often complex motor activities in an automatic and apparently effortless trend.

The literature on motor control provides various definitions of 'motor skill', such as: "skill consists in the ability to bring about some end result with maximum certainty and minimum outlay of energy, or time and energy" (Guthrie, 1952: 17); "a skilled response is highly organized, both spatially and temporally. The central problem for skill learning is how such organization or pattering comes about" (in Kelso, 1997). Therefore, a 'skilled' motor activity requires practice, a high spatial and temporal organization, and effectiveness in terms of energy and time costs. This general description of motor skill fits very well with fluent speech production mechanisms in the adult speech.

The motor skills required to produce speech are among the most sophisticated learned by humans: it requires a remarkably complex combination of 'linguistic' and motor processes. Prior to the articulation of the speech sounds, it involves rapid interactions of processes related to the planning and formulation of the intended utterance[1]. From a motor control perspective, a large number of muscles are activated to move respiratory, phonatory and articulatory structures; the nervous system must generate sets of commands that must be coordinated in time and space for the appropriate sequences of muscle activation to occur, and for speech to be produced fluently.

As all motor skills, speech movements are not innate: they require practice for an extended period of time (up to adolescence) before they reach a certain level of skill, as in adults' speech (Smith, Goffman, 1998; Smith, Zelaznik, 2004). Learning to speak is a task of immense complexity: children must learn to control their breathing sufficiently to produce the subglottal pressure necessary for speech; they have to learn that consonants and vowels tend to alternate with one another although consonants can occur in clusters. Furthermore, children will learn to produce supra-segmental constraints of intonation and timing that govern phrases while, within words, children must learn not only the order of syllables and segments, but also the details of their temporal control (Hawkins, 1984).

As for other motor skills, speech movements are organized into coordinative structures[2] (Mac Neilage, Davis, 2000): they are characterized by a high spatial and temporal organization; they are very effective because they accomplish their goals (the acoustic targets) with a minimum of effort, 'if possible' (Lindblom, 1990) and in a relatively fast pace. Speech gestures are also naturally flexible to change, in fact they adapt themselves to their linguistic context. In short, producing speech is indeed a motor skill: the articulatory organization of the adult's speech is symptomatic of a high degree of motor ability, because the speech sounds are co-produced rapidly, efficiently and with a high degree of accuracy through the phenomenon of coarticulation.

Coarticulation is a term used to describe the ubiquitous overlapping of the articulatory movements associated with separate sound segments. One of the consequences of coarticulation is therefore that speech sounds vary according to the context in which they are produced, and to the nature of sounds which precede or follow them. Coarticulation effects are often described in terms of the direction and the extent of influence. Right-to-left or anticipatory coarticulation occurs when a speech sound is influenced by a following sound, while – if a sound shows influence of a preceding sound – this is called carry-over or perseverative (left-to-right) coarticulation. Carry-over effects are often attributed to inherent kinematic characteristics of the speech organs. Anticipatory coarticulatory effects are generally regarded to be a characteristic of a skilled speech behavior. At a cognitive level, anticipatory movements are the evidence of a universal tendency for the brain to 'scan ahead' of

[1] For example, according to Levelt, Roelofs & Meje (1999), after the conceptual preparation of the linguistic message, word generation proceeds through lexical selection, morphological and phonological encoding, phonetic encoding, and articulation itself (i.e., mapping abstract intended linguistic structure to dynamic sequences of movements).

[2] "Highly evolved task-specific ensembles of neuromuscular and skeletal components constrained to act as a single unit" (Kelso, 1998: 205).

time (Lashley, 1951), and it is suggested that such anticipation may be disrupted in many types of speech disorder, affecting normal speech motor control, such as stuttering (Hardcastle, Tjaden, 2008).

1.1 Stuttering as a limitation in the speech motor skills

Stuttering is a speech motor disorder that typically arises in the first childhood: for the persistent stuttering group, the mean age at onset is 35.14 months (Ambrose, Yairi, Loucks, Seery & Throneburg, 2015). According to recent epidemiological estimates, stuttering is characterized by a high rate of spontaneous recovery: around 90% of stuttering children recover without therapeutic treatment by the fourth year from the onset of the disorder (Yairi, Ambrose, 2013). The phenomenon of recovery gives rise to important questions pertaining the differences between persistent and recovered stuttering[3] and prognosis. For this reason, recent research on stuttering focuses on the possibilities to discriminate, as soon as possible, children who will recover spontaneously from children who will become persistent, and consequently to determine if they exhibit different speech and/or non-speech characteristics even before the different developmental processes separate them (i.e., recovered vs persistent stuttering). Early prediction of the eventual course of the disorder will allow clinicians to make informed decisions about selective treatment strategies, for example, reserving immediate clinical intervention to children showing high chances of chronicity. This approach could increase the chance to recover from stuttering symptoms.

Developmental stuttering is characterized by disruptions in the production of speech sounds, also called Stuttering-Like Disfluencies (Yairi, Ambrose, 2005): monosyllabic-word repetitions, part-word repetitions, silent and audible sounds prolongations are the hallmark characteristics of the disorder.

Recent accounts of stuttering agree in defining it as a multifactorial disorder: many variables – such as language, motor, cognitive, emotional and genetic factors – are supposed to interact in complex ways in the development of the disorder and in the overt breakdowns in speech motor control that are perceived as stuttering-like disfluencies. The importance of the interactions of two of these factors, namely language and speech motor processes, is supported by a number of experimental findings according to which, for example, increases in utterances length and syntactic complexity are associated with the increased occurrence of stuttering-like disfluencies in children and adult who stutter (Buhr, Zebrowski, 2009; MacPherson, Smith, 2013). It is well known that stuttering is associated with reduced motor speech performance, and it is also true that moments of linguistic complexity tend to highly correlate with motor speech complexity. In studies where group differences emerged (Kleinow, Smith, 2000), people who stutter (henceforth, PWS) were found to show poorer performances on the timing and coordination of motor events compared to

[3] Separating the two sub-groups, should increase precision of experiments in various aspects of the disorder and provide evidence-based data to re-consider traditional view of stuttering as a unitary disorder (Subramanian, Yairi &Amir, 2003)

normally fluent speakers for a given speech task. In other words, even if most theories of the causes of stuttering postulate that many factors are involved in producing these motor breakdowns, it is clear that abnormal speech motor output is an essential component of stuttering (Olander, Smith & Zelaznik 2010).

Van Lieshout and co-workers (Van Lieshout 1995; Van Lieshout, Hulstijn & Peters 2004) proposed the Speech Motor Skills approach to explain motor aspects in PWS: speech production is a motor skill similar to any other (fine) motor skill and stuttering may arise from limitation in speech motor skills. PWS are located more toward the lower end of a presumed normal speech motor skill continuum, with people who do not stutter (henceforth PWNS) distributed across the more skilled end. From this perspective, stuttering is not viewed as a motor disorder such as dysarthria or dyspraxia, but rather as a reflection of an innate limitation of speech motor control system to prepare and perform complex motor actions in the presence of cognitive, linguistic, emotional and speech motor influences (Namasivayam, Van Lieshout, 2011).

Thus, in PWS the speech motor control system has been argued to be the critical weak link in the chain of events that lead to the production of speech. Therefore, stuttering-like disfluencies are the direct manifestation of failures of the speech motor system to address the appropriate command signals that drive the muscles involved in speech production.

Coarticulation has been one of the most studied motor aspects in relation to stuttering speech: it is a crucial mechanism for fluency, referring to the neuromuscular organization that mediates the complex and precise movements involved in the speech production. Stuttering has been hypothesized to stem from breakdown in coarticulation or difficulty in transitioning between sounds. Past research on stuttering tried to investigate coarticulatory processes in the stuttering speech using different instrumental techniques. Findings from studies investigating CV coarticulation in PWS are equivocal, especially because they vary in terms of the age group of interest, the methodology or the measures used to infer coarticulation, and the speech samples – which in the case of stuttering means whether perceptually fluent or dysfluent tokens were of interest. However, these results seem to confirm that the lingual coarticulation that accompanies a stuttering-like disfluency clearly differs from the coarticulation that characterizes normal fluency[4].

Due to space reasons, this paper will present only experimental results obtained from the acoustic and ultrasound tongue imaging data collected from the literature on the fluent speech of PWS. The rationale for studying coarticulatory patterns in the fluent productions of PWS is that differences, however subtle, between the perceptually fluent speech of PWS and normally fluent speakers may provide insight into the (disordered) speech motor control strategies of the former group. Furthermore, the more general implication is to establish that PWS are speaking abnormally even when they are not stuttering at all.

[4] The bulk of studies focusing on dysfluent utterances of PWS suggests atypical or absent F2 transitions (Harrington, 1987; Yaruss, Conture, 1993). Thus, coarticulation in dysfluencies of PWS appears to differ from normally fluent speech, at least as inferred from F2 transition characteristics.

2. *Acoustic measurements of coarticulation*

Many researchers have used relatively indirect techniques, such as acoustic analysis mainly because of the practical difficulties associated with articulatory tracking techniques.

However, acoustic analysis is still a valuable tool for exploring general contextual effects that occur as a result of coarticulatory processes. For example, second formant transitions (F2) allows for inferences to be made concerning lingual position and movement during speech production.

One method of estimating formant transition characteristics involves the use of visual criteria to determine the onset and the offset of the formant transitions, i.e., the first glottal pulse in the target vowel and the point in which a maximally steady-state is visually identified in the vowel formants on a wideband spectrograms. In this way, the entire formant transition can be measured in terms of duration[5] and frequency extent[6] (i.e., amplitude). The characteristics of these formant transitions (i.e., duration and extent) provide us the possibility to calculate the slope or 'trajectory' for the transition, and this measure allows to assess the articulatory gestures underlying coarticulation. Due to the temporal and positional aspects involved in the coarticulation of speech sounds, the slope coefficient can be regarded as 'an ordinal index of the rate of change in vocal tract geometry' (Weismer, 1991). The rate of frequency change in F2 transitions, or the speed with which formant frequencies changed during the transition, was estimated by calculating the absolute value of the extent of F2 transition (Hz) divided by the duration of the transition (msec). This measure is believed to approximate the speed with which speech articulators move from one location to the next (Yaruss, Conture, 1993). Accordingly, a large slope coefficient would reflect considerable positional and temporal movement of the tongue body inside the oral cavity following consonant release (Weismer, 1991).

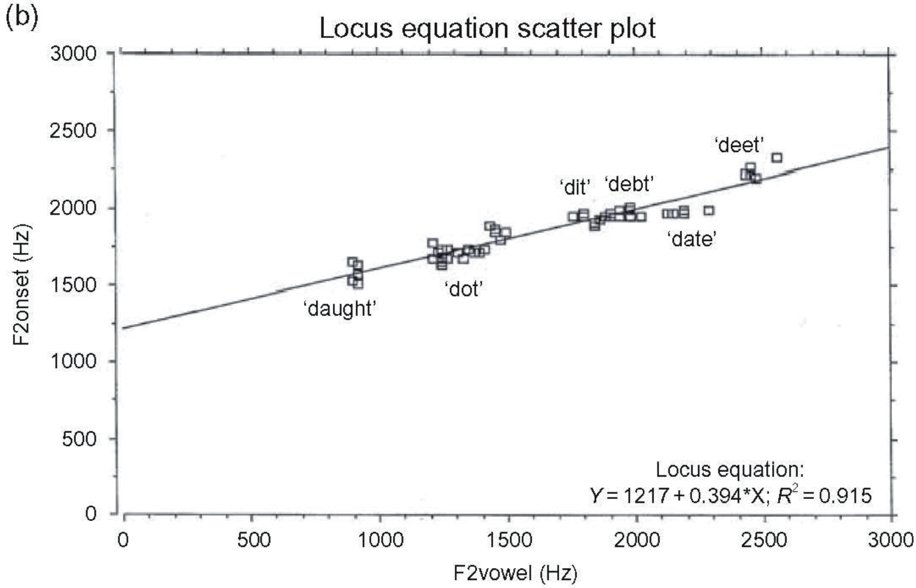
Another possibility to assess anticipatory CV coarticulation is Locus Equations metric (henceforth, LE). LEs are linear regressions of the frequency of the F2 transition sampled at its onset (on the first glottal pulse of the vowel) on the frequency of F2 when measured in the vowel nucleus (at a so-called steady-state location) (Lindblom, 1963). These frequency values are plotted for a single consonant produced with a wide range of following vowels: $F2_{onsets}$ are plotted along the y-axis and $F2_{midpoints}$ along the x-axis (cfr. Figure 1). For a given stop place category, data coordinates have been consistently shown to tightly cluster in a positively correlated distribution (Sussman, Hoemeke & McCaffrey, 1992; Sussman, Hoemeke & Ahmed, 1993), and the slope

[5] The duration of the formant transition can be estimated by calculating the difference in time (in msec) between the onset of the F2 transition at the beginning of formant movement from the initial sound, and the offset of the F2 transition at the beginning of the steady-state portion of the following sound (e.g., the vowel in CV utterances). This measure is believed to assess the amount of time the articulators spend moving from one position to another during the transition from one sound to the following one (Yaruss, Conture 1993).

[6] This measure can be estimated by calculating differences between the offset and the onset center frequencies of the transition, and is believed to approximate the overall movement of the articulators during the transition (Yaruss, Conture 1993).

of the linear regression line is said to be linked to the degree of coarticulation (Krull, 1988). The LE slope was shown to vary between the extremes of k = 0 and k = 1.0: flatter regression slopes indicated limited anticipatory coarticulation, as the C onset remains constant across varying vowel contexts, while steeper LE slopes indicated increasing extents of anticipatory coarticulation, with maximal coarticulation having $F2_{onsets}$ varying as a direct and linear function of the following vowel.

Figure 1 - *A representative locus equation scatter plot for 50 [dVt] tokens produced across 10 vowel contexts. This scatter plot is fit with a linear regression line, of the form F2onset = k*F2vowel + c, where k and c are respectively the slope and y-intercept of the line (adapted from Sussman et al., 2010: 3)*



The next section will summarise the main acoustic findings concerning the study of anticipatory coarticulation in the fluent speech of PWS.

The discovery of abnormal coarticulatory patterns in the perceptively fluent productions of PWS would corroborate the idea of a limitation in the speech motor system, which is instable and prone to interferences. In fact, to overcome interferences and achieve fluency, it is a common opinion that PWS have to adopt different articulatory strategies to accomplish the planned acoustic target (Kleinow, Smith, 2000; Van Lieshout *et al.*, 2004).

3. *Analyses of F2 transitions in the stuttering speech*

Literature on coarticulation in the stuttering speech has varied with regard to the analysis methods and speech samples; therefore, it is not surprising that a consistent picture has yet to emerge as to whether coarticulation is deviant for PWS (Hardcastle, Tjaden, 2008).

Robb, Blomgren (1997), for example, analyze CV lingual coarticulation in the speech of 5 adults who stutter (henceforth, AWS) and 5 adults who do not (AWNS). The hypothesis tested was that the perceptively fluent speech of the first group differs from AWNS in the slope of F2 transitions as a result of abnormal lingual coarticulation behavior. Acoustic results show that the F2 transitions in the stuttering group were characterized by greater frequency extents compared to the control group. These results show that the regulation of transitioning from different vocal tract configurations were handled differently: the steeper slopes found in the fluent speech of AWS would indicate greater and quicker movement of the tongue body within the oral cavity and hence a lower degree of coarticulation compared to non-stuttering speakers.

A more recent study on Farsi-speaking AWS (Dheqan, Yadegari, Blomgren & Scherer, 2016) aims to validate previous results: authors suggest that a limitation in the generalization of acoustic studies on stuttering is due to the fact that most of them have been conducted on English speakers. For these reasons, the same acoustic parameter (F2 transition slope[7]) and the same stimuli used by Robb & Blomgren were used to infer anticipatory coarticulation in the fluent speech of their subjects. Other F2 transitions features were investigated to compare articulatory dynamics: the overall frequency extent, the overall duration of the transition and the speech rate. The participants of the study were 10 Iranian AWS and 10 AWNS. The findings revealed significant differences in a number of measures between stuttering and non-stuttering speakers. Compared to control group, perceptually fluent utterances of PWS were characterized by greater F2 frequency extents during transitions, longer F2 transition which took more time to reach vowel steady state and slower speaking rate. Concerning the transition slope, results showed no differences in the overall F2 slope between the two groups of speakers but, when the comparison window was constrained to the beginning of the transition (i.e., on the first 30 ms), differences were identified. Stuttering speakers, in fact, exhibited greater absolute initial F2 slope compared to control group. Overall, these findings corroborate previous results found in the literature on stuttering (Robb, Blomgren, 1997), and suggest that PWS adopt different articulatory strategies to reach fluency, compared to normally fluent speakers. The longer F2 transitions (and the lower speaking rate) suggest that PWS need a lengthening of time to complete an articulatory movement to an intended target; the greater overall F2 transition frequency extent (or change) suggests that stuttering speakers had more lingual movement (or displacement) during the transition compared to control group. Authors conclude that there are both spatial and temporal features of F2 transitions in stuttering speakers that need to be better understood.

Sussman *et al.* (2010) focus on anticipatory coarticulation using LE metric in the fluent and disfluent productions of stop + vowel sequences in 8 PWS and 8 PWNS. Linear regression functions were performed separately for the fluent and disfluent utterances. Results show that PWS slope values, both for fluent and for dysfluent slope

[7] F2 transition slopes were calculated using the fixed-time point method (Nearey, Shammass, 1987): the offset formant frequencies were specified at the distinct time-points of 30 ms and 60 ms from the onset of the transition (corresponding to the first glottal pulse of the vowel following the C).

tokens, fall within the normative ranges. Despite this, PWS show more variation in repeated productions of the target stimuli, as demonstrated by Standard Error Estimate values collected from LE analysis. Authors suggest that this argues against 'any serious deficits in the motor planning/execution of stop + vowel anticipatory coarticulation in PWS' (Sussman *et al.*, 2010: 12). It is important to underline that, in this study, the more severe stutterers showed, both for fluent and for dysfluent utterances, the most distant slope values from the control group speakers. This means a lower degree of coarticulation compared to control group and to moderate/mild stutterers.

Therefore, the picture emerging from the study of a possible inefficiency in the stutterers' ability to plan and execute the proper degree of anticipatory coarticulation is still controversial.

A further step towards a deeper understanding of speech dynamics responsible for stuttering is to assess the speech of children close to the onset of stuttering[8]. Studying the speech of children who stutter (henceforth, CWS) allows to discriminate the core characteristics of the disorder from possible articulatory strategies that AWS could adopt to overcome interruptions in the smooth flow of speech.

Chang, Ohde & Conture (2002), for example, assess anticipatory coarticulation and Formant Transition Rate (FTR) on the fluent speech of 14 CWS and 14 non-stuttering children who do not stutter (henceforth, CWNS). Only fluent utterances were selected for the acoustic analyses. The degree of anticipatory coarticulation was evaluated through the LE metric, while the speed (the velocity) at which the tongue moves from one position in the oral cavity to another was assessed with FTR measurement. Results show that while CWS and CWNS did not differ in terms of degree of coarticulation (as measured by the LEs slope and y-intercept), they did differ in FTR measures. CWS differentiated FTR less than CWNS for place of articulation. This result indicates that CWS were slower than CWNS in executing anticipatory movement of articulators for the vowel during consonantal production. Therefore, authors suggest that there is a kinematic-based difference (speed of movement) between stuttering and control groups, rather than a direct coarticulation-based difference.

Another important rational for studying CWS's speech is that much of the actual research on developmental stuttering is focusing at the identification of clinical predictors of chronicity. One of the phonetic indexes proposed as a potential marker of chronicity was F2 transition. Stromsta (1965), for example, report an early longitudinal study in which F2 transitions were analyzed in 63 children identified by their parents as having stuttering. Disfluent segments were analyzed and results reveal that speech disfluencies characterized by abnormal formant transitions and abnormal terminations of phonation were found in the speech of those children whose stuttering became persistent, while children exhibiting normal formant tran-

[8] "coarticulatory behaviours of adults who do and do not stutter are not readily generalizable to the coarticulatory behaviours of children who stutter and children who do not stutter, due to many developmental differences in speech/language production between adults and children" (Chang *et al.*, 2002: 677).

sitions recovered. Unfortunately, details about this study are scarce and is difficult to assess the generality of its results.

Some years later, Yaruss, Conture (1993) investigate F2 transition differences among 7 young children considered at 'low-risk' of persistence and 6 children regarded as 'high-risk' for chronic stuttering. Several measures of F2 transition, such as the duration, the extent and the transition rate of F2 in sound/syllable repetitions were made in comparing the extra disfluent segment with the fluent segment (e.g., b-but; a-and). The authors show that "children who stutter do produce missing (25-29%) or atypical (10-16%) formant transitions during the first iteration of their sound/syllable repetition" (p. 893), but the presence of abnormal F2 transitions was not sufficient to differentiate the two groups. A critical issue of the study is that the validity of the classification of a child as 'high' or 'low' risk of chronicity was not verified through longitudinal observations.

A more recent study, realized by Subramanian *et al.* (2003), investigate the predictive value of F2 transitions as an early marker of chronic stuttering in children close to the onset of the disorder. Twenty CWS and 10 CWNS were audio recorded during the initial visit, when the eventual classification of a child as 'persistent' or 'recovered' was unknown. The final status of each child was evaluated after a minimum of 36 month post-onset of the disorder.

The acoustic parameters considered were the frequency change and the duration of F2 transitions; to allow the comparison with the control group, only fluent utterances were selected for the analysis.

Results suggest that "the frequency dimension of the formant transition, rather than the time dimension, is the most significant contributor to the differences between stuttering and non-stuttering children as well as between the two stuttering subgroups" (Subramanian *et al.*, 2003: 70). As we can see, in fact, in Table 1, mean values for the duration measures did not differ between the three groups, while more interesting findings emerged for the frequency change measure.

Table 1 - *Means and standard deviations (in parentheses) of the three groups for frequency change (Hz) and duration (ms) (from Subramanian et al., 2003: 68)*

Measure	*Persistent*	*Recovered*	*Controls*
Duration	58.75 (22.72)	59.07 (19.61)	66.14 (37.12)
Frequency change	395.78 (196.72)	583.99 (229.77)	502.35 (232.10)

The persistent group shows a smaller frequency change compared to recovered and control groups. This result may be interpreted to reflect restricted spatial movement of the tongue from one position to another. It seems that, to reach fluency, children with persistent stuttering have to undershoot their articulatory target, showing a higher degree of coarticulation. Authors conclude that F2 transitions could be a good marker of persistent stuttering in children close to the onset of the disorder, because results showed smaller frequency changes (reduced amplitude of the articulatory gestures) for those children whose stuttering become chronic, compared to children whose stuttering recovered.

We will focus now on an Italian project named 'Phonetic predictive indexes of chronic stuttering in preschool age children'[9]. The aim of this project is to test the prognostic value of a set of clinical predictors proposed as good markers of persistent stuttering: the Disfluencies Profile (Zmarich, Bernardini, Lenoci, Natarelli, Pisciotta, *to appear*) and the speech-associated attitudes of preschool and kindergarten children, as measured by kiddyCAT (Vanrykeghem, Brutten, 2007). Adding to these more clinical markers, the degree of CV coarticulation was also investigated.

The 13 CWS enrolled in the study were audio-recorded from the stuttering onset stage to 16-22 month post-onset. The validation of the prognosis (persistent vs recovered stuttering) was evaluated through structured telephone interviews with the parents after 3 and 4 years from the onset. A group of 26 CWNS, matched for age and sex, served as control group.

In Lenoci (2015) acoustic data are presented for a group of 5 stuttering children (3 persistent vs 2 recovered stutters), and results were compared with those of 26 non-stuttering children, selected in order to be matched for sex and age with the former group. The author collected spontaneous speech recordings at 3 different main stages: onset of the disorder, second and third semester post-onset. The last two stages were used to investigate the prognostic value of F2 transitions to discriminate persistent vs recovered stutterers. Only fluent CV and CVC sequences, containing bilabial, alveolar and velar stops in any vowel context were selected for the acoustic analyses. The metric used to investigate anticipatory coarticulation was Locus Equations. Slope values indexing anticipatory coarticulation were used for comparison analysis among the three groups (persistent stutterers, recovered stutterers and control group). Results show that for the bilabial and alveolar place of articulation, 2 of the children who later developed persistent stuttering presented lower slope values compared to the other two groups. This tendency has been observed through the three stages analysed altogether (from the onset of stuttering to the third semester post-onset). Children who later recover present slope values very similar to control group's ones for the three places of articulation. Velar slope values show less differences between the three groups, even if the more severe stuttering children present a lower degree of coarticulation compared to control group. These preliminary results suggest that speech gestures are poorly coordinated and performed with higher amplitude in the speech of children with chronic stuttering. Even if the small sample of subjects does not allow for prognostic inferences, these preliminary results corroborate previous findings from the pertinent literature (Robb, Blomgren, 1993), according to which PWS have a lower degree of coarticulation compared to control group. These results can also be interpreted with the proposal made by Van Lieshout *et al.* (2004), according to which PWS show a stronger reliance on the kinesthetic feedback compared to normal speakers, in

[9] This is a longitudinal project granted to Claudio Zmarich by the CNR in 2008 (Lenoci, Allegri, Bernardini, Chiari, Crivelli, Dadamo, de Biase, Galatà, V., Pisciotta, Polesel, Stanchina, Stocco, Vayra, Zmarich, 2012).

order to maintain stability in the speech motor control. To increase the feedback gain, probably, CWS need to increase the range of articulatory movements and this means to perform speech gestures with a low degree of coarticulation.

Despite all, acoustic results regarding coarticulation in the speech of CWS seem to be still controversial even if they provide the rationale for further investigations. For this reason, in the next section we will focus on more direct measures of speech dynamics, such as Ultrasound Imaging of the Tongue[10].

4. *Ultrasound Tongue Imaging studies on coarticulation in the stuttering speech*

Ultrasound Tongue Imaging (henceforth, UTI) has been increasingly used in speech sciences research over the last few decades and it offers a direct representation of tongue movements in speech. It is a safe and non-invasive articulatory technique, providing information about the shape and the position of the main articulator involved in the production of consonants and vowels, the tongue. Such direct measures allow to get a deeper insight into lingual articulation dynamics that are not revealed through acoustic measures. Furthermore, particular attention has been devoted to the potential power of UTI as a biofeedback tool for modifying atypical articulations in speech disordered speakers (Cleland, Scobbie & Zharkova, 2016).

So far, two ultrasound studies were realized on the speech of adults who stutter (Heyde, Scobbie, Lickley & Drake, 2015; Frisch, Maxfield & Belmont, 2016). The first study aims at corroborate the hypothesis that PWS struggle not when initiating the absolute syllable initial consonant but when transitioning from the C into the following vowel (Wingate, 1988). Fluent productions of CV syllables (C = /k/; V = /a, i, ə/) from 3 PWS and 3 PWNS were analysed for duration and peak velocity relative to articulatory movement towards (onset) and away (offset) from the consonantal closure.

Measures of displacement and velocity were collected at the point of maximum displacement of the tongue surface and results show that while the two groups had comparable onset behaviours, they do differ in offset peak velocity. PWS, in fact, displayed lower velocity at reaching the vowel target compared to controls. According to authors, results confirm the hypothesis according to which coarticulation from a sound to the succeeding one is impaired in the stuttering speech and that this could be an indicator for an underlying motor control impairment.

Frisch *et al.* (2016)'s study was specifically devoted to investigate anticipatory coarticulation in the fluent speech of 23 AWS and 23 AWNS. The study examines the coarticulation of anticipatory velar-vowel sequences in the adjustment of velar closure location for /k/[11] depending on the following vowel context (the nine Standard American English vowels /i e æ ʌ ɝ ɑ ɔ o u/). For the purposes of the study,

[10] The advantage of ultrasound over acoustic analysis is that the last one provides only indirect evidence of articulatory movements, posing some issues (for example in the detection of formants) especially for children acoustic analysis.

[11] "the target phoneme in this study was /k/, which is known to have relatively large variation in production across contexts" (p. 288).

authors selected for each velar-vowel production the UTI frames displaying maximum velar closure for subsequent analysis. Coarticulation was determined through curve-to-curve distance[12] comparison between tongue contours across the variety of all vowel contexts for each speaker following Zharkova, Hewlett & Hardcastle (2012). Differences among tokens (18 monosyllabic CV or CVC sequences) were measured, and average measures of coarticulation obtained from each speaker were then used in statistical analysis to compare overall patterns of coarticulation.

Results show that curve-to-curve distances between front and back vowel contexts did not differ for PWS and PWNS. This means that the two groups show identical patterns of coarticulation. In line with previous findings from literature (Sussman *et al.*, 2010; Smith, Sadagopan, Walsh & Weber-Fox 2010), PWS of this study were found to be more variable than PWNS in the production of the same articulatory target (i.e., the same velar-vowel sequence). Authors conclude that PWS's stuttering-like disfluencies are not attributable to immature motor planning as measured by anticipatory coarticulation (Frisch *et al.*, 2016).

4.1 An UTI study of lingual coarticulation in the speech of Italian children who stutter

An ongoing project, developed in the Laboratorio di Linguistica of Scuola Normale Superiore, focuses on studying the speech of CWS. The aim of the project is to assess the presence of differential articulatory patterns between a group of Italian stuttering children and a matched control group. The motor aspects under investigation are those underlying the anticipatory coarticulation and the stability of movements through multiple repetitions of the same item.

So far, 10 school age children (age range from 6 to 12 years) were recruited for the study: 5 CWS[13] (3 males and 2 females) and 5 CWNS, balanced for age and sex. Subjects were recorded by means of an ultrasound system (*Mindray UTI system-30 Hz*) with a microconvex probe (*Mindary probe 65EC10EA*); the UTI images were synchronized with the audio through a synchronisation unit (*Synch Bright-up unit*). In a child-friendly set up, participants were seated comfortably in a chair and the ultrasound probe was held under the chin with the head stabilization unit (*Articulate stabilisation headset*, Articulate Instruments ltd) for stabilizing the position of the transducer with respect to the head. The experimental task consisted in the production of /CV/ sequences with the consonant C corresponding to the bilabial stop /b/, alveolar /d/ and velar /g/ and the vowel V corresponding to the high front /i/, low /a/ and high back /u/. The three cardinal vowels allowed for testing diverging tongue positions. The target syllables were embedded in dysillabic pseudo-words of the type /ˈCVba/. Bilabials following the

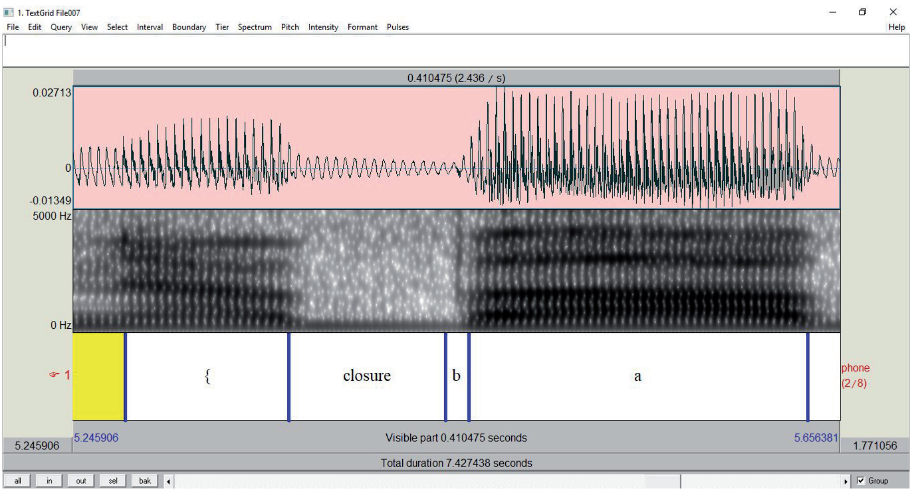
[12] "Mean distance in midsagittal tongue surface outline between tokens of the same phoneme across two different environments was taken as a measure of the phoneme's susceptibility to environment influence" (Zharkova, Hewlett, 2009: 3).

[13] Parents of stuttering children reported a history of developmental stuttering with no other language or hearing disorders. The presence of stuttering was evaluated from a speech pathologist and all stuttering children had begun therapeutic treatment in the last few years. Parents of the typically fluent group reported no history of speech, language and hearing disorders.

target syllables were used to eliminate the influence of additional lingual coarticulation within pseudo-words. Twelve repetitions of each CV sequence, embedded in short carrier phrases (ex. 'La gattina DUba salirà' the little cat DUba will go up), were collected in random order, for a total of 108 utterances for each participant. Here we will present only preliminary data for the coarticulation during the fluent speech[14] of one CWS and one CWNS. To achieve the goal we transposed measures of LE to the articulatory domain in line with previous recent studies (Noiray, Menard & Iskarous, 2013). The adaptation of LE to the articulatory domain was conducted on the lingual data simultaneously recorded with the acoustic speech signal. Instead of F2 transitions, we used the horizontal position of the highest point of the tongue at the mid-point of the consonantal closure (dependent variable) and at the mid-point of the following vowel (independent variable)[15]. We chose the mid-point of the consonant closure as dependent variable to ensure that what we were measuring was indeed 'coarticulation', as opposed to being part of the 'transition' from the C to the V articulation.

First, acoustic data were phonetically segmented and labelled using PRAAT. Three intervals were selected on the spectrograms of each CV sequences (see Figure 2): the consonantal closure (from the offset of the preceding vowel to the burst of the stop); the VOT of the C (from the burst to the first glottal pulse of the following vowel); the V-target (from the first glottal pulse to the offset of the vowel: see Figure 1).

Figure 2 - *Example of a PRAAT window, with the waveform, the spectrogram and the textgrid containing the segmentation and labelling of a [ba] sequence*



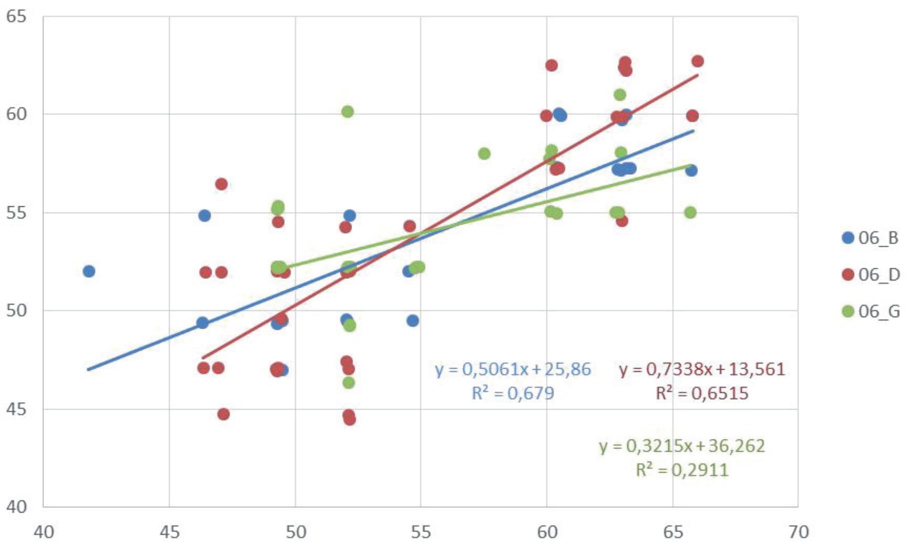
[14] CWS's speech was judged as perceptually fluent on the basis of audible acoustic data. The absence of syllable repetitions, blocks and sound prolongations (cfr. Stuttering-like disfluencies) was used as criteria to judge the speech as perceptually fluent.

[15] Since several studies have associated the frontness of the body of the tongue in the vocal tract to F2, Iskarous, Fowler & Whalen (2010) measured, with EMMA, the horizontal position of the tongue body at the release of consonant and in the middle of the vowel, in order to investigate the articulatory origins of LE. Authors found the same linear relations present in the acoustic domain.

The annotations were imported in the software used for the articulatory analysis (*Articulate Assistant Advanced* - Articulate Instruments ltd) for the selection of the ultrasound relevant frames within each C-V intervals. A semi-automatic tongue contour splining was performed in the acoustic interval spanning from the beginning of the consonantal closure to the end of the following vowel.

As already described, relevant ultrasound frames were selected for each CV sequence at two time points: the midpoint of the consonantal closure and the midpoint of the vowel. Each tongue curve (spline) was converted in terms of x,y coordinates and the x value corresponding to the highest y point on each curve was selected for subsequent linear regression analysis. Linear regression functions were calculated for each place of articulation and for each child[16]. Figures below display the regression lines obtained from the fluent utterances of one CWNS (Figure 3) and one CWS (Figure 4).

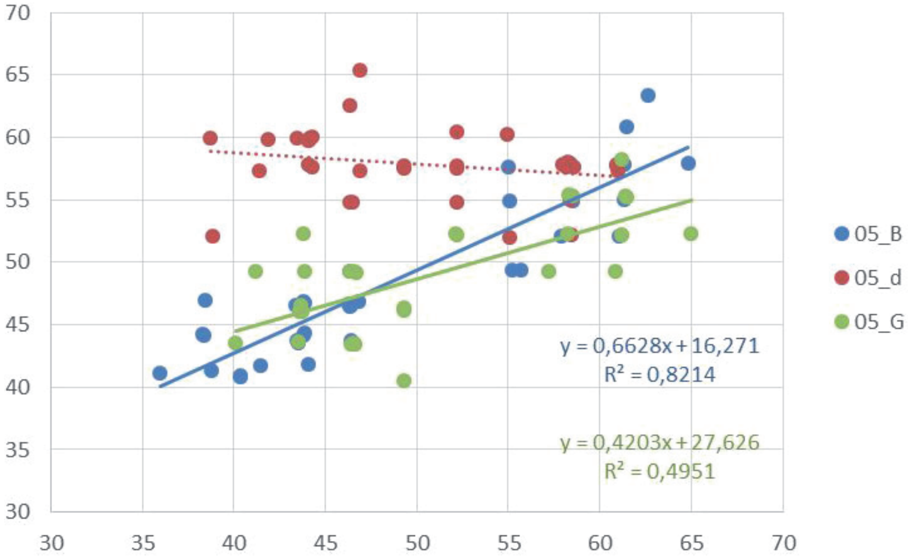
Figure 3 - *Linear regressions between consonant-closure midpoint and vowel midpoint for one child who do not stutter. Slopes are b: 0.50, d: 0.73, g: 0.29 respectively for the bilabial (blue), alveolar (red) and velar (green) place of articulation*



For CWNS we can observe that, for each place of articulation, data points tightly cluster across the regression lines. This means that, in line with previous literature, the linearity of LE originates in linearity in articulation between the horizontal position of the tongue dorsum in the consonant and to the horizontal position of the tongue dorsum in the vowel (Iskarous *et al.*, 2010).

[16] The experimental procedure and the statistical analyses were developed with the technical support of Irene Ricci, from the Laboratorio di Linguistica of Scuola Normale Superiore.

Figure 4 - *Linear regressions between consonant-closure midpoint and vowel midpoint for one stuttering child. Slopes are b: 0.66, g: 0.42 respectively for the bilabial (blue), alveolar (red) and velar (green) place of articulation*



As for the CWS, we can observe that for the bilabial and velar place of articulations this CWS shows slightly higher slope values compared to the CWNS. For the alveolar consonants instead, the regression line did not fit the data (red dots). This means that for the CWS there is no interaction between the position of tongue dorsum at the mid-point of the consonantal closure and the midpoint of the following vowel. Consequently, in terms of LEs metric, there is no anticipatory effect in the production of /dV/ sequences. Interestingly enough, a linearity emerged when we used a different time point as dependent variable: the consonantal offset[17] instead of the mid-point of the closure. According to Zharkova *et al.* (2012), a possibility to measure the extent to which a speech sound varies systematically according to the identity of the following one, and to compare the effects between speakers, is to compare the size of any coarticulatory effect at a selected time point. The result obtained in our study could mean that anticipation of the upcoming V gesture starts later in the CWS compared to CWNS. Another aspect to underline is that even though a coarticulatory effect emerged for the stuttering child, the slope value was considerably lower compared to the normally fluent child. These preliminary results suggest that, even if no significant differences emerge for the degree of coarticulation of /bV/ and /gV/ sequences between the two children, some differences emerged when the synergistic use of different part of the same articulator (the tip and the body of the tongue) is required. Our opinion is that, this result can be inter-

[17] We selected as dependent variable the horizontal position of the tongue on the UTI frame included between the burst and first glottal pulse of the vowel.

preted as a less mature motor control system for the CWS who need more time to reach the articulatory target for the C and, consequently, cannot anticipate earlier the vowel gesture.

5. *Conclusion*

A still controversial picture emerges from this review on the acoustic and articulatory studies investigating coarticulation on the speech of adults and children who stutter. The ambiguous results may be due to the different age samples, speech samples and methodologies used across studies here reviewed. Overall, most studies have shown that PWS differ in respect to coarticulatory dynamics compared to PWNS, even during their perceptually fluent speech. Sometimes, however, PWS and PWNS seem to be very similar, showing only extremely subtle differences, and these results corroborate Van Lieshout *et al.* (2004)'s motor skill perspective, according to which there is an individual variability along a continuum for *all speakers*, and PWS stay in the low end of that continuum. Anyway, further investigations especially on CWS are needed in order to both assess the real manifestations of the disorder, and to equip clinicians with the instruments for a better rehabilitation. Ultrasound Tongue Imaging analysis is a valuable tool for both these purposes. Preliminary data on two Italian children show, for example, that some differences can characterize the fluent speech of the stuttering child: for the alveolar stops, for example, we observed a lower degree of coarticulation, which emerges later compared to the control peer. The lower degree of coarticulation suggests that for the stuttering child, the alveolar consonants adapt to the vowels less than the control child. A possible interpretation is that the CWS have not learned how to achieve the articulatory differentiation between two parts of the same articulator, the tip and the body of the tongue, and how to coordinate them in the proper way, as the normally fluent peer does. Furthermore, this articulatory result corroborates previous acoustic results obtained for a group of pre-school Italian children who stutter (Lenoci, 2015) and it is an aspect that need to be investigate further.

Bibliography

Ambrose, N.G., Yairi, E., Loucks, T.M., Seery, C.H. & Throneburg, R. (2015). Relation of motor, linguistic and temperament factors in epidemiologic subtypes of persistent and recovered stuttering: Initial findings. In *Journal of Fluency Disorders*, 45, 12-26.

Buhr, A.P., Zebrowski, P.M. (2009). Sentence position and syntactic complexity of stuttering in early childhood: A longitudinal study. In *Journal of Fluency Disorders*, 34, 155-172.

Chang, S.E., Ohde, R.N. & Conture, E.G. (2002). Coarticulation and formant transition rate in young children who stutter. In *Journal of Speech, Language, and Hearing Research*, 45, 676-688.

Cleland, J., Scobbie, J. & Zharkova, N. (2016). Insights from ultrasound: Enhancing our understanding of clinical phonetics. In *Clinical Linguistics & Phonetics*, 30, 171-175.

Dehqan, A., Yadegari, F., Blomgren, M. & Scherer, R.C. (2016). Formant transitions in the fluent speech of Farsi-speaking people who stutter. In *Journal of Fluency Disorders*, 48, 1-15.

Frisch, S.A., Maxfield, N. & Belmont, A. (2016). Anticipatory coarticulation and stability of speech in typically fluent speakers and people who stutter. In *Clinical Linguistics & Phonetics*, 30, 277-291.

Guthrie, E.R. (1952). *The psychology of learning.* New York: Harper and Row.

Hardcastle, B., Tjaden, K. (2008). Coarticulation and Speech Impairment. In Ball, M.J., Perkins, M.R., Muller, N. & Howard, S. (Eds.), *The Handbook of Clinical Linguistics*. Malden: Blackwell Publishing Ltd., 506-524.

Harrington, J. (1987). Coarticulation and stuttering: an acoustic and electropalatographic study. In Peters, H.F.M., Hulstijn, W. (Eds.), *Speech Motor Dynamics in Stuttering.* Wien: Springer Verlag, 381-392.

Hawkins, S. (1984). On the Development of Motor Control in Speech: Evidence from Studies of Temporal Coordination. In Lass, N.J. (Ed.), *Speech and Language: Advances in basic research and practice.* New York: Academic Press, 317-374.

Heyde, C.J., Scobbie, J.M., Lickley, R. & Drake, E.K.E. (2015). How fluent is the fluent speech of people who stutter? A new approach to measuring kinematics with ultrasound. In *Clinical Linguistics & Phonetics*, 30, 292-312.

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

Iskarous, I., Fowler, C. & Whalen, D. (2010). Locus equations are an acoustic expression of articulator synergy. In *Journal of the Acoustical Society of America*, 128, 2021-2032.

Kelso, J.A.S. (1997). *Dynamic patterns. The self-organization of brain and behavior.* Cambridge, MA: MIT Press.

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

Krull, D. (1988). Acoustic properties as predictors of perceptual responses: A study of Swedish voiced stops. In *Phonetic Experimental Research at the Institute of Linguistics*. Stockholm University: PERILUS X, 87-101.

Lashley, K.S. (1951). The problem of serial order in behaviour. In Jeffress, L.A. (Ed.), *Cerebral Mechanisms in Behavior.* New York: Wiley, 112-36.

Lenoci, G. (2015). Balbuzie evolutiva e indici fonetici predittivi di cronicità. Tesi di Dottorato, Scuola Normale Superiore, Pisa.

Lenoci, G., Allegri, S., Bernardini, S., Chiari, F., Crivelli, N., Dadamo, V., de Biase, M., Galatà, V., Pisciotta, C., Polesel, L., Stanchina, S., Stocco, D., Vayra, M. & Zmarich, C. (2012). Il progetto di ricerca longitudinale “indici fonetici predittivi di balbuzie cronica in età prescolare: primi risultati”. In Paoloni, A., Falcone M. (Eds.), *La voce nelle applicazioni*, Roma. Bulzoni Editore: Roma, 19-34.

Levelt, W.J.M., Roelofs, A. & Mejer, A.S. (1999). A theory of lexical access in speech production. In *Behavioral and Brain Sciences*, 22, 1-75.

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In Hardcastle, W., Marchal, A. (Eds.), *Speech production and speech modeling*. Amsterdam: Kluwer Academic Publishers.

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

MacNeilage, P.F., Davis, B.L. (2000). The Motor Core of Speech: A Comparison of Serial Organization Patterns in Infants and Languages. In *Child Development*, 71, 153-163.

MacPherson, M.K., Smith, A. (2013). Influences of Sentence Length and Syntactic Complexity on the Speech Motor Control of Children Who Stutter. In *Journal of Speech, Language, and Hearing Research*, 56, 89-102.

Namasivayam, A.K., Van Lieshout, P. (2011). Speech Motor Skill and Stuttering. In *Journal of Motor Behavior*, 43, 477-489.

Nearey, T., Shammass, S. (1987). Formant transitions as partly distinctive invariant properties in the identification of voiced stops. In *Canadian Acoustics*, 15, 17-24.

Noiray, A., Menard, L. & Iskarous, K. (2013). The development of motor synergies in children: Ultrasound and acoustic measurements. In *Journal of Acoustical Society of America*, 133, 444-452.

Olander, L., Smith, A. & Zelaznik, H.N. (2010). Evidence that a motor timing deficit is a factor in the development of stuttering. In *Journal of Speech, Language, and Hearing Research*, 53, 876-886.

Robb, M., Blomgren, M. (1997). Analysis of F2 transitions in the speech of stutterers and nonstutterers. In *Journal of Fluency Disorders*, 22, 1-16.

Smith, A., Goffman, L. (1998). Stability and patterning of speech movement sequences in children and adults. In *Journal of Speech, Language, and Hearing Science*, 41, 18-30.

Smith, A., Sadagopan, N., Walsh, B. & Weber-Fox, C. (2010). Increasing phonological complexity reveals heightened instability in inter-articulatory coordination in adults who stutter. In *Journal of Fluency Disorders*, 35, 1-18.

Smith, A., Zelaznik, H.N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. In *Developmental Psychobiology*, 45, 22-33.

Stromsta, C. (1965). A spectrographic study of disfluencies labeled as stuttering by parents. In *De Therapia Vocis et Loquellae*, 1, 317-320.

Subramanian, A., Yairi, E. & Amir, O. (2003). Second formant transitions in fluent speech of persistent and recovered preschool children who stutter. In *Journal of Communication Disorders*, 36, 59-75.

Sussman, H.M., Byrd, C.T. & Guitar, B. (2010). The integrity of anticipatory coarticulation in fluent and non-fluent tokens of adult who stutter. In *Clinical Linguistics & Phonetics*, 25, 169-186.

Sussman, H.M., Hoemeke, K.A. & Ahmed, F. (1993). A cross-linguistic investigation of locus equations as a relationally invariant descriptor for place of articulation. In *Journal of the Acoustical Society of America*, 94, 1256-1268.

Sussman, H.M., Hoemeke, K.A. & McCaffrey, H.A. (1992). Locus Equations as an index of coarticulation for place of Articulation Distinctions in Children. In *Journal of Speech and Hearing Research*, 35, 769-781.

Van Lieshout, P.H.H.M. (1995). *Motor planning and articulation in fluent speech of stutterers and nonstutterers*. Nijmegen: The Netherlands University Press Nijmegen.

Van Lieshout, P.H.H.M., Hulstijn, W. & Peters, H.F.M. (2004). Searching for the weak link in the speech production chain of people who stutter: A motor skill approach. In Maassen, B., Kent, R., Peters, H.F.M., Van Lieshout, P.H.H.M & Hulstijn, W. (Eds.) *Speech motor control in normal and disordered speech*. Oxford: Oxford University Press, 313-355.

Vanryckeghem, M., Brutten, G.J. (2007). *The KiddyCAT: communication attitude test for preschool and kindergarten children who stutter*. San Diego: Plural Publishing.

Weismer, G. (1991). Assessment of articulatory timing. In Cooper, J. (Ed.), *Assessment of speech and voice production: Research and clinical applications*. Bethesda, Maryland: National Institute on Deafness and other Communication Disorders.

Wingate, M.E. (1988). *The structure of stuttering: A psycholinguistic analysis*. New York: Springer Verlag.

Yairi, E., Ambrose, N. (2005). *Early Childhood Stuttering: for clinicians by clinicians*. Austin, TX: Pro-Ed.

Yairi, E., Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. In *Journal of Fluency Disorders*, 38, 66-87.

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

Zharkova, N., Hewlett, N. (2009). Measuring lingual coarticulation from midsagittal tongue contours: Description and example calculations using English /t/ and /a/. In *Journal of Phonetics*, 37, 248-256.

Zharkova, N., Hewlett, N. & Hardcastle, W.J. (2012). An ultrasound study of lingual coarticulation in /sV/ syllables produced by adults and typically developing children. In *Journal of the International Phonetic Association*, 42, 193-208.

Zmarich, C., Bernardini, S., Lenoci, G., Natarelli, G. & Pisciotta, C. (in press). Could the frequencies of stuttering-like-dysfluencies predict persistent stuttering in children who have just started to stutter? In Savy R., De Meo, A. & Alfano I. (Eds.), *La fonetica sperimentale nell'insegnamento e nell'apprendimento delle lingue straniere*.